

LONDON
SCHOOL *of*
HYGIENE
& TROPICAL
MEDICINE



**Analysing the balance between primary care providers and
hospitals in China**

Jin Xu

Thesis submitted in accordance with the requirements for the degree of

Doctor of Philosophy of the

University of London

January 2017

**Department of Global Health and Development
Faculty of Public Health and Policy
LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE**

Funded by China Scholarship Council

Declaration by candidate

I, Jin Xu, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

I have read and understood the School's definition of plagiarism and cheating given in the Research Degrees Handbook. I have read and understood the School's definition and policy on the use of third parties (either paid or unpaid) who have contributed to the preparation of this thesis by providing copyediting and, or, proof reading services. I declare that no changes to the intellectual content or substance of this thesis were made as a result of this advice, and, that I have fully acknowledged all such contributions.

I have exercised reasonable care to ensure that the work is original and does not to the best of my knowledge break any UK law or infringe any third party's copyright or other intellectual property right.

Signed:

A handwritten signature in black ink, appearing to read 'Jin Xu', is written over a faint, light-colored rectangular stamp or watermark.

Date: 6th January 2017

Full name: Jin Xu

Abstract

The primary aim of the thesis is to understand how to strengthen primary care in China, by looking at the effectiveness and functioning of a gatekeeping pilot, and more broadly at the history of the balance between hospitals and primary care providers. A secondary aim is to explore how multiple methods can be used to study a complex system-related issue. The thesis developed comprehensive metrics for structural and functional balance between primary and hospital care (1949-2013); developed a dynamic path dependence analytical framework to study the coevolution of the two sectors (1835-2013); used a difference-in-differences analysis to identify the impact of a gatekeeping pilot; and developed a qualitative systems analysis to understand the functioning of the pilot.

The quantitative analysis suggested more patients did visit primary care facilities (increased by 38.7%) due to the pilot, but without obvious extra-spending. Evidence from the qualitative study suggested this seemed to be caused by patients visiting for referrals. The intended effects of gatekeeping in changing patients' utilization pattern of care were made unattainable mainly by the existing weak conditions of primary care, feedback loops that further weakened primary care development regarding service capacity, human resources, and patient trust, as well as unintended consequences of other related policies. The dynamics between hospitals and primary care providers in the pilot were the contemporary manifestation of a long-term hospital-centric structure, where primary care providers were the de-professionalized antithesis of hospitals and relatively weakly institutionalized. The institutional complex was path-dependent and has gone through three cycles since 1835.

The thesis has demonstrated the feasibility and value of using multiple analytical theories and research methods to address a complex health system issue. The findings of the study suggest the importance of building a strong primary care profession and

the value of a political coalition for primary care strengthening.

Acknowledgements

I am greatly indebted to my supervisor, Prof. Anne Mills. Anne is not only a public health giant of our time, but also a superb teacher and a great human being. With the breadth and depth of her knowledge, and the boldness and pragmatism of her thinking, she has helped me navigate through the past four years, and this multi-methods doctoral study. Indeed, Anne is a highly committed and supportive supervisor. Talking with her provides intellectual stimulation and mental encouragement. In her very busy schedule, she always finds time to meet me regularly and discuss my study. Over the past year, she has also taken my thesis drafts all around the world in her travels, taking every opportunity to read and comment, not infrequently on trains and in flights, and sometimes even in her holidays. To have been a student of Anne (who happens to be also an amazing gardener, cook and host) is such a supreme fortune that Chinese would call it “the luck for three incarnations”.

My thankfulness also goes to my advisors, to Prof. Martin Gorsky, who has not only taught me important knowledge about the field of history of health system, but painstakingly help me use them in studying China; to Prof. Andy Haines, who has shared very important insights on primary care and the topic of the balance between primary and hospital care; to Dr. Tim Powell-Jackson for his unfailing support in impact evaluation; to Dr. Karl Blanchet, for his guidance regarding systems theories and methods; to other staff member at the School, including but not limited to Profs Kara Hanson and Pauline Allen for their friendship and beneficial comments on my thesis; to Prof. Stephen Peckham and Dr. Gerry Bloom for being my examiners, reading through my thesis and providing invaluable comments.

I am grateful to my country and my countrymen: to the China Scholarship Council, for providing the financial support for the four years; to Prof. Han Qide for endorsing my application for the PhD and the scholarship; to Prof. Meng Qingyue, for providing

continuous and enormous intelligent and logistical support; to Dr. Lei Haichao, for his crucial help regarding fieldwork; to my old colleagues, Profs Liu Xiaoyun and Fang Hai, Dr. Zhu Weiming and others for sharing with me their experiences and advice; to Dr. Su Jing for her advice regarding sources related to history of medicine in China; and to the health workers and patients who were kind enough to be interviewed by me during my fieldwork.

My thanks also goes to my friends in the UK: to my colleagues at the London School: Daniel Cauchi, David Lugo, Iftikhar Khan, Jingying Xu, Jorn Jacobson, Julie Evans, Lishi (Henry) Li, Noppakun Thammatacharee, Yoko Lawrence, and many others; and to my fellows in the London Research Group on China Health. Their friendship has been helping me through both sunshine and rain in study and in life.

My deepest gratitude also goes to my family: to my grand farther and parents, for their continuous encouragement for me and financial assistance in caregiving for my son; to my wife, for holding up more than “half the sky”, making sacrifice during the first three years of my PhD for our family, and helping with editing and proofreading during the final year of my study; to my son, for rewarding me with so much joy for the very limited fatherhood I have provided--usually over WeChat and eight time zones away; to my mother-in-law, for giving amazing parenting for my son during my study, enduring pain with him when he was ill and raising up his glorious boyhood. This work is not possible without their support and sacrifice. It is through this pursuit of study far away from home that I have learnt the real strength of my family.

Contents

Declaration by candidate.....	2
Abstract.....	3
Acknowledgements.....	5
Contents.....	7
List of tables.....	12
List of figures.....	14
List of abbreviations.....	16
Part I Conceptualization.....	17
Chapter 1 Introduction.....	17
1.1 Importance of balance between hospitals and primary care facilities.....	17
1.2 Moving towards stratified health service delivery in China.....	18
1.3 The approach of the thesis: from gatekeeping to history.....	20
1.4 Organisation of thesis.....	24
Chapter 2 Literature review.....	26
2.1 Introduction.....	26
2.2 Review process and search results.....	26
2.3 Conceptual review.....	31
2.3.1 The concepts related to the balance between primary care and hospital care.....	32
2.3.2 The concept of primary care.....	35
2.3.3 Commentary.....	42
2.4 Measurement of the balance between hospitals and primary care providers.....	43
2.4.1 Structural aspects.....	44
2.4.2 Functional aspects.....	44
2.4.3 Commentary.....	46
2.5 Findings regarding strategies to shift the balance between primary and hospital care.....	47
2.5.1 Effectiveness of gatekeeping in strengthening primary care functioning.....	47
2.5.2 Issues of complexity in shifting the balance between hospitals and primary care providers.....	50
2.5.3 Commentary.....	53
2.6 Review of approaches used in thesis.....	55
2.6.1 Impact evaluation.....	56
2.6.2 Systems theory and causal loop analysis.....	60
2.6.3 Historical institutionalism and path dependence analysis.....	63
2.7 Conclusions.....	74
Chapter 3 Aims, objectives and overview of methods.....	75
3.1 Aims and objectives.....	75
3.2 Overview of methods.....	76

3.2.1	The overall study design	76
3.2.2	Describing historical trends of the balance between primary and hospital care ...	78
3.2.3	Analysing the historical logic of evolution of the balance between hospitals and primary care providers	78
3.2.4	Impact evaluation of the gatekeeping pilot	79
3.2.5	Systems analysis of the functioning of the gatekeeping pilot	79
3.2.6	Ethics.....	80
3.3	Summary	80
Part II	History	81
Chapter 4	Historical trends (1949-2013)	81
4.1	Introduction.....	81
4.2	Background about the health delivery system in China.....	81
4.2.1	Hospitals	85
4.2.2	Primary care providers.....	86
4.3	Methods.....	88
4.3.1	Overall rationale for developing the metrics.....	88
4.3.2	Data	91
4.3.3	Justification for indicators and their operational definitions.....	93
4.4	Results.....	100
4.4.1	Human resources.....	100
4.4.2	Assets, equipment and infrastructure	110
4.4.3	Service utilisation.....	115
4.4.4	Financing.....	120
4.5	Discussion.....	121
4.5.1	Strengths and limitations.....	122
4.5.2	Discussion about findings	124
4.6	Conclusion	128
Chapter 5	Historical roots of hospital dominance (1835-1949).....	130
5.1	Introduction.....	130
5.2	Methods.....	131
5.2.1	Analytical framework	131
5.2.2	Protocol	132
5.2.3	Unit of analysis: institutional complex behind the coevolution of primary care and hospital sectors.....	133
5.2.4	Chronological scope and periodization.....	135
5.2.5	Issues about collection and critical reading of texts.....	136
5.2.6	Sources	139
5.3	Overall historical context (1835-1949)	141
5.4	Organisation of health services in traditional China (1572-1835).....	143

5.5 Introduction of Western medicine (1835-1928).....	148
5.5.1 Introduction of mission hospitals (1835-1844).....	149
5.5.2 Further development of mission hospitals (1844-1901).....	151
5.5.3 Consolidation of hospital-based professionalization (1901-1928).....	156
5.6 Two tracks of health services development (1928-1949).....	162
5.6.1 The emerging agenda of state medicine (1928-1934).....	163
5.6.2 Splitting into two tracks (1935-1949).....	168
5.7 Summary.....	177
Chapter 6 Path dependence of hospital-centrism (1949-2013).....	178
6.1 Introduction.....	178
6.2 Overall historical context (1949-2013).....	178
6.3 Development of the two-sector system of service delivery (1949-1978).....	180
6.3.1 Launching exclusively urban welfare (1949-1953).....	181
6.3.2 Self-reinforcing development of hospital dominance (1953-1965).....	183
6.3.3 Turning towards rural primary care (1965-1968).....	190
6.3.4 Resistance in implementation and development in rhetoric (1968-1978).....	192
6.4 Adaption and path dependence during reforms (1978-2013).....	197
6.4.1 Changes favouring hospitals and technology (1978-1985).....	198
6.4.2 Evolution of the two-sector structure (1985-2002).....	201
6.4.3 Turning towards universal health coverage (2002-2009).....	205
6.4.4 Implementing health system reform (2009-2013).....	207
6.5 Discussion.....	210
6.5.1 Weaknesses and strengths.....	210
6.5.2 Coevolution of hospitals and primary care providers.....	211
6.6 Conclusion.....	218
Part III A contemporary gatekeeping pilot.....	220
Chapter 7 Impact evaluation of the gatekeeping pilot.....	220
7.1 Introduction.....	220
7.2 Methods.....	221
7.2.1 The gatekeeping pilot.....	221
7.2.2 Difference-in-differences study design.....	226
7.2.3 Data.....	227
7.2.4 Econometric analysis.....	229
7.3 Results.....	230
7.4 Discussion.....	236
7.4.1 Strengths and limitations.....	236
7.4.2 Findings.....	239
7.5 Conclusion.....	240
Chapter 8 A qualitative systems analysis of the gatekeeping pilot.....	242

8.1 Introduction.....	242
8.2 Methods.....	243
8.3 Results.....	248
8.3.1 Intended dynamics	251
8.3.2 Low incentive of performance-based salary	253
8.3.3 Vicious cycle of primary care	254
8.3.4 Challenges from hospitals.....	260
8.3.5 Gatekeeping backfires.....	263
8.4 Discussion.....	265
8.4.1 Limitations and value of the approach.....	265
8.4.2 Findings regarding gatekeeping and implications beyond.....	267
8.5 Conclusion	270
Part IV Discussion and conclusions.....	272
Chapter 9 Discussion	272
9.1 Methodological issues regarding synthesizing findings from sub-studies.....	272
9.1.1 Conceptualization of balance between hospitals and primary care providers.....	273
9.1.2 Synthesis within the historical part	274
9.1.3 Synthesis within the contemporary part.....	275
9.1.4 Synthesis across the two parts.....	275
9.1.5 Summary	279
9.2 Discussion on results.....	279
9.2.1 Situation and evolution of the balance between hospitals and primary care providers	280
9.2.2 Can the gatekeeping pilot shift the balance towards primary care.....	285
9.2.3 Diagnosis, prognosis and solutions	287
9.3 Relevance and generalizability of findings of sub-studies.....	300
9.3.1 Trend analysis of structural and functional balance of care.....	300
9.3.2 Historical-institutionalist analysis of coevolution of hospitals and primary care providers	301
9.3.3 Qualitative systems analysis and impact evaluation of the gatekeeping pilot	302
9.4 Methodological reflection: a cross-disciplinary mixed-methods study approach in health policy and systems research	303
9.4.1 Overall study design.....	303
9.4.2 Sub-study 1: historical trend analysis of the balance between hospitals and primary care providers	305
9.4.3 Sub-study 2: Historical-institutionalist analysis of the coevolution of hospitals and primary care providers	306
9.4.4 Sub-study 3: Impact evaluation.....	307
9.4.5 Sub-study 4: Qualitative systems analysis of gatekeeping pilot functioning.....	307

Chapter 10 Conclusions and recommendations	309
10.1 Critical reflection of overall methods	309
10.2 Summary of findings.....	310
10.2.1 Findings of the historical trend analysis.....	310
10.2.2 Findings of the historical-institutionalist analysis.....	311
10.2.3 Findings of the impact evaluation of the contemporary gatekeeping pilot	312
10.2.4 Findings of the qualitative systems analysis of the gatekeeping pilot	312
10.2.5 Findings regarding combined research approaches	313
10.2.6 Overall interpretation of results	313
10.3 Policy implications.....	314
10.3.1 Policy implications for China	314
10.3.2 Policy implications for other low- and middle-income countries	319
10.4 Areas for further research	321
10.4.1 Metrics and trend analysis of balance between hospitals and primary care providers	321
10.4.2 History.....	322
10.4.3 Strategies to shift balance towards primary care.....	322
10.4.4 Qualitative systems analysis	323
10.4.5 Impact evaluation.....	323
10.4.6 Using history and systems analysis for health policy and systems research.....	323
10.4.7 Research agenda for health policy and systems research.....	324
10.5 Conclusions.....	324
References.....	327
Appendices.....	360

List of tables

Table 2-1. Electronic databases and search engines.....	27
Table 2-2. Search strategy for literature on Ovid.....	28
Table 2-3. Search strategy for Chinese literature on CNKI and Wanfang.....	30
Table 2-4. Representative definitions of primary care and primary health care.....	39
Table 2-5. Complexity issues and methodological challenges.....	54
Table 2-6. Justification for selection of research approaches.....	56
Table 2-7. A multi-step analytical method for critical juncture analysis.....	70
Table 3-1. Objectives and research approaches.....	75
Table 4-1. Indicators of structural and functional balance between primary and hospital care.....	93
Table 4-2. Comparison across indicators over time.....	125
Table 5-1. Analytical steps for path dependence analysis.....	133
Table 5-2. Key issues of documentary analysis.....	136
Table 5-3. Comparison of the three types of medical missionary work.....	152
Table 5-4. Comparison between results of mission hospital surveys 1919-1933/1934.....	167
Table 5-5. Comparison of distribution of hospital resources and county health centres	174
Table 6-1. Policies on ownership of primary care facilities.....	188
Table 6-2. Historical logic of coevolution of hospitals and primary care providers in China (1835-2013).....	214
Table 7-1. Outcome variables of enrolees in pilot and non-pilot townships before and after the start of the gatekeeping pilot.....	233
Table 7-2. Regression results on visits.....	234
Table 7-3. Regression results on expenditures.....	235
Table 8-1. Overall process of the qualitative systems analysis.....	244
Table 8-2. List of interviewees and their characteristics.....	246

Table 8-3. Symbolic representation in the causal loop diagram..... 249

List of figures

Figure 2-1. Flowchart of search process of international literature.....	29
Figure 2-2. Flowchart of search process of Chinese literature.....	31
Figure 2-3. Main types of path dependence analysis identified in the literature review	70
Figure 3-1. Summary of overall study design	77
Figure 4-1. Main types of health service providers in China in 2013	83
Figure 4-2. Number of hospitals and primary care facilities in China (1949-2013)	85
Figure 4-3. Number of staff at hospitals and primary care facilities (1952-2013).....	103
Figure 4-4. Percentage of professional doctors in all practitioners in urban and rural health facilities (1952-1988)	104
Figure 4-5. Percentage of professional doctors in all doctors in rural primary care facilities and county hospitals (selected years between 1952 and 1988)	106
Figure 4-6. Percentage of fully licenced doctors in all doctors at various types of health facilities (2002-2013)	107
Figure 4-7. Percentage of health professionals with a university degree in selected years between 2002 and 2013.....	108
Figure 4-8. Average annual income of staff at various health facilities (2004-2013) .	110
Figure 4-9. Value of assets of selected facilities (1 billion <i>yuan</i>) (2004-2013).....	111
Figure 4-10. Number of beds in hospitals and township health centres (1949-2013).	113
Figure 4-11. Number of valuable devices at hospitals and primary care facilities (2003- 2013)	114
Figure 4-12. Building spaces of hospitals and primary care facilities (1988 and 2003- 2013)	115
Figure 4-13. Number of patient visits to hospitals and primary care facilities (1981-2013)	117
Figure 4-14. Number of admissions at hospitals and township health centres (1981-2013)	118

Figure 4-15. Visits to admissions ratio in hospitals and primary care facilities (1981-2013)	120
Figure 4-16. Total revenue of hospitals and primary care facilities (billion <i>yuan</i>) (2004- 2013)	121
Figure 5-1. A model of dynamic path dependence	132
Figure 5-2. Location of institutions receiving aid from the China Medical Board	161
Figure 6-1. Three paintings featuring rural health workers in the 1960s and the 1970s	197
Figure 7-1. Schematic representation of study design.....	227
Figure 8-1. Causal loop diagram for gatekeeping pilot.....	250
Figure 8-2. R1', an intended reinforcing feedback loop regarding the effects of gatekeeping for balance of care.....	252
Figure 8-3. B1', an intended balancing feedback loop concerning referral balance ...	253
Figure 8-4. R1a, a reinforcing feedback loop concerning a vicious cycle of primary care capacity.....	256
Figure 8-5. R1b, a reinforcing feedback loop concerning the sustainability of primary care human resources	258
Figure 8-6. R1c, a reinforcing feedback loop concerning primary care losing patients' faith	259
Figure 8-7. B2, a balancing loop concerning brain drain from primary care.....	261
Figure 8-8. R2, a reinforcing feedback loop, concerning syphoning of HR, patients and resources.....	262
Figure 8-9. R3, a reinforcing feedback loop concerning hospital bargaining power ..	263
Figure 8-10. B1, a balancing feedback loop concerning resistance to the gatekeeping policy.....	264

List of abbreviations

CMS	Cooperative Medical Scheme
CLD	causal loop diagram
CNKI	China National Knowledge Infrastructure
CPC	Communist Party of China
GIS	Government Employees Insurance Scheme
GP	general practitioner
INUS	Insufficient but Necessary parts of a condition which is itself Unnecessary but Sufficient
LIS	Labour Insurance Scheme
LNHO	League of Nations Health Organization
NCMS	New Rural Cooperative Medical Scheme
NCDs	non-communicable diseases
PHC	primary health care
PUMC	Peking Union Medical College
WHO	World Health Organization

Part I Conceptualization

Chapter 1 Introduction

1.1 Importance of balance between hospitals and primary care facilities

A world-wide momentum has been growing to push for progress towards universal health coverage, enshrined in the United Nations' 2030 agenda for sustainable development (United Nations, 2015). With increasing financial resources being committed, what is needed “now more than ever” are health systems that focus on primary care--“person-centredness, comprehensiveness and integration, and continuity of care, with a regular point of entry into the health system” (World Health Organization, 2008, p. 42). A primary-care orientation has been associated with important contributions to health systems performance ranking on health indicators, efficiency and population satisfaction with health services in rich countries (Starfield, 1994). Ecological studies have showed that stronger primary care is associated with better population health, not only in high income countries (Macinko, Starfield, & Shi, 2003), but also in low- and middle-income countries (Macinko, Starfield, & Erinosh, 2009). Primary care appears to mitigate the negative health effects of income inequality (Shi et al., 2002) and contributes to quality of care (Zahradnik, 2008). Primary care has also been demonstrated as a sound platform for cost-effective services that could reduce the need for expensive specialist care (Walker et al., 2010).

The country reality, however, usually runs counter to the vision of a primary-care-oriented health system, with a widely existing phenomenon known as “hospital-centrism”--“disproportionate focus on hospitals and sub-specialization has become a major source of inefficiency and inequality, and one that has proved remarkably resilient” (World Health Organization, 2008, p. 11). Compared to primary care providers, hospitals are more likely to enjoy advantageous resources (Shi, 2012):

specialists that usually are more prestigious and better paid than generalists, advanced medical equipment, advanced and expensive pharmaceuticals, etc. In low- and middle-income countries, hospitals and their professionals tend to have evolved into powerful entities that challenge reforms (Mutizwa-Mangiza, 1999; World Health Organisation, 2008). Disproportionate increase in the hospitals' share of resources means reduced resources available for primary care strengthening (Mills, 1990). As the World Health Organization (2014, p. 6) pointed out, hospital spending accounts for half to two thirds of total government expenditures on health. World Bank Group President Jim Kim (2013) also criticised the “epidemic of hospital-building”: spending primarily on building hospitals while investment should be made in upstream prevention and disease management to prevent those complications that need hospitalization. In low- and middle-income countries, strengthening the primary care level is believed to be “the only efficient and effective way” (Chan, 2011) to achieve early detection, treatment and management of the growing burden of chronic diseases.

1.2 Moving towards stratified health service delivery in China

Although the issue of the distribution of resources between hospitals and primary care is of critical importance to most low- and middle-income countries, the geographical focus of the thesis is China. The reasons for studying this issue in the context of China are manifold. First, the balance between primary and hospital care is a central policy issue in contemporary China. Recently, “primary care level as the emphasis (基层为重点)” was not only included but also placed as the first among the government's six guiding principles for health work. These principles were applied in the government's middle and long term plan for health—“Healthy China 2020” (Xinhua, 2016b). In a top level national health congress, the last of which was held two decades ago in 1996, policy makers placed the establishment of “*fenji zhenliao*” (分级诊疗), roughly translated as “stratified health service delivery”, as a top priority for health system reform and development (Xinhua, 2016a).

Second, the role of primary health care in China's health system development has been of interest to the international community. A key element in primary health care movements in China has been about the organisation of primary care providers to extend medical and public health services, yet relatively little is known about China's history regarding the balance between hospitals and primary care providers (as the thesis will later address, primary care is not identical to primary health care). Back in the 1930s, the Ding County health system (involving a county health centre, district health stations and village health personnel) was regarded as one of the origins of the primary health care concept (Litsios, 2015). In the 1970s, development of primary health care in China involving barefoot doctors was reported in *Health by the People* (Newell, 1975), laying the empirical foreground of the Primary Health Care Movement launched in Alma Ata in 1978. Afterwards, China also featured as a success model in *Good Health at Low Cost* (Halstead, Walsh, & Warren, 1985). By the 1990s and the early 2000s, the country was heavily criticized for its neglect of primary health care under market-oriented reforms. In recent years, the existence of a network of primary health care providers in rural China has contributed to its successful development of the New Rural Cooperative Medical Scheme (Meng & Xu, 2014a). With expanded health insurance coverage in recent years (Meng et al., 2015), the country's renewed effort to increase the primary care might offer lessons that would be of interest to international audiences.

Third, the size and the global role of China matter. The country was home to about one fifth of the world's population (1.4 billion out of 7.3 billion) in 2015, and it had more people than the United States of America (321 million), the European Union (508 million), Russia (144 million), Japan (127 million), and Brazil (257 million) combined (World Bank, 2016b). China is also the world's second largest economy (World Bank, 2016a) and the world's leading exporter (World Trade Organization, 2015, p. 25). Just as improved health of Chinese has direct impacts on global health,

increasing efficiency of health spending in China would have consequences for the global economy.

Fourth, China shares a strong element of colonial/semi-colonial history with other non-Western countries. Introduction of Western medicine in such countries was associated with foreign powers and ideas, and usually through foreign actors (e.g. missionaries), who preferred to build hospitals (Hardiman, 2006). China is an example with such early focus on hospital construction, as I will address in Chapter 5. China also has a shared history with post-communist countries in terms of various political and economic elements. Many of these countries were considered to have a history of relatively poor primary care providers, as priority was given to specialists and hospitals (Saltman, Rico, & Boerma, 2005). These countries have also gone through similar economic reforms, which profoundly influenced development of primary health care. China therefore offers an opportunity to consider how such history influenced the relative development between hospitals and primary care facilities, as different styles of medicine were forced to co-exist and/or compete with each other.

1.3 The approach of the thesis: from gatekeeping to history

The central question of this study concerned the challenges of how to strengthen primary care in China. The study started with an interest in the value of measures to retain patients at primary care level, especially via gatekeeping. There has been a nearly dominant narrative in the country that patients' irrational use of resources has led to a high concentration of patients at large and sophisticated hospitals, even though a large percentage of patients have relatively simple and mild conditions that could have been handled at primary care facilities (Lei et al., 1996; Song & Yin, 2010; Xu, Yin, & Jing, 2012). There have also been claims in Chinese and international academic papers that gatekeeping would be an effective solution to the issue of underuse of primary care in China (Chen & Zhou, 2007).

Evidence on the effectiveness of gatekeeping in China has been lacking, which will be apparent in the review in Chapter 2. As the ideas for the PhD study developed, a pioneering gatekeeping pilot was launched and became available for study. The thesis took the first step to study the effectiveness of this gatekeeping pilot programme using an impact evaluation, which identifies the causal effects of specific interventions (Gertler et al., 2011, pp. 7-8). For this purpose, this component of the thesis adopted the technique of difference-in-differences analysis, which has been found to facilitate studying net treatment effects by constructing counterfactuals when randomization is not feasible (Gertler et al., 2011: Chapter 6), which is usually the case in health systems research. Sutton et al. (2012) and Kristensen et al. (2014), for example, used the method to examine whether a pay-for-performance scheme reduced in-hospital mortality in England.

The literature review in Chapter 2 on gatekeeping and other options to balance the relative growth of hospitals and primary care providers, as well as general development in methods of evaluating complex interventions, suggested multiple dimensions of complexity associated with strengthening primary care. Some of these issues were “within” the programmes of particular interventions like gatekeeping, others were related to a wider range of “contextual” factors, some of which can be historically shaped and restrict the development of particular institutions. Systems theories and system thinking have been advocated as useful for health systems and policy research (Adam & de Savigny, 2012; Bloom, 2014; Gilson, 2012). In order to address what became clear as the important issues of complexity, the thesis borrowed from recent development of different approaches to systems thinking in health systems research, and expanded the study to analyses of both the contemporary system, and of the historical system that the pilot was embedded within.

To better understand the functioning of the gatekeeping pilot programme in the contemporary system where the gatekeeping pilot was embedded, the thesis drew on

a qualitative systems analysis tool (causal loop diagrams), a health systems building blocks categorization tool, and a qualitative data analysis tool. Combining qualitative methods with systems thinking has been shown to add depth to analysis of health systems issues, and visualization can help convey complex interpretations and findings (Adam, 2014). Causal loop analysis is a method for systems analysis that maps out and qualitatively models the dynamics between factors within a system (Williams & Hummelbrunner, 2010a: Chapter 1). It has been applied and found valuable in health systems research topics including studying an immunization system (Rwashana, Williams, & Neema, 2009), analysing factors contributing to stagnant neonatal mortality rates (Rwashana et al., 2014), studying medical dual practice (Paina et al., 2014), and identifying systems determinants of integrated Community Case Management of malaria (Sarriot et al., 2015).

To understand the historical logic of the evolution of the balance between primary and hospital care, the study took two steps. The first step was to describe the historical trends of the balance between hospitals and primary care providers. Metrics have been proposed for assessing health systems features in relation to the concept of the balance between primary and hospital care. White, Williams, and Greenberg (1961) studied the utilisation of different levels of care. Mills (1990) analysed the distribution of health expenditures by the level of care in a cross-national analysis. Starfield (1994) developed a comprehensive system of metrics of primary care orientation of health systems. More recent development includes the Primary Health Care Performance Initiative (2015) measuring percentage of government spending on primary health care. This component of the study followed these measurement tools previously developed and extended them further.

The second step of the historical analysis was to explain how the changes took place. For this purpose, the thesis conducted a historical-institutionalist analysis of changes of institutional factors related to the relative balance between primary care

and hospital care. The historical approach has been advocated as useful for studying the institutional factors that underlie the evolution of health systems (Bloom, 2014; Peters, 2014) (including the changing the balance between primary and hospital care). For example, a framework of path dependence analysis within the political school of historical institutionalism has been shown useful for institutional development (Pierson, 2004) and health policy analysis (Hacker, 1998). A recent series of historical case studies on health system development used the framework of path dependence to facilitate an investigation on factors contributing to successful health system development (Balabanova, McKee, & Mills, 2011).

As the scope of the thesis expanded to tackle an increasingly broad range of factors affecting the balance between hospital and primary care, the framing broadened from the initial focus on gatekeeping as a means to strengthen primary care, to a dynamic contemporary system containing interrelated policies and elements, and a historical system containing structures rooted in an institutional legacy that evolves over time. Correspondingly, the thesis employed four approaches, which constituted a set of analytical tools for a multi-level analysis of the issue of balance between hospitals and primary care providers. Trend analysis and historical-institutionalist analysis provided an empirical and theoretical basis for making sense of the development of hospitals and primary care providers over time and the forces behind such development. The historical-institutionalist analysis also encompassed multiple policies and institutions that closely interacted with one another, which would be much less meaningful if restricted to factors directly related to gatekeeping. The contemporary impact evaluation provided needed evidence on the effectiveness of a particular strategy. Combined in a case study, impact evaluation and qualitative systems analysis generated rich understanding of the barriers facing each intervention. As the thesis later argues, such a mix of multiple approaches has particular strengths in helping deepen our understanding of complex issues like the balance between hospitals and

primary care providers in the context of low- and middle-income countries.

1.4 Organisation of thesis

The remaining chapters of Part I further lay out existing knowledge and explain the conceptual basis of the thesis. Chapter 2 reviews literature related to the concepts and measurements relevant to the relative development between hospitals and primary care facilities, the evidence regarding gatekeeping in retaining patients at primary care facilities both internationally and in China, as well as complexity issues related to studying gatekeeping and other similarly-intended interventions. The chapter also reviews the relevant useful research approaches to study such interventions. The chapter concludes with a summary of gaps in knowledge. Chapter 3 explains the aim and objectives of the thesis, and provides an overview of methods and sub-studies corresponding to each objective.

Although sub-studies were conducted in parallel, the results chapters are arranged in temporal sequence starting with quantitative results before qualitative analysis. Part II provides a historical account of the relative development of hospital and primary care facilities in China. Chapter 4 provides a description of the overall trends of the relative development of the hospital sector and the primary care sector in China between 1949 and 2013, in terms of human resources, assets, equipment and infrastructure, service utilization, and financing. Chapter 5 lays out the analytical approach of a path dependence framework and then uses it to analyse the emergence of the two sectors in China and their division in the context of the introduction of biomedicine from the West before the establishment of the People's Republic of China in 1949. Chapter 6 studies how the two sectors, including their institutional embedding, evolved from 1949 onward, when the country periodically underwent drastic political changes.

Part III examines a contemporary gatekeeping pilot in China. Chapter 7 provides

an impact evaluation of the pilot programme on visits and expenditures at different levels of care based on difference-in-differences analysis using claims data. Chapter 8 uses an approach of qualitative systems analysis based on causal loop diagrams, “Framework” analysis and the WHO categorization of health system building blocks, and uses the approach to analyse the systems factors involved in the functioning of the gatekeeping pilot.

Part IV consists of a discussion chapter and a concluding chapter. Chapter 9 discusses methodological advantages and limitations of the approach of the thesis, overall results of sub-studies (Chapters 4-8), and relevance and generalisability of thesis findings. Chapter 10 concludes the thesis by summarizing the key findings and contribution to knowledge, and suggesting implications for policy-making and further research priorities.

Within each chapter, material under a primary subheading (e.g. 2.1) is referred to as a section, while material under a secondary subheading (e.g. 2.1.1) is referred to as a sub-section.

Chapter 2 Literature review

2.1 Introduction

The first purpose of the chapter is to review literature with a focus on knowledge and previous studies about the subject matter. Specifically, the review inquired about: the two closely related concepts--primary care and primary health care, and the definitions relevant to the balance between primary care and hospital care; the ways in which the balance between primary care and hospital care has been measured; the evidence regarding effectiveness of strategies to shift the balance between primary care and hospital care and particularly regarding gatekeeping (definition provided later); and the findings of previous Chinese studies regarding effectiveness of strategies to shift the balance between primary care and hospital care. The first stage of review was done through a systematic approach.

The second purpose of the chapter is to identify useful approaches to study strategies to shift the balance between primary care and hospital care from the field of health policy and systems research. The review identifies three particularly promising approaches, namely, impact evaluation (difference-in-differences analysis), qualitative systems analysis (causal loop analysis), and historical institutionalism (path dependence analysis). The second stage of the review was conducted using an explorative approach. Finally, this chapter concludes with a summary of findings of the review and gaps in knowledge.

2.2 Review process and search results

A mix of automated and manual searches was used for reviewing knowledge about the subject matter, conducted in three parts. The first part of the search focused on the concept of the balance between primary care and hospital care. A number of electronic databases were used in the search via the platform of Ovid, based on suggestion from

the library of the London School of Hygiene & Tropical Medicine on relevant databases (see Table 2-1). A search was run for the period prior to September 2016. A list of all searched electronic databases and search engines are shown in Table 2-1. Search terms included key words for “balance of care” or for terms related to “hospital”, “primary care” and “balance” (see Table 2-2). A site search via google was conducted for the website of the World Health Organisation. A manual search was conducted to include key book titles on primary care.

Publications that discussed the definition and/or the measurement of the balance between primary care and hospital care (including the related term “balance of care”), and/or the strategies to shift the balance between primary care and hospital care were included. Excluded were studies that 1) only provided viewpoints without description or evidence, 2) only dealt with either primary care sector or hospital sector, without addressing the balance between primary care and hospital care, or 3) were methodologically inappropriate (e.g. cross-sectional studies) or unspecified.

Table 2-1. Electronic databases and search engines

Databases searched on Ovid*
✓ Global Health <1910 to 2016 Week 34>
✓ HMIC Health Management Information Consortium <1979 to July 2016>
✓ LSHTM Journals@Ovid
✓ Northern Light Life Sciences Conference Abstracts <2010 - 2016 Week 34>
✓ Econlit <1886 to August 2016>
✓ Embase Classic+Embase <1947 to 2016 September 02>
✓ Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) <1946 to Present>
✓ Social Policy and Practice <201607>
Website searched on Google
✓ World Health Organization (http://www.who.int)
Databases searched on China National Knowledge Infrastructure (CNKI)
✓ China Academic Journals Full-text Database (1994 onwards)
✓ China Doctoral Dissertations Full-text Database (1984 onwards)
✓ China Masters' Theses Full-text Database (1984 onwards)
✓ China Proceedings of Conference Full-text Database (1953 onwards)
Databases searched on Wanfang Data

- ✓ China Science Periodical Database (1998 onwards)
- ✓ China Dissertation Database (1980 onwards)
- ✓ China Conference Paper Database (1983 onwards)

Note: *access to Ovid was provided by London School of Hygiene & Tropical Medicine.

Table 2-2. Search strategy for literature on Ovid

No.	Search terms
1	balance of care.mp.
2	balance.mp.
3	hospital*.mp. or hospital/
4	community care.mp.
5	community health care.mp. or community care/
6	community health service*.mp.
7	health centre.mp. or health center/
8	primary care.mp. or primary medical care/
9	primary health care.mp. or primary health care/
10	primary healthcare.mp. or primary health care/
11	general practitioner/ or general practice/ or general practi*.mp.
12	4 or 5 or 6 or 7 or 8 or 9 or 10 or 11
13	2 and 3 and 12
14	1 or 13
15	gatekeep*.mp.
16	14 and 15
17	remove duplicates from 14 ^a
18	remove duplicates from 16
19	limit 18 to yr="2010-" ^b

Note: a) search for balance of care; b) search for gatekeeping.

The automated search identified 1872 results on Ovid (after automatic de-duplication) (see Figure 2-1). After further excluding duplicates ($n=17$), the title and/or abstracts (if needed) of the resulting articles were read through to exclude 1744 irrelevant articles. The full text of the remaining 111 articles were retrieved and scanned through, which reduced the number of included articles to 5. The search of the website of the World Health Organisation generated 455 results, which were reduced to 28 after a screening through the search results based on title/abstract. Further screening of the full text of these reduced the number of articles to 12.

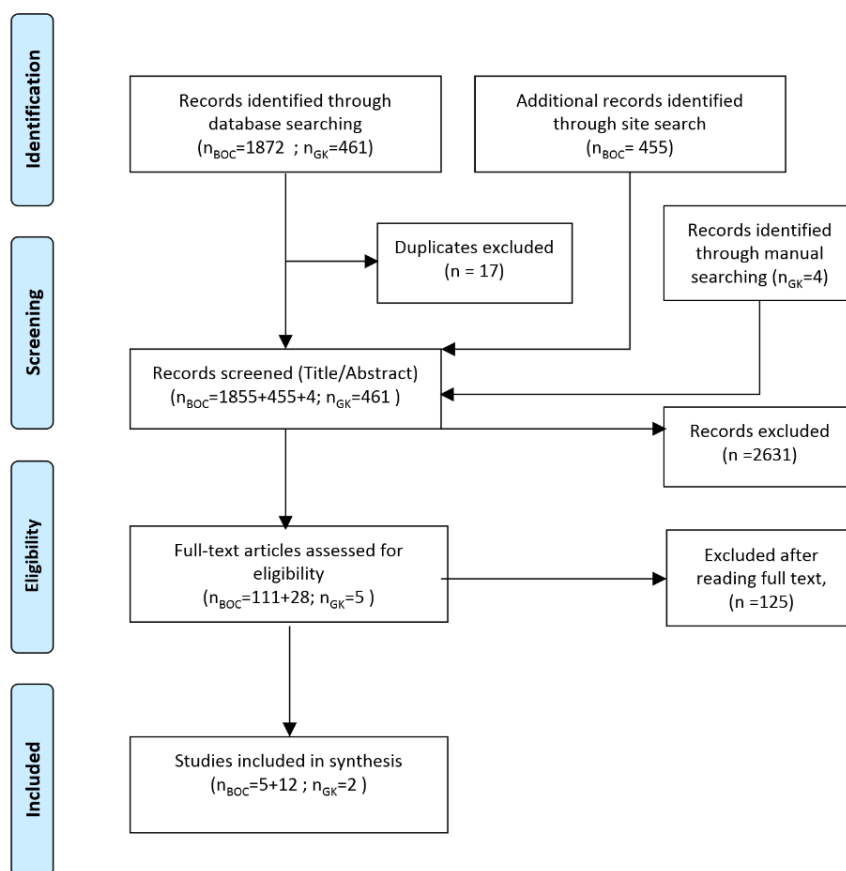


Figure 2-1. Flowchart of search process of international literature

The second part of the search focused specifically on gatekeeping (a concept that is elaborated later and became central as an opportunity to study a pilot emerged). A similar search process was used with different search terms related to both sectors (hospitals and primary care providers) and gatekeeping (see Table 2-2). As there has been a systematic review for gatekeeping which is considered to be of good quality (Centre for Reviews and Dissemination), the review here covered the articles included in the systematic review and conducted an additional update of the literature since 2010. The search identified 461 results, which were reduced to 5 after screening titles and abstracts, and further reduced to 2 (including a systematic review, the studies reviewed by the systematic review were also read).

The third part of the search focused on studies in Chinese on the balance between primary and hospital care including gatekeeping (see Table 2-3 for search strategy) in two main Chinese academic publication databases, namely, China National Knowledge Infrastructure (CNKI) and Wanfang Data. The search identified a total of 612 results in the two databases (218 in CNKI, 394 in Wanfang Data, see Figure 2-2). After excluding duplicates, 466 articles were screened by reading through the title and/or abstracts, which excluded 459 irrelevant articles. The full text of the remaining 7 articles were retrieved and read, after which 1 was included in the synthesis.

Table 2-3. Search strategy for Chinese literature on CNKI and Wanfang

No.	Subject
1	基层
2	卫生院
3	社区
4	村卫生室
5	村医
6	家庭医生
7	全科医生
8	1 or 2 or 3 or 4 or 5 or 6 or 7
9	首诊
10	守门人
11	患者分流
12	病人分流
13	分级诊疗
14	9 or 10 or 11 or 12 or 13
15	效果
16	8 and 14 and 15

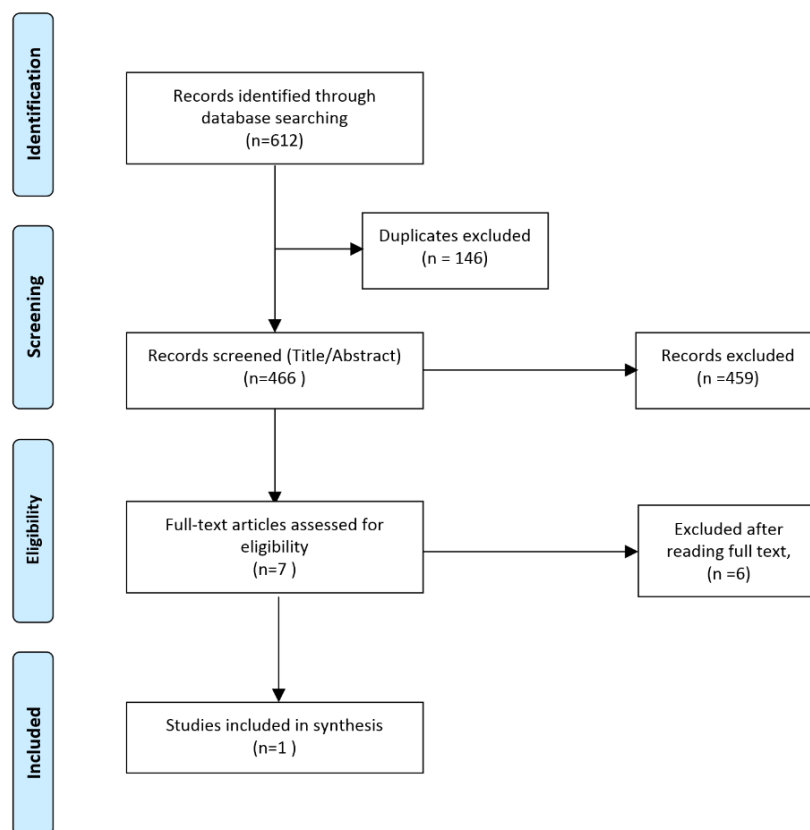


Figure 2-2. Flowchart of search process of Chinese literature

Review regarding theories and methods was conducted using an explorative approach recognized as important for the nascent field of health policy and systems research (McPake & Hanson, 2016; Sheikh et al., 2011) and involved two stages. In the first stage, I searched and read the key texts written on methodology related to health policy and systems research. The search led to three potentially useful approaches, namely, impact evaluation, systems analysis and historical institutionalism. The second stage of review then focused specifically on the theories and methods in relation to these three approaches.

2.3 Conceptual review

The section starts with articles using the specific term “balance of care” and others using terms related to the idea, then extends to primary care. The section concludes

with a summary of the findings from the conceptual review and clarification of the use of the concept in this thesis, particularly for the further review regarding measurement of the balance between primary care and hospital care in Section 2.4.

2.3.1 The concepts related to the balance between primary care and hospital care

The search identified three different understandings and uses of the concepts related to the balance between primary care and hospital care. The first group of publications focus specifically on the balance between primary care and hospital/specialty care (or secondary and tertiary care) (Battistella & Chester, 1973). White et al. (1961) studied the proportion of adults going to different levels of care and presented the demand for and utilisation of health care resources in a diagraph he called “the ecology of medical care”. McMillan et al. (1970) studied the proportions of general practitioners and specialists among physicians. Mills (1990) studied the allocation of funds among various levels of care. Starfield (1994) constructed a comprehensive index to measure the degree of primary care orientation of health systems. Overall, these studies represented focuses on the services utilised, the resources allocated, or the orientation of health systems.

Godber, Robinson, and Steiner (1997) also discussed “shifting balance of care towards primary care”, which was defined as the situation when a service acquired additional attributes of primary care (defined as direct access, generalist care, longitudinal care and delivery in a community setting, the concept of primary care will be elaborated later in Sub-section 2.3.2). Depending on how many of the four attributes were displayed before and after the change, the shift can be either full or partial. Based on a workshop with staff working at the primary-secondary care interface and a literature review, Godber and colleagues identified the main forms of full shift as: substitution for hospital care by primary care (general practitioners (GPs) providing minor surgery, employing specialists), primary care access to diagnostic services, and

day surgery and/or early discharge arrangements; and the main forms of partial shifts as: shared/integrated care; home care, domiciliary care and/or early discharge with follow up by secondary care providers; and specialist outreach clinics.

The second group of studies are a strain of literature more specifically associated with the “balance of care approach”. A recent systematic review (Tucker et al., 2013; Tucker et al., 2015) on the use of the approach summarised the lessons from this body of literature. Hence, here I draw from the review and the articles referenced by the two articles that came out of the systematic review. The concept of “balance of care” in this sense appeared in 1974 in relation to the development of mathematical models for resource allocation planning by the British government, and referred to “a balance between the major types of care and groups of patient” (McDonald, Cudderford, & Beale, 1974). The focus of the “balance of care approach” was the margin of care where alternative forms of care existed for specific patient groups. As Mooney (an early architect of the balance of care approach) explained: “to determine the balance of care it is not necessary to determine the costs and benefits of care for all individuals but merely those likely to be affected by small shifts in the existing boundaries of care, that is those of relatively high dependency levels in residential homes and those of relatively low dependency levels in hospital.” (Mooney, Russell, & Weir, 1986, p. 32). Due to this focus on the “margins of care” (Hughes & Challis, 2004), the field was also referred to as “marginal analysis” (Mooney & Drummond, 1982). Indeed, the overwhelming majority of studies using the balance of care approach were about older people and the settings of hospital, care homes, and community services, while some other special population groups like children, people with mental health problems, and people with HIV/AIDS were also included (Tucker et al., 2013). The balance of care approach is considered to have an advantage in helping service planners to consider the potential cost and outcomes involved in changing the delivery of services, which cannot be achieved by cost-effectiveness analysis of individual technologies (Tucker

et al., 2013). This approach arguably is limited in that it does not take into consideration the potential of effective primary care in preventing illnesses and dependency or reducing unnecessary referral for investigation in patients at low risk of serious disease.

The third group of studies also used the concept of “balance of care” as a broad health system strategy, referring to balance across multiple dimensions in health (and social) care. Indeed, the first article that mentioned “altering the balance between primary, secondary, and tertiary levels of care” also discussed the balance “between curative and ‘carative’ services” (Battistella et al., 1973). The use in this case has also been more general in the dimensions of “balance”. As quoted below, Royston (1998) asked “why, when, where, how, what and who” about balance between various levels of care and observed that both changes in demand (driven by demographic, epidemiological and social-economic changes) and changes in supply (driven by medical, information and wider technological development) pushed for shift of “balance of care”:

“When care is delivered: the balance between prevention and treatment.

Where it is delivered: the balance between institutional and community care.

How it is delivered: the balance between patient and professional involvement in care.

What is delivered: the balance between knowledge and habit based care.

Who is cared for: the balance between care of one group and another.”¹

A report by the Scottish Executive (2004) pushing for shifting “balance of care” used the concept in a similar sense. Although “shifting the balance of care” was not defined explicitly, it was used to refer to reducing reliance on hospital-based treatment

¹ Underlines were used throughout the thesis for cited words italicized in original sources.

(increasingly through emergency admissions) to preventive medicine, supporting self-care, greater targeting of resources on those at greatest risk and “a more proactive approach in the form of anticipatory care services”. The Scottish government later published a review of evidence on “shifting the balance of care” (Johnston, Lardner, & Jepson, 2008). In the report, “shifting the balance of care” included changes on three dimensions: shifting the focus of care towards long-term conditions, shifting the location of care towards more community-based facilities, and changing roles and responsibilities of patients and professionals.

2.3.2 The concept of primary care

As will be reflected below, primary care is a complicated concept with varied interpretations over time and space. The sub-section first briefly introduces the origin of the term, then presents the different approaches in which primary care has been defined, and finally discusses the relationship between primary care and primary health care.

1) Origins

The notion of “primary care” (and closely related terms) is generally believed (Frenk, 2009; Starfield, 1998) to have formally originated in *the Interim Report on the Future Provision of Medical and Allied Services* in 1920. In the report, “Primary Health Centre(s)” were proposed as “institution[s] equipped for services of curative and preventive medicine to be conducted by the general practitioners of that district, in conjunction with an efficient nursing service and with the aid of visiting consultants and specialists” (Lord Dawson, 1920: Section II). According to this report, these primary health centres would be treating more simple cases by general practitioners, while difficult cases would be referred to “Secondary Health Centres” staffed by consultants and specialists, and then to tertiary hospitals. As historical studies suggest, Dawson’s proposal of primary health centres was a reorganisation of general practice

and other facilities at that time (Honigsbaum, 1979: Chapter 6; Webster, 1993). Besides general practitioners and visiting consultants and specialists, the primary health centres would also include “[m]edical, surgical and maternity beds” (Lord Dawson, 1920: Section VII). This proposal did not materialise.

In the United States, White claimed in an interview to have invented the concept of “primary care” (White, 1998), which originally appeared in 1961 as “primary, continuing medical care” (White et al., 1961). The article provided no definition for the concept. According to what White said during the interview, the term was coined as a euphemism for general practice, which was suffering from stigmatization as an inferior subject of medicine. In the same interview, White acknowledged the earlier existence of the ideas about primary care, which included the systematically organised rural health system in Ding County in China established by Chen Zhiqian, as well as the regionalization of health care in China lead by Chen’s colleague and teacher John Grant.

2) Various definitions

Over the years, the concept of primary care has had a range of definitions (see Table 2-4 for a list of representative definitions). This literature review found that most definitions of primary care were proposed by experts in North America (Institute of Medicine, 1996: Chapter 2; Muldoon, Hogg, & Levitt, 2006; Shi, 2012; Starfield, 1998)², though other definitions were also discussed by experts with the World Health Organisation and from European countries (Saltman et al., 2005; Vuori, 1985; World Health Organization, 1978, 2008). The review identified several approaches to which

² A report by the King’s Fund (a British health policy think-tank) used the definition based on (Starfield, 1998).

these definitions belonged.

First, primary care can refer to a range of clinical professions or disciplines: closely associated with first contact care: general practice (Royal College of General Practitioners, 1972), or family medicine, general internal medicine, general paediatrics, obstetrics and gynaecology, and nurse practitioners (Shi, 2012).

Second, primary care can refer to a level of care, a type of setting or a system of service delivery. As a level of care, primary care is a point of entry to a health delivery system, along with secondary care, tertiary care, and emergency care (Starfield, 1998, pp. 8-9). Primary care is also referred to as health services located in the settings of communities (Institute of Medicine, 1996, p. 32). According to a systematic review of the core dimensions of primary care (Kringos et al., 2010), primary care is also seen as a multi-dimensional system structurally consisting of governance, economic conditions and workforce development.

Third, primary care has been used to refer to a set of services that are “person-focused (not disease-oriented)” and cover “all but very uncommon or unusual conditions” (Starfield, 1998, pp. 8-9). In terms of services, the definition of primary care has gradually evolved from being a synonym with general practice (or indeed a euphemism of it) as defined by White et al. (1961) to a multi-disciplinary team that provides the first contact care, satisfies comprehensive health needs centring on persons (rather than diseases and specialities), and functions as a hub for services at other levels of care (World Health Organization, 2008).

The fourth group of definitions saw primary care as services with a set of normative attributes which can be used to measure services at micro level: “accessible, comprehensive, coordinated, continuous, and accountable” according to the Institute of Medicine (1978), or “first contact, accessibility, longitudinality, and

comprehensiveness” according to Starfield (1992).

Fifth, in relation to the previous category’s set of normative functions but at a more macro level, some have suggested primary care as a set of principles for health system or health system strengthening (Starfield, 1998, p. 9).

Among the five approaches, the first three are more empirical definitions (focusing on service ranges or their providers) and the last two are more normative definitions (focusing on aspirational attributes). A single definition can cover more than one of these approaches listed above. The different approaches of definitions sometimes interrelate to each other. For example, the first approach of definitions based on primary care disciplines essentially describes the clinical staff under the disciplines that provide services falling under the third approach of definitions. One approach may also contradict another. For example, according to Saltman et al. (2005, p. 14), primary care can include services provided by secondary care and tertiary care providers, which seems to be incompatible with the definitions under the second approach based on levels of care. In relation to the various definitions, multiple criteria have been applied and/or proposed for distinguishing primary care from other types of health care. As Shi (1995) argued, primary care and specialty care (including secondary and tertiary care) can be distinguished by the time (first contact or not), the focus (diseases/organs or person as a whole) and the scope (comprehensive or episodic) of services. Godber et al. (1997) mentioned three other criteria for drawing such lines: key inputs (GPs vs. specialists), location of care (community settings vs. hospitals), and essential service attributes.

3) Distinguishing concepts

The review identified interchangeable use of the terms of “primary care” and “primary health care”, as well as efforts to distinguish them. The second column of

Table 2-4 provides some representative definitions of primary health care for comparison. The concept of primary health care was popularised by the Declaration of Alma Ata at the International Conference on Primary Health Care held in 1978. Like primary care, primary health care has also been defined in multiple ways, namely, as services with certain normative characteristics (World Health Organization, 1978), a range of specific services (World Health Organization, 1978), a development strategy for health services (Vuori, 1985), a philosophy of health (Vuori, 1985), and a set of policy orientations and reforms towards health for all (World Health Organization, 2008, p. xvi).

Table 2-4. Representative definitions of primary care and primary health care

Definitions of primary care	Definitions of primary health care
<ul style="list-style-type: none"> ➤ “<i>accessible, comprehensive, coordinated and continual care delivered by accountable providers of personal health services.</i>” (Institute of Medicine, 1978) ➤ “<i>care that is characterized by first contact, accessibility, longitudinally, and comprehensiveness</i>” (Starfield, 1992) ➤ “<i>Primary care is the provision of <u>integrated, accessible health care services</u> by <u>clinicians</u> who are <u>accountable</u> for addressing a large <u>majority of personal health care needs</u>, developing a <u>sustained partnership</u> with <u>patients</u>, and practicing in the <u>context of family and community</u>.” (Institute of Medicine, 1996, p. 32)</i> ➤ “<i>Primary care is that level of a health service system that provides entry into the system for all new needs and problems, provides person-focused (not disease-oriented) care over time, provides care for all but very uncommon or unusual conditions, and coordinates or integrates care provided elsewhere or by others.</i>” (Starfield, 1998, pp. 8-9) ➤ “[P]rimary care is an approach that forms the basis for and determines the work of all other 	<ul style="list-style-type: none"> ➤ “<i>Primary health care is essential health care based on practical, scientifically sound, and socially acceptable methods and technology made universally accessible to individuals and families in the community through their full participation and at a cost that the community and country can afford to maintain at every stage of their development in the spirit of self-reliance and self-determination. It forms an integral part of both the country’s health system, of which it is the central function and the main focus and of the overall social and economic development of the community. It is the first level of contact of individuals, the family and community with the national health system bringing health care as close as possible to where people live and work, and constitutes the first element of a continuing health care process.</i>” (World Health Organization, 1978) ➤ a set of activities outlined in the Declaration of Alma-Ata, including “at least <i>education concerning prevailing health problems and the methods of preventing and controlling them; promotion of food supply and proper nutrition;</i>

-
- levels of health systems. Primary care addresses the most common problems in the community by providing preventive, curative, and rehabilitative services to maximize health and well being. It integrates care when there is more than one health problem and deals with the context in which illness exists and influences the responses of people to their health problems. It is care that organizes and rationalizes the deployment of all resources, basic as well as specialized, directed at promoting, maintaining, and improving health.*” (Starfield, 1998, p. 9)
- *“Primary care refers to directly accessible, first contact ambulatory care for unselected health (related) problems; offers diagnostic, curative, rehabilitative and palliative services in response to the bulk of these problems; offers prevention to individuals and groups at risk in the population served; takes into account the personal and social context of patients; is provided by a variety of disciplines, either within primary care, secondary care or related sectors; assures patients continuity of care over time as well as between providers.”* (Saltman et al., 2005, p. 14)
 - *“a multidimensional system... [including] ten core dimensions that constitute a primary care system. The structure of a primary care system consists of three dimensions: 1. governance; 2. economic conditions; and 3. workforce development. The primary care process is determined by four dimensions: 4. access; 5. continuity of care; 6. coordination of care; and 7. comprehensiveness of care. The outcome of a primary care system includes three dimensions: 8. quality of care; 9. efficiency care; and 10. equity in health.”* (Kringos et al., 2010)
- an adequate supply of safe water and basic sanitation; maternal and child health care, including family planning; immunization against the major infectious diseases; prevention and control of locally endemic diseases; appropriate treatment of common diseases and injuries; and provision of essential drugs.”* (World Health Organization, 1978)
- *a strategy under which services would be “accessible; relevant to the needs of the population; functionally integrated; based on community participation; cost-effective; and characterized by collaboration between sectors of society; and a new distribution of resources between primary, secondary and tertiary care... a reorientation of health personnel (in terms of numbers, training and activities) towards primary health care.”* (Vuori, 1985)
 - *a philosophy of health: “social justice and equality; international solidarity; self-responsibility; and an acceptance of a broad concept of health.”* (Vuori, 1985)
 - *a set of policy orientations and reforms needed for moving towards health for all: “moving towards universal coverage; shifting service delivery to people-centred primary care; ensuring health in all policies; promoting inclusive leadership and governance.”* (World Health Organization, 2008, p. xvi)
-

Note:

- 1) Muldoon et al. (2006) listed some Canadian definitions of primary health care that are close to or included in the definition in the Declaration of Alma Ata and are not included in the table.
- 2) Underlines were used for originally italicized words.

The two terms have been used interchangeably by some. The Institute of Medicine (1978) originally used the term “primary health care” when referring to what they eventually called “primary care”. Starfield (1998) occasionally used primary care to refer to primary health care, but also to an intermediate stage of “conventional primary medical care striving to achieve the goal of primary health care”. Atu (2004) used primary care and primary health care interchangeably. Treating primary health care and primary care as synonyms, the World Health Report 2003 pointed out that primary health care refers to the first level of care in high- and middle-income countries, while the same term is “seen more as a system-wide strategy” in low-income countries (World Health Organization, 2003: Chapter 7).

Others see differences in the two terms and have tried to distinguish them. The Institute of Medicine (1996, pp. 30-31) later made the distinction that primary health care consists of population-oriented public health services (with the examples of sanitation and ensuring clean water for populations), and primary care focused on delivery of “personal health services in a context of family and community health”, while acknowledging increasing intersections and changing connections between public health and personal health care delivery (examples were immunization for children and tuberculosis outreach services). Saltman et al. (2005, p. 6) also noted that primary care was an intermediary concept, which was broader than the care delivered by a general practitioner or a family doctor, yet “more restricted than intersectoral concept of primary health care” in the sense of Alma Ata. Muldoon et al. (2006) argued that the two concepts were very different but unfortunately similar in language. They proposed that primary care be used to refer to “family doctor-type” services for individuals, while primary health care be used to describe “public health-type” services for both individuals and population. The World Health Organization Regional Office for Europe (2016) referred to primary care as a subset of primary health care. The case to clarify the definition of primary care was made strongly in the World Health Report

2008, which particularly criticized the unacceptable over-simplification of primary care in resource-constrained settings as:

“only deal[ing] with a few ‘priority diseases’; ... reduced to a stand-alone health post or isolated community-health worker; ... restricted to a one-way delivery channel for priority health interventions; ... just be about treating common ailments; ... synonymous with low-tech, non-professional care for the rural poor who cannot afford any better; ... financed through out-of-pocket payments on the erroneous assumption that it is cheap and the poor should be able to afford it” (World Health Organization, 2008, p. 11).

2.3.3 Commentary

As the section has shown, the balance between primary and hospital care (or between levels of care more generally) has been studied in the context of margins of care, across a range of health services, and between the hospital (or specialist care) sector and the primary care sector. This finding suggests that the various approaches to conceptualization of the balance between primary and hospital care have been grounded in the policy issues of interest. Therefore, this thesis needs to pay attention to the situation in China in order to conceptualize balance of care in an appropriate way that facilitates addressing key policy issues. The development of the concept of primary care over the years has given rise to different understandings and emphases in different studies. While some have defined primary care more from a positive perspective, others have defined it in a normative sense. Again, the definitions of primary care appear to have been based on specific contexts and policy issues at stake. The review has also revealed that the origins of primary care were closely linked with efforts to improve the balance between primary and hospital care within health systems.

The focus of literature on primary care seems to have at times slipped away from a system-wide perspective involving the balance between primary and hospital care to a narrower focus on primary care as a stand-alone field, which for instance constituted a solution to the overuse in hospitals. Such thoughts were clearly reflected in the

conclusion that White (1998) drew from the WHO International Collaborative Study of Medical Care Utilization: “no matter how many beds per thousand population, if there was adequate primary care and people had access to a general physician, the use of hospital beds was much lower.” On the other hand, there also seems to be some lasting attention to the issue of the balance between primary and hospital care. For example, in the World Health Report 2008, primary health care was defined as a multi-dimensional orientation for health systems including “shifting service delivery to people-centred primary care” (World Health Organization, 2008). This thesis follows this latter tradition of understanding the balance between primary and hospital care as a system-wide concept.

2.4 Measurement of the balance between hospitals and primary care providers

This section reviews the tools used in previous studies to measure the balance between hospitals and primary care providers. The review has not identified any comprehensive measurement or categorisation for dimensions of the balance between hospitals and primary care providers. Instead, the section draws on a related set of measurement--primary care orientation of health systems--to inform the categorization of available measurement tools in this thesis. To measure the primary care orientation of health systems, Starfield (1994) developed structural indicators--five “health system characteristics”, which reflect the capacity to achieve the intended features of primary care; and performance/process indicators--six “practice characteristics”, which reflect the degree of success in attainment of these features. Inspired by this dichotomy, the review used the categorisation of indicators identified from the literature as indicators representing structural aspects of the balance between primary and hospital care--referring to the comparison of the capital and the capacity between hospitals and primary care facilities, and functional aspects--referring to the comparison of the functioning and the performance between hospitals and primary care facilities.

2.4.1 Structural aspects

Three out of the five system/structural characteristics contained in the scoring system for primary care orientation developed by Starfield (1994) could be seen as related to structural balance between hospitals and primary care providers: type of physicians designated as primary care doctors; the professional earnings of primary care doctors compared to specialists; and the number of primary care doctors compared to specialists. In other words, the quantity and quality of human resources have been used as structural aspects of the balance between hospitals and primary care providers.

2.4.2 Functional aspects

This review identified two main groups of studies that measured the functional aspects of the balance between primary and hospital care. One group looks at service utilisation/provision. The other group focuses on resource utilisation/allocation.

General distribution of patient visits (i.e. types of provider that patients visit) has been frequently used as a key indicator of the functional balance between primary and hospital care. A landmark study in this area is *the Ecology of medical care* (White et al., 1961) mentioned earlier. Later White also led the WHO International Collaboration Study on Medical Care Utilization (Kohn & White, 1976), which described and analysed the geographical variations of service utilisation across a number of countries. Closely related to White's studies of service utilisation were the six practice characteristics in the scoring system for primary care orientation developed by Starfield (1994). Four of these characteristics could be seen as directly pertinent to the functional aspects of the balance between primary and hospital care: "the extent to which people first sought care from their primary care physician"; "the strength of relationships between people and primary care physicians"; "the extent to which primary care practice dealt with common needs"; and "the degree of coordination between primary care and other health services".

Other indicators/tools have been used to reflect the function of primary care or hospital care, or the effectiveness of referral systems, and are less obviously related to the balance between the two sectors. The remaining two practice characteristics contained in the assessment for primary care orientation of the scoring system developed by Starfield (1994) were “family orientation of primary care” and “community orientation of primary care”. Another related approach regarding the services utilised is the appropriateness of the setting where care is provided: the degree to which appropriate services are provided in appropriate settings (Lavis & Anderson, 1996). Widely used tools for measuring the appropriateness of service setting include the Appropriateness Evaluation Protocol (Gertman & Restuccia, 1981) and the Intensity-Severity-Discharge review system (Inglis et al., 1995). Both tools measure the appropriateness of a hospital admission or a day of stay in hospital. Sanders et al. (1998) studied how well case-mix (in terms of severity of illness) matched the level of health care facilities in Zimbabwe. Another approach uses ambulatory care sensitive conditions (ACSCs) to reflect the function of primary health care. ACSCs are a set of diseases and health problems, for which good quality primary health care were considered able to reduce hospital admissions (Caminal et al., 2004). However, this approach cannot reflect the balance between the two sectors, indeed there is inconsistent evidence on how it well ACSCs are negatively associated with improved primary health care (Gibson, Segal, & McDermott, 2013; Rosano et al., 2013). Overall, these measurements reflect an attention to the appropriateness and effectiveness of services.

Regarding financial resources utilized, a cross-national analysis by Mills (1990) mentioned earlier investigated the proportions of hospitals (and other types of facilities for some indicators) in total health sector expenditures, total health sector expenditures of government, annual development (capital) expenditures on health, national government health facility expenditures, and district health facility expenditures.

Nabyonga, Munguti, and Thomas (2002) studied public sector expenditures by level of care (hospitals, outpatient care centres, etc.) and the percentages of total expenditures on hospitals vs primary care facilities in particular in seven African countries. Mugisha and Nabyonga-Orem (2010) analysed the proportion and amount of recurrent expenditures that went to hospitals and the primary health care sector in Uganda from 1997/1998 to 2007/2008. The Primary Health Care Performance Initiative (2015) reported 25 PHC Vital Signs indicators, one of which was “Percent of government health spending dedicated to PHC”.

2.4.3 Commentary

The review identified a lack of a comprehensive system of metrics for the balance between hospitals and primary care providers. Previous studies measuring the balance between hospitals and primary care providers used indicators mentioned above that can be grouped into either structural or functional aspects. Indicators for the structural aspects include quantity and quality of human resources. Indicators for the functional aspects include service utilisation and provision, and resource utilisation and allocation.

Left out in the dichotomy of structural and functional aspects seem to be the institutional factors that underlie the balance between hospitals and primary care providers. Two indicators in the five system/structural characteristics contained in the index for primary care orientation developed by Starfield (1994) are cases in point: the extent of regulations that allow health professionals and facilities to be distributed according to need; and risk pooling for health care services. Measurement for such institutional factors are obviously different from the quantitative measurement of the capacity and performance aspects of the balance between primary and hospital care and deserves special treatment. The thesis distinguishes between structural and functional aspects of the balance between primary and hospital care on the one hand and policies and interventions to change the balance (e.g. empowerment and

incentivisation of primary care providers as gatekeepers) on the other hand.

2.5 Findings regarding strategies to shift the balance between primary and hospital care

The section presents the findings on effective strategies, particularly gatekeeping, in shifting the balance between primary and hospital care, and extends to a synthesis of important issues related to complexity of the balance.

2.5.1 Effectiveness of gatekeeping in strengthening primary care functioning

This section reviews knowledge about “what works” in shifting the balance. Before proceeding, it should be clarified that the focus is whether and/or to what extent interventions are (likely to be) effective in causing a shift of the balance between primary and hospital care to happen, not whether and/or to what extent a shift in the balance is effective for a secondary purpose. It is obvious that this strict criterion would exclude a broad range of relevant studies focusing on strengthening primary care without incorporating analysis of hospital care.

Findings from a systematic review in several high-income countries suggest that a range of interventions could contribute to shifting the balance between hospitals and primary care providers (as defined in this thesis): assessment of elderly people before case management, disease management (especially in relation to long-term conditions), early supported discharge with community-based rehabilitation for stroke and other patients, rehabilitation in the community for a range of conditions, care at home and hospital at home interventions (Johnston et al., 2008). Indeed, various interventions and policy approaches are possible in shifting the balance towards primary care. The remainder of this sub-section focuses on gatekeeping because of the claims made about its potential contributions and the need to keep the review manageable.

Gatekeeping has been defined as an arrangement between primary care providers

and specialists which involves a generalist (primary care doctors, family medicine doctors, general practitioners, etc.) who acts as a gatekeeper for specialist care and coordinates care for patients, i.e. that in order to access specialist care patients have to be referred by a primary care practitioner, and is supposed to work as a mechanism to enhance greater efficiency in use of health care resources (Starfield, 1994). Historically, gatekeeping probably emerged in two ways: first employed by health insurers to approve sickness claims, and then to excess restriction of access to hospital care financially or administratively (Stone, 1978). The search on gatekeeping identified a recent systematic review (Garrido, Zentner, & Busse, 2011) considered by independent assessors at Centre for Reviews and Dissemination (2011) to be of good quality, which has provided a summary of empirical studies about gatekeeping before 2010. For years between 2010 and 2016 (September), the review here included another publication from Egypt (Ward, 2010), covering three indicators: visits to primary care providers, visits to specialists, expenditures for ambulatory specialist care.

One study (Ferris, Perrin, et al., 2001) included in the systematic review was dropped due to severe selection bias³. Only three studies in the systematic review were thus suitable for comparison and provided limited information. A US study by Escarce et al. (2001) reported an additional 50.9% of primary care visits on average in insurance plans with gatekeeping compared to a plan without such requirements. Schwenkglens et al. (2006) reported, based on a study in Switzerland, no change in primary care visits after introducing gatekeeping. While Escarce et al. (2001) reported a 77.4% reduction of visits to specialists in gatekeeping plans, a German study by Ose et al. (2008) reported little effect of gatekeeping in reducing specialist care utilisation-

³ In the study of Ferris, Perrin, et al. (2001), beneficiaries were allowed to choose between plans requiring gatekeeping and without such requirement. As the author acknowledged, the reduction of utilisation and expenditures care in both primary care and specialist care was likely mainly due to selection bias since beneficiaries who expected lower utilisation of care tended to opt for the gatekeeping plan.

-a gatekeeping scheme saw a decrease of 4% of visits to specialists per year on average, while the reduction was 5.8% in a comparison group without gatekeeping. Using a survey from a pilot district with gatekeeping and a control district without gatekeeping, Ward (2010) reported about 63% reduction in visits to hospital outpatient departments, with 10.75% increase in visits to primary care facilities (in contrast to a decrease of 29% in a comparison area without gatekeeping).

Gatekeeping studies identified here could be generally categorized into three groups. The first group focused on gatekeeping as a feature of managed care, and mainly corresponds to the rise and fall of the managed care movement mainly in the US but also in other countries without gatekeeping policies such as Germany and Switzerland between 1980s and early 2000s (e.g. Ferris, Chang, et al. (2001) and Perneger, Etter, and Rougemont (1996)). The second group of studies were in countries with previously strict gatekeeping policies trying to respond to greater calls for choice, particularly in the case of National Health Service Scotland (Perneger et al., 1996). The study from Egypt (Ward, 2010) was the only study in the third group, addressing the challenge of hospital-centred care in developing countries.

In the Chinese literature, despite repeated reports showing a large proportion of cases could be treated more cost-effectively at lower levels of care (Lei et al., 1996; Wang, Gusmano, & Cao, 2011a; Xu, Diwan, & Bogg, 2007), studies about the effectiveness of interventions aimed at shifting the balance between primary and hospital care were rare. The only article included was of low quality. Guo, Liang, and Jin (2012), pooled two cross-sectional surveys of residents in the catchment of a community health centre in Beijing that had implemented a payment reform which replaced fee-for-service with global budget, rendering budget independent from providing services. Without using a control group, the study found that the proportions of patients visiting primary care facilities for initial contact rose from 32.6% at baseline to 61%. The quality of the study was further compromised by not providing sufficient

details of methods.

2.5.2 Issues of complexity in shifting the balance between hospitals and primary care providers

As the guideline on evaluating complex interventions by the Medical Research Council (2008) pointed out, “[f]ew interventions are truly simple, but there is a wide range of complexity”. The guideline listed five dimensions that make an intervention complex, and all dimensions seem to apply to interventions on the balance between primary and hospital care:

- “ • *Number of and interactions between components within the experimental and control interventions*
- *Number and difficulty of behaviours required by those delivering or receiving the intervention*
- *Number of groups or organisational levels targeted by the intervention*
- *Number and variability of outcomes*
- *Degree of flexibility or tailoring of the intervention permitted”*

In response, the guideline suggested that 1) good theoretical understanding be developed about how the intervention causes changes, in order to address potential “weak links in the causal chain”, 2) the process evaluation be used to identify implementation problems (which could explain lack of impact); 3) sample size might need to be increased to be able to account for variability, while cluster- rather than individually-randomized designs might need to be considered; 4) studies should employ a range of measures and to identify unintended consequences; 5) interventions might need to adapt to local setting instead of strictly following protocols. The positions suggest a need to clarify complexity issues and reflect them in research approaches, though no example was found which used these guidelines in studying the balance between primary and hospital care.

This review has identified some complexity issues involved in shifting the balance

between primary and hospital care (gatekeeping in particular) and their implications for studying strategies to shift the balance between primary and hospital care.⁴ First, interventions under similar categorization are implemented in varied ways. All gatekeeping studies have involved some other changes in what could be seen as complementary arrangements. The gatekeeping study from Egypt, for example, involved increased charges and co-payment for patients bypassing primary care providers, a marketing campaign to raise awareness, and an exemption policy for the poor who received care at the primary care level or received care at hospitals with referral from primary care (Ward, 2010). Indeed, the findings of the Organisation for Economic Co-operation and Development (OECD) Health System Characteristics Survey 2012 showed varied arrangements of gatekeeping and cost-sharing policies for accessing outpatient specialist care across rich countries (Organization for Economic Co-operation and Development, 2013). Studies on pro-coordination reforms including gatekeeping also found that coordination reforms were not self-sustaining but required supporting conditions at systems, professional, and practices levels (Rico et al., 2003; Saltman et al., 2005).

Second, there are issues related to the context within and beyond health systems. A study in the Netherlands suggested that GPs use a “demand-satisfying” attitude when it comes to gatekeeping, even though they think patients are receiving unnecessary care (Wammes et al., 2014). Important contextual changes that drive policy changes include the increasing demand on coordination caused by rising prevalence of non-communicable diseases, as well as pressures related to cost-containment (Rico et al., 2003).

Third, there are issues related to the various and interrelated impacts of the

⁴ The review draws from Greenfield, Foley, and Majeed (2016) and Rico, Saltman, and Boerma (2003).

gatekeeping policies. Financial incentives for general practitioners in the United Kingdom (UK) to reduce specialist referrals caused concerns about conflicting roles and ethical issues (Lauridsen, 2009; Matthews-King, 2016). While pro-gatekeeping financial incentives seemed to reduce access to specialist care in France, particularly for poor and uninsured people (Dourgnon & Naiditch, 2010), gatekeeping also reduced overall health care inequities in European countries (Biro, 2013; Reibling & Wendt, 2013; Schnitzer et al., 2011), provided disadvantaged groups with assistance in decision making, and reduced specialist use by advantaged populations, who were the more frequent users of specialist care (Reibling et al., 2013). There have been different opinions regarding the influence of gatekeeping on the traditional divide between GPs and specialists, particularly in relation to the flow of information between the two groups (Bjornsson et al., 2010). Restrictions on patient choice are associated with lower patient satisfaction (Bjornsson et al., 2010; Dusheiko et al., 2007; Greenfield et al., 2012), though not everywhere (Gervas, Ferna, & Starfield, 1994), while low patient satisfaction has been shown to be associated with compromised outcome and compliance (Roter & Hall, 2006). Studies have also indicated that primary care gatekeeping may actually be associated with poorer cancer survival potentially due to delayed diagnosis (Crawford, 2014). In short, the results of policies to shift the balance between primary and hospital care (particularly gatekeeping) appear to include a range of interrelated consequences, some of which were not necessarily the intended outcomes of the interventions, highlighting the complexity of the interventions.

Fourth, history and institutions also influenced the implementation of strategies to shift the balance between hospitals and primary care providers. A comparative study (Rico et al., 2003) referenced above analysed trajectories of development in relation to strengthening the role of primary care both in service provision and coordination (including gatekeeping) across health systems in European countries during the 1990s. The study showed that health systems in which primary care historically had strong

informal political power (particularly private ownership) and state had monopsonic control of health services tended to find it easier to implement primary care-focused pro-coordination reforms. The distribution of informal political power and state control were historically formed, and affected by policy feedback and societal pressures (cost-containment, market competition and increasing prevalence of chronic illnesses). In other words, this study indicated the importance of power dynamics in determining the outcome of policy interventions related to the balance between primary and hospital care. Such power appeared to come from historically shaped institutional factors that determined distribution of power and formation of coalitions among groups, as well as the political structure that affected the influence of these coalitions in driving or preventing policy changes.

2.5.3 Commentary

Section 2.5 has reviewed findings regarding strategies to shift the balance between primary and hospital care with a focus on gatekeeping. Research about gatekeeping has been especially rare, particularly in low- and middle-income settings. Studies on effectiveness of strategies to shift the balance in China have been rare and of low quality. The point about gatekeeping involving some concurrent arrangements deserves particular attention.

The review has also identified complexity issues related to shifting the balance between primary and hospital care. Table 2-5 summarizes these complexity issues and extends to include their corresponding methodological implications.

Table 2-5. Complexity issues and methodological challenges

No.	Complexity issues	Methodological implications
1	Varied implementation of interventions under the similar categories	Research on effective strategies needs to be able to isolate the impact of the component of interest, and be complemented with understanding of implementation details.
2	Variations in the context both within and beyond health systems	Research needs to deal with the variations within the context, including concurrent changes and unobservable factors. Challenges of isolating the effectiveness of interventions from complex contexts
3	Multiple consequences, including interrelating and unintended consequences	Research needs to address not only the immediately intended outcomes, but also to examine the consequences not directly intended and the interrelationships between consequences of the intervention.
4	Historically formed institutional factors shape and interact with how and whether contemporary strategies might work	Research needs to understand how historically formed institutional factors influence the power dynamics affecting policy and system changes.

The issue of variations in implementation of intervention programmes suggests the need to go beyond the generic brands of certain intervention programmes and to identify the specificity of implementation. Methodologically, this means more ground-level understanding of implementation details, including the barriers to and facilitators of implementation in the context of interest. The issue of variations in the context both within and beyond health systems also requires research designs to explore/control for the variations within the context, which may include concurring changes and unobservable factors.

The issues about multiple consequences demand research methods to reveal the multitudes of interrelating consequences, and to be sensitive to interrelationships and unintended consequences. The important influences of historical and institutional factors on implementation of interventions also require a research approach to conceptualize and study such factors.

To sum up, it is difficult if not impossible to isolate effects of one intervention (or multiple interventions under one brand) embedded in multiple dimensions of complexity. Such complexity issues as discussed above pose methodological challenges that deserve attention to each one of them.

2.6 Review of approaches used in thesis

Besides the need to increase knowledge on measurement of the balance between hospitals and primary care providers, the review above suggests potential knowledge gains from three elements.

First, because interventions happen as changes embedded in concurrent development of health systems, identifying the impact (defined as the causal effect) of policy interventions requires isolation of intervention mechanisms from implementation issues and more or less separable contextual factors. In this regard, an impact evaluation was conducted to identify the causal effect of the pilot programme.

Second, understanding the mechanisms of interventions implemented, their contexts, and the interaction between intervention mechanisms and contexts contributes crucially to understanding why interventions had varied effects. To obtain in-depth information regarding the implementation processes, variation in context and the interaction between process and context, the study further adopted a systems analysis to understand the functioning of gatekeeping in the context of multiple on-going changes in the health system.

Third, understanding history can help elucidate the power dynamics within the parameters of institutional factors formed historically—this is crucial as the power dynamics determined the likelihood and directions of substantial changes in strengthening primary care. To address such power dynamics, the study also included an analysis using historical institutionalism.

The remainder of this section reviews the approaches used in the thesis corresponding to each of these elements. The rationale for choosing the research approaches is provided in Table 2-6. While an overview of how the thesis used the methods is provided in the next chapter, and the details of methods are provided in sub-study chapters, the remainder of this section reviews the literature regarding previous uses of the three approaches.

Table 2-6. Justification for selection of research approaches

Impact evaluation	<p>A critical justification of impact evaluation is the availability of a gatekeeping pilot in China. The pilot used a cluster design that matched the evaluation requirement.</p> <p>Difference-in-differences analysis in particular establishes the counter-factual situation to assist calculating the net effects of an intervention. By the creation of a counter-factual, the method reduces the risk of distorting the effects of an intervention by including the effects of other changes and inherent trends.</p>
Qualitative systems analysis	<p>Impact evaluation on its own provides little information about the actual causal pathways, implementation process and unintended consequences.</p> <p>Causal loop analysis (a sort of qualitative systems analysis) combined with qualitative research methods allows the study to clarify the intended mechanisms, identify the implementation process, elucidate the dynamics among multiple components involved in interventions, and gather information about multiple and unintended consequences.</p>
Historical institutionalism	<p>Historical institutionalism emphasises the power dynamics set in forth in a wider temporal-spatial context, which drives policy change and build up institutional resilience.</p> <p>Path dependence analysis has been applied to health systems analysis and found potentially valuable.</p>

2.6.1 Impact evaluation

Previous findings regarding the effectiveness of gatekeeping in shifting the balance between primary and hospital care are of generally low quality and focused on high-income countries, while evidence regarding effectiveness of gatekeeping in the context of China is also thin, as shown above. Indeed, given the existence of multiple

complexity issues that the previous section has alluded to, the effectiveness of gatekeeping in shifting the balance between primary and hospital care is likely to be influenced by the local context and implementation details. The mechanisms themselves may be compromised by the context, which also affects implementation. A particular challenge for evaluating effects lies in isolating the impact of interventions from the changing backgrounds, owing to the existence of multiple and interrelated influences of gatekeeping and concurrent changes.

Evaluations are defined as “periodic, objective assessment of a planned, ongoing, or complete project, program, or policy” to answer questions regarding design, implementation and/or results (Gertler et al., 2011). In general, evaluations can address descriptive questions (e.g. what is happening, in what processes, under what conditions, relationships between actors and how people see the process), normative questions (e.g. gaps in inputs, implementation and outputs) and cause-and-effect questions (what difference in the outcomes of interest is caused by the intervention) (Imas & Rist, 2009). Impact evaluation has a clear focus on the last group of questions and on demonstrating the existence (or absence) of causal relationship between an intervention (a programme, a policy, etc.) and its intended and unintended effects (Gertler et al., 2011, pp. 7-8).

Impact evaluation essentially measures the impact (causal effects) of an intervention as the difference in outcomes for the same subjects with or without enrolment in the intervention programme. However, subjects are not identical, and one subject cannot be a participant of the programme and a non-participant at the same time. In other words, the outcome of a subject without participation of the intervention programme cannot be observed from a participating subject. To fill this information gap, the impact evaluation methods estimate the counterfactual typically by using comparison groups which have identical characteristics to the treatment group, except for the absence of the intervention of interest. There are two commonly applied though

usually problematic (sometimes called “naïve”) methods of constructing comparison groups: before-and-after (comparing the outcomes of participants pre- and post-introduction of an intervention program, i.e. using the pre-intervention outcome of subjects as an estimate of the counterfactual), and with-and-without (comparing the outcomes of enrolled and non-enrolled, i.e. using the outcome of non-enrolled subjects as an estimate of the counterfactual). Before-and-after comparison risks neglecting the effects of changing environmental conditions, while with-and-without comparison risks neglecting the inherent differences between the two groups and potential selection bias. Neither methods provide a sound estimate of the counterfactual (Gertler et al., 2011: Chapter 3).

Among the methods of impact evaluation, the difference-in-differences method offers an approach to study impact by constructing a counterfactual when randomization is not feasible (Gertler et al., 2011: Chapter 6). The difference-in-differences method compares the changes in the outcome variables over a certain period between a treatment group—subjects that participate in a programme, and a control group—subjects that do not participate in the programme. The method essentially combines investigation of the before-and-after difference in the outcomes of the treatment group (through which factors constant over the period of study are controlled for) and the before-and-after difference in the outcomes of the control group (through which time-varying factors are captured). The most basic difference-in-differences method assumes that the control group and treatment group exhibit “equal trends in the absence of treatment”, and constructs a counterfactual based on the change in outcomes of the control group. Therefore, net effects of intervention are calculated by “cleaning” the time-varying effects (the second difference) away from the before-and-after change in the treatment group (the first difference). By considering the time-varying effects, difference-in-differences analysis is capable of generating a better estimate of the counterfactual than simple before-and-after studies.

Difference-in-differences is widely applied. Using difference-in-differences, Duflo (2001) analysed the impact of school construction on the schooling and labour market in Indonesia; Di Tella and Schargrotsky (2004) studied whether an increase in staffing of police forces contributed to crime reduction. Examples of difference-in-differences have also been seen in health systems research. For instance, Sutton et al. (2012) and Kristensen et al. (2014) used the method to examine whether the effects of a pay-for-performance scheme reduced in-hospital mortality in England. Thus, difference-in-differences provides a rigorous method to test the causal effect of a single intervention that would not be available in cross-sectional study, which is lacking in current studies about effectiveness of gatekeeping. None of the gatekeeping studies found used difference-in-differences analysis.

Limitations of difference-in-differences methods include the assumption of equal trends, and calculation of an average treatment effect for all enrolled subjects, i.e. without sensitivity to variations of treatment effects amongst individuals (though this focus on overall impact could also be seen as a strength) in calculating average changes. Furthermore, the attention to causal effects might be restricted with a narrow view on the complexity involved in a programme. As Mackie (1965) has explained, causes are at a minimum, conditions which are “Insufficient but Necessary parts of a condition which is itself Unnecessary but Sufficient” (INUS) for their effects. A classical example of INUS conditions given by Mackie is that an electrical short-circuit has caused the burning of a house. The short-circuit was a necessary part of this accident as it was indispensable in causing the fire, but it was not the necessary condition for the house to catch fire, as other conditions (e.g. a lighted oil stove as overturned) could have caused the fire. The short-circuit was not a sufficient condition for the house to catch fire, as other conditions (e.g. the absence of inflammable materials nearby) could have prevented the house from catching fire, but the short-circuit combined with a set of conditions constituted a complex condition which became sufficient for the house

to catch fire. As causes dealt by impact evaluation (and other similar approaches to causal inference) are INUS conditions, the evaluation technique focuses on only part of the causes and might end up providing oversimplified or even misleading results. In other words, impact evaluation should be complemented by more in-depth knowledge of the process of intervention embedded in a system of interrelated elements.

2.6.2 Systems theory and causal loop analysis

Issues related to the various arrangements, various contexts and multiple, interrelated impacts mentioned above, as well as the nature of gatekeeping which involves the two sectors of hospitals and primary care providers suggested the need for an approach that allows sufficient sensitivity about and synthesis of the multiple factors in their interrelating dynamics. Knowledge derived from systems theory and system thinking (also known as “the systems field” (Williams, 2015), as “systems theories” vary) have been advocated to be useful for health systems and policy research (Adam et al., 2012; Bloom & Standing, 2008; Gilson, 2012). Systems thinking has been described as a mind-set which sees systems and sub-components of systems as interrelated to one another, and interpreting the interrelationships as the key to knowledge about how things function (Adam, 2014). Two recent supplement issues published in peer-reviewed journals *Systems thinking for health systems strengthening in LMICs: seizing the opportunity* (2012) (Adam et al., 2012) and *Advancing the application of systems thinking in health* (2014) (Adam, 2014) illustrate the developing application of systems thinking in health systems research. Based on these two journal issues, systems thinking appears to facilitate addressing a growing range of questions associated with the impacts of interventions. As the 2012 supplement demonstrated (Adam et al., 2012), systems thinking has proved valuable in revealing key elements of success and failure in implementing complex interventions, the role of relationships, the role of actors in health systems, the importance of environmental factors, the

importance of anticipating potential unintended consequences, and systematically evaluating the implementation process and reacting to feedbacks within the systems. As the 2014 supplement further showed (Adam, 2014), combining qualitative methods with systems thinking can add depth to analysis of health systems issues, and visualization can help convey complex interpretations and findings.

Causal loop analysis is a method for systems analysis that maps out and qualitatively models the dynamics between factors within a system using causal loop diagrams (CLDs) (Williams et al., 2010a: Chapter 1). Through constructing CLDs, causal loop analysis helps identify the key elements/variables in the situation, the linkages/interrelationships among elements, and the feedback among interrelated elements. The basic unit of a causal loop diagram is a causal link. Each causal link between two variables has a direction and a polarity. Direction denotes the cause and the effect within a link, illustrated by arrows departing from the cause and arriving at the effect.

There are two types of link polarity in CLDs: positive and negative. A positive link means that, all else being equal, a change of the cause variable will lead to a change of the effect variable in the same direction, compared to the situation when the cause variable is held unchanged; in contrast, a negative link means that, all else being equal, a change of the cause variable will lead to a change of the effect variable in the opposite direction, compared to the situation when the cause variable is held unchanged. Connecting the links generates feedback loops (closed circular causal relationships) that may be linked to relevant variables that do not fall in any feedback loop.

There are two main types of feedback loops, namely, reinforcing loops, when the sum of negative causal links within the loop produces an even number, and balancing loops, when the sum of negative causal links produces an odd number (Richardson,

1986; Sterman, 2000: Chapter 5; Williams & Hummelbrunner, 2010b: Chapter 1). Compared to the linear model usually used for analysis of policy interventions, CLDs allow researchers to put together, in a visually accessible way, a multitude of interacting factors with feedback loops.

Causal loop analysis has been found valuable in health systems research. Rwashana et al. (2009) used data collected through various sources including preliminary interviews, questionnaires, documents, and field studies to create CLDs and used them to generate a wide and holistic perspective of the issues related to an immunization system for stakeholders to prioritize and develop policies. Rwashana et al. (2014) developed CLDs about demand and supply side issues associated with neonatal mortality in two divisions of a district with high neonatal mortality rates in Uganda, based on literature reviews and interviews. The study found synthesis of concepts through constructing CLDs facilitated understanding of the interactions and feedback loops contributing to the stagnant neonatal mortality rates, and provided an initial step to develop potential intervention strategies.

Paina et al. (2014) used CLDs to illustrate the interactions and feedback among actors involved in the development of medical dual practice in Uganda, and demonstrated the value of CLDs in visualizing the changes of policies and practice in relation to dual practice. Sarriot et al. (2015) constructed CLDs to identify systems determinants of integrated Community Case Management (iCCM) of malaria, pneumonia and diarrhoea from the perspective of the national health system. The study utilised CLDs to develop high-probability future scenarios and policy recommendations. In these studies, CLDs make explicit cause-and-effect relationships and facilitate understanding and interpretation of interacting factors and feedback loops that contribute to important policy issues in the field of health systems and policy research.

Williams neatly summarized three core concepts of systems theories in general: interrelationships, perspectives, and boundaries (Williams et al., 2010a; Williams, 2015). Systems scholars started with the focus on a multitude of interrelationships in which an intervention is embedded; they also emphasised the existence of multiple perspectives (e.g. different interpretations of an apparently similar situation); more recently, the issue about decisions regarding the boundaries of a systemic inquiry or intervention has also been highlighted. The three concepts are dimensions where systems methods could help improve understanding of a complex intervention within a dynamic system like a health system.

The advantage of CLDs has been clear in identifying and analysing interrelationships, however, s. Studies have also included multiple perspectives (e.g. Paina et al. (2014) examined the multiple perspectives of stakeholders) and broadened boundaries (e.g. Rwashana et al. (2009) examined the unintended consequences of the immunization system in Uganda). Causal loop analysis on its own does not provide a study design and is usually used in combination with qualitative methods (e.g. Varghese et al. (2014) used CLDs along with content analysis). In short, CLDs generally constitute an additional powerful tool to systematically incorporate systems thinking in along with qualitative methods, and therefore facilitate studying complex health systems issues. They have not yetso far been used to study balance of caregatekeeping or the balance between primary and hospital care.

2.6.3 Historical institutionalism and path dependence analysis

1) Rationale

The following conceptual discussion and review regarding historical institutionalism and path dependence arises from the assumption that China's configuration of hospitals and primary care providers cannot be understood without investigating the deep historical roots and dynamics through proper tools. The

historical approach has been gaining attention in the field of health policy and systems research. Bloom (2014) argued for the importance of situating health systems development in the context of its history and political economy. In the same commentary, Bloom drew from Leach et al (2011) to emphasise the need to understand the risk of powerful interests narrowing the options of questions asked about framing the complex reality of health systems. The historical approach focuses on the analysis of the process of change, and therefore has the potential to reveal the political economy of the health sector and the role of various actors. Path dependence theory and its close relative—“punctuated equilibrium” theory (in social theories) have been recognized as useful systems thinking theories for health systems research (Peters, 2014). A recent historical case study on health system development used the framework of path dependence to facilitate a study on factors contributing to successful health system development (Balabanova et al., 2011). No study has used historical institutionalism and path dependence systematically for studying the balance between primary and hospital care.

2) Key theories and concepts

Before the rise of historical institutionalism, there were multiple competing approaches to historical analysis of health and welfare policies. The structural functionalist approach explained the emergence of the welfare state as a result of industrialization, the Marxist explanation for the welfare state was based on class interests and the rise of the labour movement, the pluralist approaches to explain the welfare state focused on social interests treating the state as a neutral presence responding to external social actors (Skocpol, 1985). These approaches, however, have been found inadequate to explain the timings and configurations of national welfare states in health and in general. Historical-institutionalists particularly emphasised states as autonomous actors with bureaucrats of potentially varied capacities, who were not just strictly executing in the interests of certain social groups, or obediently

answering the call of certain ideologies (Hall & Taylor, 1996; Skocpol, 1985; Skocpol & Amenta, 1986). A group of historical-institutionalists focused on the structures of politics, and how much such structures allow interest groups to veto policies that run against their interests (Immergut, 1990; Steinmo & Watts, 1995).

Although the state may have an autonomous role in policy making, it does so bounded by limitations imposed by earlier choices. Path dependence is a way of thinking systematically about how policy choices are circumscribed by past decisions (Pierson, 2000). In a sense, historical institutionalism fuses the explanation of the emergence of institutions and the constraining effects of institutions (Fioretos, Falletti, & Sheingate, 2016: Part I). The theory of path dependence originated as researchers tried to explain the endurance of inefficient technologies and found that technologies could be locked in due to accidental historical events: the initial technologies created network externalities, learning effects, adaptive expectations and high fixed costs that increased the cost of changing technologies by so much that it was rational for individual actors to stick to the inefficient technologies (Arthur, 1989; David, 1994).

Later researchers also borrowed from the model of “punctuated equilibrium” (Gould & Eldredge, 1972) in evolutionary biology which challenged more incremental models of evolution (Krasner et al., 1984). The theory has been defined in different ways, some broader than others. Path dependence has been used to refer generally to “that what happened at an earlier point in time will affect the possible outcomes of a sequence of events occurring at a later point in time” (Sewell, 1996, pp. 262-263). It has also been defined narrowly as: “that once a country or region has started down a track, the costs of reversal are very high” (Levi, 1997, p. 28).

Pierson (2004) pointed out that policy actors who might want to change the established institutions or policies would likely face three type of challenges: coordination difficulties, asset specificity and positive feedback. Coordination

difficulties involve the interaction among multiple policies widely seen in the health delivery arena, and are likely to exist in relation to the issue of the balance between primary and hospital care. For instance, all major reforms of professionalization-related policy need to coordinate not only medical schools that produce graduates but also facilities who employ them. Coordination difficulties come from the adaptive expectation that in a fragmented decision-making environment, actors may have different opinions about what is the best outcome but would be eager to come to a shared acceptable outcome. Once this point is reached, the likelihood of deviating from the consensus would diminish, as it would entail formidable challenges (Hardin, 1989). A stronger form of coordination difficulties lies in the existence of multiple veto points in underpinning political institutions or institutions related directly to the balance between primary and hospital care. They constitute a conservative force that impedes further reform. The degree of coordination difficulties helps explain the fragility or resilience of institutions. Thus, the political institutions (both formal and informal structures) are important to path dependence analysis.

Specificity of previously accumulated assets poses difficulties for change. Joskow (1988) drew from Williamson's analysis of transactional costs and identified four types of asset specificity: site specificity (value of the asset is tied to its location), physical asset specificity (specificity is tied to its role as an integrated part of a transaction or product), human asset specificity (skills and knowledge that are specific to the position, or developed trust), dedicated asset specificity (value derives from continuing particular exchange). Assets in institutional studies take a wide range of meanings (relationships, expectations, privileges, knowledge of procedures) and are not readily reallocated.

The above two points are more related to the structural or the nature of individual elements of policies. Policies are not only more difficult to roll back due to such institutional constraints, but are also likely to trigger a process of increasing returns

(Arthur, 1994; David, 2001; Pierson, 2004), or a self-reinforcing process. While the first two points tend to be partial on setting barriers to change in a certain direction, change that satisfies previous structures would further develop. Hence, the three mechanisms can be mutually complementary.

Besides the reinforcing processes, reactive sequences were also identified, as chains of events that were temporally ordered and connected to one another in a causal relationship. Each event in a reactive sequence is “both a reaction to antecedent events and a cause of subsequent events.” Reactive sequences are characterized by “backlash processes which *transform[sic]* and perhaps *reverse[sic]* early events.” (Mahoney, 2000)

Building upon the theory of path dependence, researchers in the school of historical institutionalism tend to divide history into long periods of stability when institutions structure political outcomes, and intersecting short periods during which events external to the institutions in question make those institutions more malleable (Fioretos et al., 2016, p. 9) (illustrated as Type 1 in Figure 2-3). Two main concepts of such short periods featured prominently, namely critical juncture and conjuncture. The concept of critical juncture has been used to explain the formation or origin of institutions. According to Pierson (2004, p. 135), critical junctures “are ‘critical’ because they place institutional arrangements on paths or trajectories, which are then very difficult to alter.” Various definitions exist for critical junctures, some emphasise structural factors constraining the options of actors (Lipset & Rokkan, 1967). Others emphasise the existence of contingency during and divergence after critical junctures (Capoccia & Kelemen, 2007). Wilsford (1994) discussed “conjunctures” which focused on the coming together of events (war, crisis, major change of economic policy, etc.) that lead to change and the unravelling of institutions (illustrated as Type 2 in Figure 2-3). To sum up, critical junctures have been defined mainly to account for the locking of path dependence and the origins or formation of institutions, while

conjunctures have been defined mainly to account for the unlocking of path dependence and the termination or reforms of institutions (illustrated as Type 3 in Figure 2-3).

Among historical-institutionalists in the field of health policy and systems research, studies that link research on health (usually as a welfare) politics and health service research seem to have moved beyond traditional territories of political scientists and ventured into an interdisciplinary field of health policy and systems research from a historical perspective. Two particular examples are illustrative. Hacker (1998) studied the historical logics of National Health Insurance in UK, US and Canada with the main question being why the United States, given its similarities with UK and Canada, did not have national health insurance. The study contained analysis of not only the critical junctures *per se*, but also the downstream increasing return, i.e. the delay in United States in establishing national health insurance at an early period of health system development created interests that frustrated the agenda of establishing national health insurance. Hacker also attended to political structures and argued that the different nature of federal structures in the US and Canada played a key role in the timing and extent of Medicare. Hacker observed that all these factors matter and that political institutions are not unchangeable and have varied impacts on policy development. As Hacker showed, it is the sequence of key historical events that really has determined the influence of all these factors and the outcome of political processes.

The other example is the study by Tuohy (1999) on the effects of the 1990s policy period on the health systems of the same three countries. Tuohy proposed a dichotomy of “policy episodes” (similar to conjunctures) and “system logics” (see Type 4 in Figure 2-3), which indicated the discrepancy between the world of policy and the world of health systems. According to Tuohy’s construction, big changes of policies were only likely during short policy episodes opened by external forces. After the

policy episodes, the system logic plays out, which may or may not lead to major/intended changes in the health system. Tuohy's framework highlighted the difference between the world of policy-making and the world of implementation. Particularly, Tuohy's study demonstrated that even with major changes in policy (in the case of introduction of an "internal market" in the National Health Service), systems could be rather resilient and reduce the scale of impact of the changed policies in implementation. In short, the findings from studies on system changes pre-juncture and post-juncture reaffirm the common knowledge in health system and policy research that systems changes are somewhat independent from policy change, or it can be said that effects of policy change are bounded by entrenched factors within health systems.

A potential limitation of reinforcing-path-dependence theory (including critical junctures) is its lack of appreciation for endogenous changes—the possible windows of change are relegated to special exogenous shocks. Besides juncture points, other theories about institutional changes propose mechanisms of change. These include the role of previously marginal groups, the significance of overlapping processes, and the role of institutional entrepreneurs. The literature on institutional change also identifies a range of typical processes of institutional changes: layering, institutional conversion, diffusion (Pierson, 2004, pp. 137-138). Problems of such mechanisms are that they have to defy the resilience of institutions to the degree they have been established. While attention to these "catalysts" of change are important, Pierson suggested that the focus should be on the resilience (ability to survive policy changes) of institutions and the long term process of their build up (Pierson, 2004: Chapter 5). The inclusion of reinforcing and reactive sequences and in the same theoretical framework may potentially address the issue of lack of appreciation of endogeneous changes. Essentially, path dependence analysis allows policy analysis to connect history to contemporary policy studies not just by applying social science theories to empirical data to explain

history, nor by methodologically providing available or additional cases for developing theories. Instead, as Pierson (2004, pp. 4-5) pointed out, the use of history is theoretical and adds “a distinctly temporal dimension” to the analysis of institutions and institutional changes.

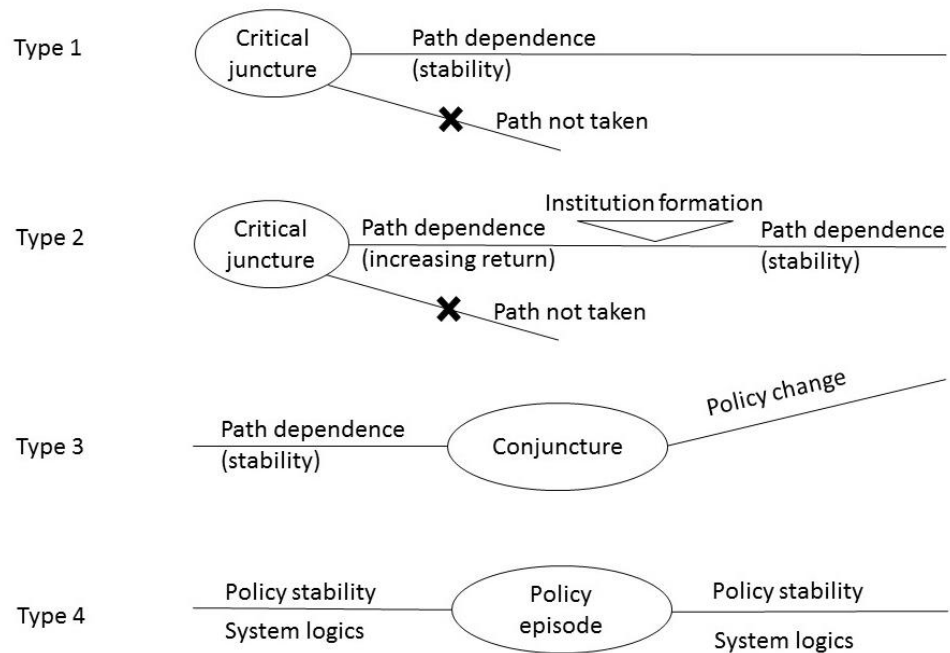


Figure 2-3. Main types of path dependence analysis identified in the literature review

3) Analytical methods

Path dependence and critical juncture have been invoked by researchers in a vague sense when it comes to the analysis of historical processes. Recognizing this issue, Capoccia (2015) developed a multi-step analytical method to systematically study critical juncture (summarised in Table 2-7). This approach offers a sound basis for developing a rigorous approach for path dependence analysis in this thesis.

Table 2-7. A multi-step analytical method for critical juncture analysis

Step	Content
1	Clarify the unit of analysis

2	Identify “candidate” critical junctures
3	Test for structural effects
4	Select for analysis the most “critical” critical junctures
5	Reconstruct the political process
6	Counterfactual analysis
7	Demonstrate lasting legacy of the institutions selected
8	Cross-sectional comparison of critical juncture (at all steps if possible)

Source: summarized from Capoccia (2015)

First, by clarifying unit of analysis, critical juncture analysis can focus clearly on one institution or a set of institutions, rather than using critical juncture in a vague sense to indicate change. Second, by identifying candidates for selection, the approach also seeks to avoid bias in selection. Structural effects might rule out the possibility for alternatives, either by eliminating all but one option, or by selecting the actors and forcing them to adopt just one choice. According to some definitions (Capoccia, 2015), critical juncture should allow contingency (or a plausible alternative to the institutional outcome)—otherwise it would not be critical. Therefore, testing for structural effects (Step 3) would help narrow the range of candidate critical junctures. In the same vein, Step 4 further narrows the range to a very small number of the most “critical” candidates of critical junctures. Two criteria have been specified for judging criticalness: the probability that the institution acquires durable path-dependent characteristics by the end of the critical juncture; the relatively short duration of the critical juncture compared to the duration of its path-dependent legacy.

As a main part of the analysis, Step 5 reconstructs the political dynamics during the selected critical junctures. This should be done by identifying the main actors and their interaction, reconstructing politically viable alternatives for institutional outcomes and the forces behind each options, identifying which decisions are most influential in forming the final intuitional legacy. Steps 6 to 8 include counterfactual analysis, demonstration of institutional legacy and cross-sectional comparison of critical junctures, which are all mainly about reconstructing the criticalness of the

critical junctures selected by comparing the path-dependent institutional outcome with the “the paths not taken”. For these steps, Capoccia (2015) emphasised the importance of negative cases (critical junctures where changes failed to happen).

4) Relevance for China’s context

The literature of historical institutionalism has developed based on experiences as well as concepts derived from Western societies. Just before concluding the literature review, it is worthwhile to ponder about the usefulness or relevance of the theories for the context of China. First, the polity of China has been apparently different from that of the West. As just discussed, the theory of path dependence and the school of historical institutionalism developed largely based on the assumption of fragmentation within an autonomous state. By contrast, political power in China has been centralized, and the country has had only one ruling party after 1949 without competition among different political parties. However, fragmentation of power always existed in the Chinese system including in the period after 1949 (this point will be elaborated on in Chapter 6). Those in power (or with access to power) had their own preferences, which tended to favour certain societal groups more than others. Preferences of certain policy actors do not necessarily agree with those of others.

Another challenge to applying path dependence analysis to China is the degree of changes in the country as compared to the West. Theories from historical institutionalism like path dependence originated with the need to explain the observation that changes rarely happened in some countries, and then were extended to reveal mechanisms for both entrenchment of institutions and changes. Therefore, such theories (particularly when it comes to critical juncture analysis) were not limited to Western societies. That said, the parameters of the institutional changes, i.e. political and societal contexts, were generally stable. Although massive social changes took place in the West, which might be considered as critical junctures, including major

events like the two world wars, civil rights movement and so on, such changes seem less disruptive as compared to the breadth and depth of social changes in China.

A review of studies on organisational and institutional change about China (Zhou, 1999) concluded that there were generally two main tendencies in this field. Some researchers tried to explain the situation in China with existing Western theories. Their efforts tended to generate some “awkward” explanations, while also produced some novel insights. Other studies were based on the background and conditions of China. Such studies led to analysis with greater depth and appropriate analysis, but were generally limited in the power and range of theoretical explanation. Research that combined the advantages of both groups were exceptional. On the other hand, Zhou’s review also argued that social institutional infrastructure (socialist ideology, cultural tradition, pluralistic ownership) and complicated reform paths (incremental, multi-layer, multi-aspects) of China were unique, and provided important opportunity for theoretical innovation.

Studies that explicitly used historical institutionalism for China suggested different results: that its high degree of authoritarianism overruled path dependence--in the case of health insurance reform in Shanghai (Luk, 2014); and that historical institutionalism worked well to explain important parts of the history of health policy in China--a significant element of continuity existed in health service delivery after the launch of major economic reform in 1978 (Huang, 2000), and somewhat contradictorily, in the case of the different fates of continuity of urban risk pooling and collapse of rural health risk pooling over the same period of economic reform (Duckett, 2012). Lessons from these previous studies seem to be twofold. There are risks of projecting Western theoretical propositions on China’s potentially different realities, but also potential value of building upon the Western-origin theories to explain a phenomenon in a varied context with some common denominators. While using theories and methods derived from Western societies, research on China should

perhaps focus on the areas where parallels exist, and maintain enough sensitivity to the differences and anticipate the variations caused by such differences.

2.7 Conclusions

This final section summarizes the key points reviewed and highlights the gaps in knowledge.

First, this chapter has reviewed the concepts related to the balance between primary and hospital care, as well as primary care. The two concepts were found to be strongly related to each other. Both concepts have multiple interpretations, and this thesis adopts a systems-level definition for both.

Second, there is a lack of a systematic measurement tool for the balance between hospital and primary care providers. Existing measurements generally fall into two groups: structural measurements and functional measurements.

Thirdly, evidence regarding effectiveness of gatekeeping in changing the balance between primary and hospital care is limited overall and concentrated in high-income countries. Studies on the effectiveness of gatekeeping and other policies to shift the balance between hospitals and primary care providers in China have been rare and of low quality. Gatekeeping involves various dimensions of complexity that requires more than a “simple” evaluation on average difference between outcomes.

Finally, the review has identified three particularly promising analytical approaches and the gaps they can fill: impact evaluation to study the effectiveness of intervention, qualitative systems analysis to understand the intervention in a system, and path dependence analysis to understand the institutional constraints. These approaches are yet to be applied to the research question raised in the thesis, though their application in related fields suggests potentially significant value.

Chapter 3 Aims, objectives and overview of methods

3.1 Aims and objectives

The primary aim of this thesis is to understand how to strengthen primary care in China, by looking at the effectiveness and functioning of a gatekeeping pilot, and more broadly at the history of the balance between primary and hospital care, in order to contribute to policy development in China as well as other similar settings. Through applying a mix of health system research approaches (trend analysis, historical analysis, qualitative systems analysis and impact evaluation), a secondary aim of the thesis is to explore how mixed methods can be used to study a complex system-related issue.

Building upon the findings from the literature review, the thesis developed the research aims into six objectives (see Table 3-1). The ordering of objectives follows a temporal logic, which seemed best suited to develop the argument and analysis through the thesis. Objective 1 and Objective 2 concern the history of the balance between primary and hospital care; Objectives 3 and 4 focus on a contemporary gatekeeping pilot, which was a pioneering gatekeeping pilot in China available for study. Objective 2 and Objective 4 also provide qualitative explanations for the quantitative analysis in Objective 1 and Objective 3, respectively. Objectives 5 and 6 draw from the four sub-studies to discuss methodological, substantive and policy implications.

Table 3-1. Objectives and research approaches

No.	Objectives	Research approaches	Main corresponding chapters
1	to describe the historical trend of the balance between hospitals and primary care providers in China	Trend analysis	Chapter 4

2	to understand the historical logic of the coevolution of hospitals and primary care providers in China	Historical institutionalism	Chapter 5 and Chapter 6
3	to evaluate the effectiveness of a gatekeeping pilot in strengthening primary care functioning	Impact evaluation	Chapter 7
4	to understand how gatekeeping functions in the pilot in the context of China's recent reforms	Qualitative systems analysis	Chapter 8
5	to discuss the overall findings of the four sub-studies	Drawing on 1-4	Chapter 9
6	to identify lessons for the policies on shifting the balance between primary and hospital care in China and elsewhere	Drawing on 1-5	Chapter 10

3.2 Overview of methods

The section provides an overview of the research approaches and methods used in the thesis to address each objective, while full methodological details are provided in corresponding chapters or parts. Each objective is dealt with using a specific research approach, except for the last two objectives (methodological reflection and policy implications) which draw from all other sub-studies (see Table 3-1).

3.2.1 The overall study design

Figure 3-1 is a graphical summary of how the four sub-studies constituted a multi-approach study on the balance between primary and hospital care. The thesis addresses the conceptual complexity of the balance discussed in Chapter 2 by dividing the issues involved into three levels (shown on the left side of the diagram). The three levels of issues involved in shifting the balance between primary and hospital care include: an identifiable policy intervention (represented as the thinner oval); a contemporary system (represented as the wider oval) involving a broad range of inter-connected health systems factors directly or indirectly related to the intervention, the interaction among which can be captured in a cross-section; and a historical system (represented as the shape of a large arrow, with a shaded arrow below representing the underlying

axis of time), where temporal sequence is critical in setting a long term dynamic. The two ovals and the large arrow overlap and are linked with circular arrows in the diagram, reflecting the embeddedness of intervention in the contemporary system and the contemporary system in the historical system. The diagram represented the recognition that the three levels are closely interrelated and need to be investigated together to form a comprehensive assessment of the situation and future strategies.

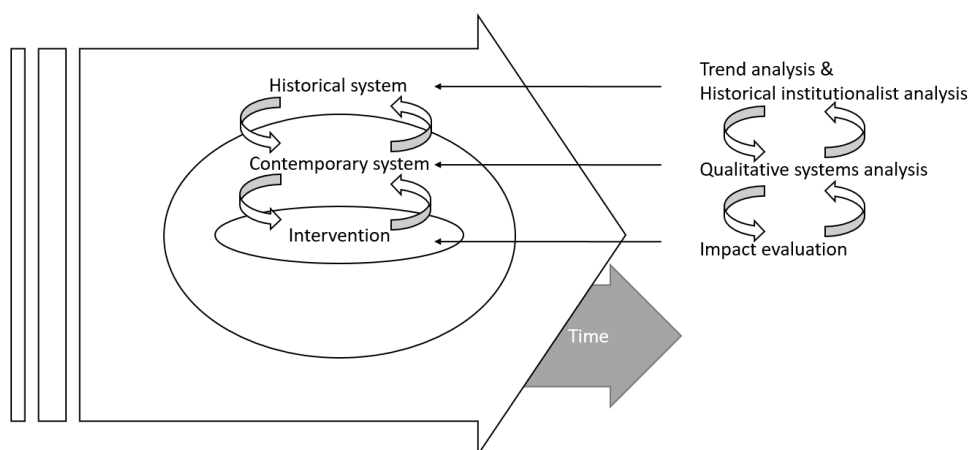


Figure 3-1. Summary of overall study design

Having conceptualized shifting the balance between primary and hospital care as involving three levels, the thesis relied on four sub-studies to address the three levels (shown on the right side of the diagram), and two “integrative analyses” to synthesize findings of all sub-studies. The first sub-study (trend analysis) measured the changing structural and functional aspects of the balance to provide a background for the historical-institutionalist analysis. The second sub-study (historical-institutionalist analysis) tried to explain the historical evolution of the balance, drawing from the theoretical lens of path dependence theory, and treating factors involved in the coevolution of primary and hospital care over the long term as a historical system. The third sub-study (impact evaluation) investigated the effectiveness of the gatekeeping

pilot, seen as an intervention. The fourth sub-study (qualitative systems analysis) dealt with the functioning of the gatekeeping pilot, seen as a contemporary system. While the four sub-studies were conducted in parallel, the remaining two integrative analyses drew on the findings of all four sub-studies, focusing on their interconnectedness (represented by circular arrows) and reflected on the methodological significance, substantive findings and policy implications.

3.2.2 Describing historical trends of the balance between primary and hospital care

To address Objective 1, a descriptive study assessed the historical trend of the balance between primary and hospital care in China. Due to data availability issues, the period studied was limited to the years from 1949 to 2013. As a systematic way of measuring the balance between primary and hospital care did not exist, the sub-study built on the findings of the literature review in the previous chapter, and developed a system of metrics, involving 11 indicators, to measure the structural and functional domains of the balance. The structural domain included indicators on human resources, assets, equipment and infrastructure. The functional domain include indicators that measure service utilisation and revenue. Data for both hospitals and primary care facilities were sourced from official national health statistics, including a statistical report that covered the 40 years between 1949 and 1988, the series of national health yearbooks that covered the years between 1949 and 2008, and the series of national health statistical yearbooks that covered the year between 1949 and 2013.

3.2.3 Analysing the historical logic of evolution of the balance between hospitals and primary care providers

To address the second objective of understanding the historical logic of the coevolution of primary and hospital care in China, a historical-institutionalist sub-study was conducted. The sub-study covered the years from 1835 to 2013, divided into

three periods segmented by 1949 and 1978. For the analysis, a dynamic path dependence analytical framework was developed, building on the critical juncture analysis approach (Capoccia, 2015). Data about main actors were collected from an extensive range of sources, including: relevant historical studies/accounts; official documentations of history; local archives; anthologies, biographies, and memoirs; journals and newspapers; and three propagandist paintings.

3.2.4 Impact evaluation of the gatekeeping pilot

To address Objective 3—evaluating the effectiveness of a pioneering gatekeeping pilot in shifting the balance between primary and hospital care, a difference-in-differences study design was applied. This impact evaluation used the pilot townships as a treatment group and non-pilot townships as a control group and compared the different trends of key outcome variables both before and after the start of the pilot and across intervention and control areas. Claims and enrolment data were used in econometric analysis. The study covered a population of around 21,000 in 17 townships, in the 12 quarters from 2012 to 2014. The sub-study was limited by the small scale: two small rural townships in a vast and varied country.

3.2.5 Systems analysis of the functioning of the gatekeeping pilot

To address Objective 4—understanding the functioning of the gatekeeping pilot, the thesis developed and used an approach of qualitative systems analysis, which combined a categorisation of health system building blocks, a qualitative method for policy analysis and causal loop analysis. Over two years (2014 and 2015) during the gatekeeping pilot programme, 20 stakeholders (including patients, doctors, facility managers and health administrators) were interviewed. The health systems categorisation was used to develop question guides for the interviews. The “Framework” approach of policy analysis was used to analyse the data from the field work. A causal loop diagram was drawn and used to assist the analysis and present the

results. Again, the sub-study was limited by the small scale of implementation.

3.2.6 Ethics

Ethical approval for this study (impact evaluation and qualitative systems analysis) was provided by the Institutional Review Board for Biomedical Ethics at Peking University (IRB00001052-14024-MIAN) and the Research Ethics Committee at the London School of Hygiene & Tropical Medicine (PR/123/123).

3.3 Summary

This chapter has laid out the aims and objectives, as well as the methodological framework of the thesis. The study design included four sub-studies, which constituted a contemporary evaluation of a pioneering gatekeeping pilot and a historical study of the development of the balance between hospitals and primary care providers in China. Chapter 9 and 10 will revisit methodological issues of this study design.

Part II History

Chapter 4 Historical trends (1949-2013)

4.1 Introduction

The main objective of this chapter is to describe the historical trend of the balance between hospitals and primary care providers in China using a comprehensive system of metrics. As data before 1949 were neither routinely nor consistently collected and reported, the chapter is limited to the years between 1949 and 2013. Section 4.2 presents a brief introduction to the historical and contemporary health delivery system in China. Section 4.3 lays out the rationale and process for selecting indicators, the nature of the data, and the justification and data availability for each indicator. Section 4.4 describes the trend regarding each indicator. Section 4.5 discusses the methods and results. The final section concludes with key findings as well as their implications.

4.2 Background about the health delivery system in China

This section provides a brief background regarding the history and current situation of the health delivery system in China, mainly since 1949 but also with brief reference to earlier periods. Figure 4-1 is an overview of the main types of health facility in 2013, which provides a basis to explain the evolving structure over history. It is noticeable that almost each level of government had health facilities under its direct supervision, and these were classified by the government as facilities “under the health sector” as reflected in government statistical reports (Ministry of Health, 1989; National Health and Family Planning Commission, 2014). In other words, local territorial governments rather than a centralized health ministry ran health facilities. Besides, industrial sectors, collectives and other sectors also ran their own facilities

(classified as non-health sectors). The effects of such a structure will be revisited in later chapters.

The health service providers consisted of the public health delivery system (shown with green background) and the medical care delivery system (shown with blue background). The two sub-systems converged at the primary care level, with primary care providers providing a mixture of public health and medical care services. Within the public health delivery system, there was a national system of centres for disease prevention and control (CDCs). The local CDCs used to be epidemic control stations and were restructured in the outbreak of severe acute respiratory syndrome (SARS) in 2002, when the country's Academy of Preventive Medicine became the national CDC (Centre for Disease Prevention and Control, 2012). Correspondingly, there was a nation-wide system of maternal and child health centres (MCH centres), except that the national centre for maternal and child health was a department under the national CDC. The precursors of CDCs and MCH centres were established nation-wide by the 1950s. Some of these public health facilities might also provide specialized medical services (birth delivery, treatment for tuberculosis, etc.) and might have a simultaneous title of hospitals and are reported here as hospitals in such circumstances. In the post-1949 Chinese health delivery system, Chinese medicine (i.e. traditional Chinese medicine) and biomedicine (i.e. Western medicine) are integrated at the facility level.

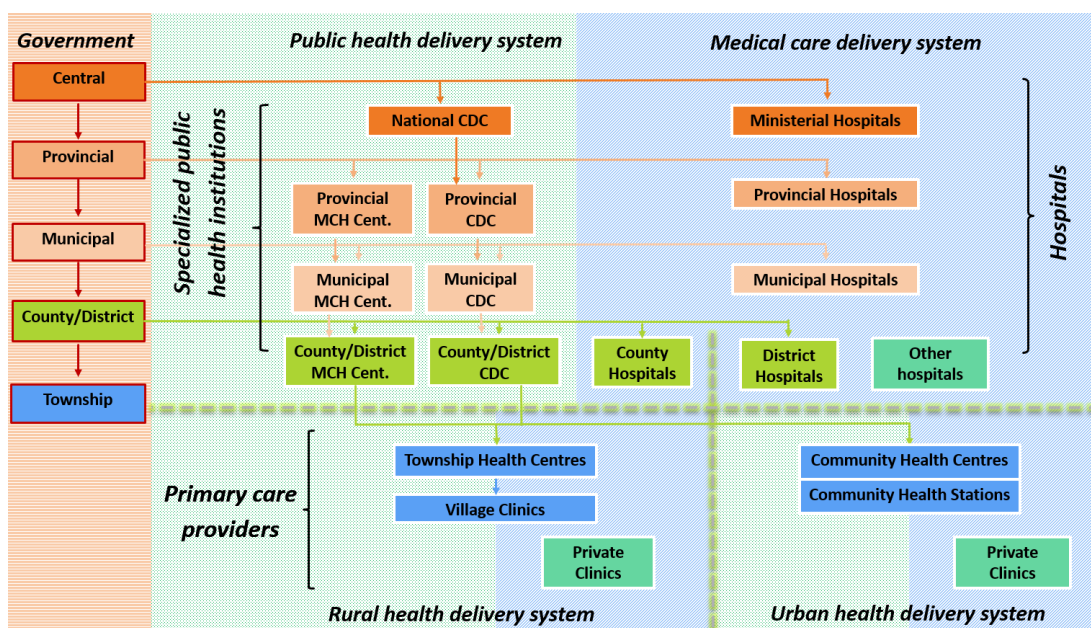


Figure 4-1. Main types of health service providers in China in 2013

Note:

- a) CDC=Centre for disease prevention and control;
- b) MCH=Maternal and child health.
- c) The horizontal dash line represents the division between hospitals and primary care facilities.
- d) The vertical dashed line represents the division between rural and urban medical care delivery.
- e) Sub-district health centres were not shown in the graph due to limited space and their small role.
- f) The areas with blue slashed shading represent the medical care delivery system; the areas with the green dotted shading represent the public health delivery system; and the area with orange horizontal line shading represents the government hierarchy.

The medical care delivery system in 2013 consisted of various levels of hospitals and primary care facilities. The categorical division between hospitals and primary care facilities here is based mainly on the current categorization of facilities by policy-makers and official statistical reports (National Health and Family Planning Commission, 2014). The root of this distinction, however, goes back for at least eight decades. As Chapter 5 will further describe, a nation-wide county-based regional health delivery network had been advocated as a national policy around the 1930s to provide wide coverage of health services. The structure was fleshed out and evolved over the decades, while the centre of this structure in each region was always a county

hospital⁵. Two levels of facilities extended resources and services beyond the county-seat and reached out to sub-county territorial units, i.e. townships/communes and villages. The medical services facilities at the three levels of county, township and village, along with the county MCH centres and CDCs, formed the rural health delivery system. A parallel urban system consisted of district hospitals as counterparts of county hospitals, and community health centres and stations as counterparts of township health centres and village clinics. The urban system also included three higher territorial levels of hospitals (i.e. municipal, provincial and ministerial), and other hospitals beyond the health sector. This organisational division between hospitals at county level (and above) and primary care at sub-county levels was established in the 1950s in official policy documents that specified sub-county/sub-district facilities as “grassroots health organisations” (*jiceng weisheng zuzhi* 基层卫生组织) (Ministry of Health, 1957a). Categorization of health facilities including hospitals and primary care facilities was done in the statistical reports according to the types of facility registered (Ministry of Health, 1994b; National Health and Family Planning Commission, 2014). In reality, the two categories of facilities provided an overlapping range of services. This will be a repeating theme of this and following chapters. Figure 4-2 provides the number of hospitals and primary care facilities in China between 1949 and 2013.

⁵ In fact, county hospitals were initially established as county health centres, which incorporated public health services and administration in a county-level regional referral hospital. The county health centres later split into county hospitals and county epidemic prevention stations (Ministry of Health, 1956a).

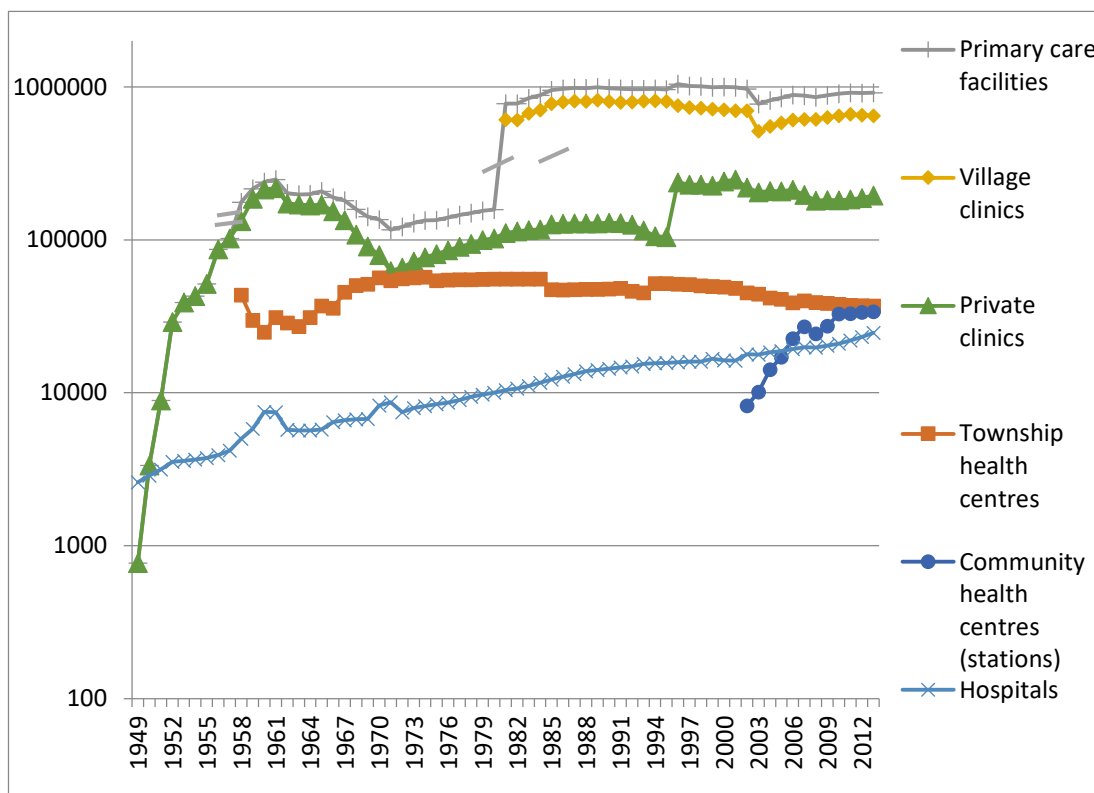


Figure 4-2. Number of hospitals and primary care facilities in China (1949-2013)

Data sources: (Ministry of Health, 1989; National Health and Family Planning Commission, 2014)

Note:

- Number of primary care facilities before 2000 was calculated by summing up numbers of all types of primary care providers with available data. The number of village clinics was not reported until 1981.
- A log scale is used for the Y axis in this diagram.

4.2.1 Hospitals

After 1949, the state took over mission hospitals and built new public hospitals; industrial and other sectors (mainly factories), collectives and individuals also set up their own hospitals. The meaning of hospitals in official statistical reports also changed over the years. Statistical reports about hospitals published up to 1997 (including data up to 1996) included township and sub-district health centres; after 1997, township and sub-district health centres were separately reported (Ministry of Health, 1996, 1997). The analysis here either used later sources that adopted the post-1997 classification, or used data for only the hospitals at or above the level of county if

reclassified data were not available.

Since the 1990s, hospitals have been classified, according to their technical capacity, into hospitals at primary level, secondary level, tertiary level and unassigned level. Starting from 2,600 in 1949, the number of hospitals increased steadily over the decades and reached 24,709 in 2013 (see Figure 4-2). Among all hospitals, the tertiary and secondary hospitals are the most important, as they jointly provided more than 80% of all hospitalisations in 2013. In the same year, there were 1,787 tertiary hospitals, generally consisting of ministerial, provincial and municipal hospitals, and teaching hospitals of medical colleges. There were 6,709 secondary hospitals, mainly consisting of county and district level hospitals (including traditional Chinese medicine hospitals). Although the hospitals (particularly the secondary and tertiary hospitals) were supposed to be referral hospitals that received only complicated cases, patients generally had direct access to hospitals. Hospitals also provide a significant amount of ambulatory services (a point elaborated in later sections).

4.2.2 Primary care providers

The primary care facilities/providers in China have always been a pluralistic group of providers. In recent national health statistical reports (from 2010), the term “grassroots health care facilities” (*jiceng yiliao weisheng jigou* 基层医疗卫生机构) was used for the first time in national health (statistics) yearbooks to describe the categories of primary care providers (Ministry of Health, 2011). Those primary care facilities under government sponsorship are “grassroots *public* health care institutions” (*jiceng yiliao weisheng shiyan danwei* 基层医疗卫生事业单位) (Ministry of Human Resources and Social Security, 2009). These public facilities have been the main target of government policies on primary care and are spread around the urban and rural communities of the country. The study here used the contemporary definition of the term primary care facilities to refer to the types of provider that have been

retrospectively identified as primary care providers.

In 2013, the primary care facilities included township health centres and village clinics in rural areas, community health service centres and stations in urban areas, sub-district health centres, as well as other clinics (as these clinics were mostly private, they are referred to in the chapter as private clinics). Being the foundation for both the public health delivery system and the medical care delivery system, they provided the community-level work related to health education, preventive services, rehabilitation, family planning and treatment of common and prevalent illnesses (including a significant amount of inpatient services) (National Health and Family Planning Commission, 2014).

The reported number of primary care facilities grew from 769 in 1949 to 0.9 million in 2013. As shown in Figure 4-2, the increase was driven first by the rising number of clinics in the 1950s along with the organization of “united clinics” and then the inclusion of village clinics in 1985, when the number of primary care facilities reached almost 0.8 million. A nation-wide network of rural facilities similar to that of today came into being in the late 1950s with the emergence of commune health centres, and in the late 1960s with the popularization of barefoot doctors. In 1958, there were 43,579 commune health centres (shown as township health centres in Figure 4-2 and the remainder of the text and all graphic representations in the chapter), in addition to 132,894 private clinics (including united clinics and solo practices). After a decade-long fluctuation, the number of commune health centres reached 50,341 in 1968. Commune health centres were reorganized as township health centres soon after 1978, and their number remained generally stable until 2001. Since then, the number of township health centres has declined and reached 37 thousand in 2013 (Ministry of Health, 1989; National Health and Family Planning Commission, 2014). In regard to village clinics, data were not available before 1985. Their number dropped from 0.8 million in 1985 to 0.6 million in 2013 (Ministry of Health, 1989; National Health and

Family Planning Commission, 2014).

Urban primary care facilities consisted of private or united clinics immediately after 1949, as mentioned above. Sub-district health centres (not shown in Figure 4-1) were formed during the 1950s (Ministry of Health, 1989). The number of sub-district health centres was small and increased from 446 in 1996, when first reported, to 593 in 2013. Community health centres and stations came into being around the beginning of the 21st Century (Bhattacharyya et al., 2011). Community health centres were the urban equivalent of township health centres, and community health stations were generally outreach facilities of community health centres. The total number of community health centres and stations was 8,211 in 2002 (when the number was first reported) and reached 33,965 in 2013.

4.3 Methods

Metrics involving multiple indicators were developed and used to measure the balance of care in China between 1949 and 2013. This section explains the overall rationale and process of selecting and categorizing these indicators, the sources and nature of data, and finally the justification for all indicators and operational decisions.

4.3.1 Overall rationale for developing the metrics

Based on the findings from the literature review and conceptual discussions in Chapter 2, the two domains of functional and structural balance between primary and hospital care were established. Under each domain, sub-domains were developed, with corresponding indicators chosen for each sub-domain. The structural domain concerned with human capital and physical capital, each comprising an independent sub-domain. The functional domain concerned with service utilisation and financing, each comprising an independent sub-domain. The criteria for selecting indicators were that the indicators could reflect the nature of sub-domains and were available for

sufficiently long periods to allow meaningful analysis over time. The process of selection was iterative and resulted in multiple indicators under each of four sub-domains shown in Table 4-1. The table also shows the sources and years for which data were available.

Before getting into the details, one has to admit that no systematic measurement of the balance between hospitals and primary care providers over a long period of time can be satisfactory. First, the boundaries between hospitals and primary care providers were not always straightforward. This has been reflected in the literature review (in Section 2.3) on the various and overlapping definitions of primary care. The metrics used a level of care as the unit of analysis, zooming out and seeing primary care facilities and hospitals each as a whole, rather than zooming in and focusing on the margins of care. However, there could still be multiple approaches to operationalize the boundaries between the two levels of care. What the analysis needed to achieve was a feasible measurement with policy implications.

Second, besides the issue of debatable boundaries, both primary care providers and hospitals involved a complex range of facilities. By categorizing these facilities into two sectors the variation within each category was lost. Furthermore, the value of each indicator would also be affected by the changing availability of data. Hence, the analysis included data on sub-groups of facilities when possible to allow greater flexibility and to capture some important details.

Third, in relation to the debatable boundaries between hospitals and primary care providers, there was an inherent challenge related to the coexistence of similarities and differences between hospitals and primary care providers in the context of China, in terms of both the resources they utilised and the services they delivered. These similarities and differences challenged quantification of various dimensions of the balance between hospitals and primary care providers. As mentioned earlier, primary

care facilities and hospitals provided a significantly overlapping range of services. While primary care facilities provided inpatient services, hospitals provided a large proportion of ambulatory care. On the other hand, there were also quite substantial differences between the two sectors. Hospitals provided to a greater extent inpatient services and used technologies (including devices) much more intensively. Even for similar categories of services, medical treatment provided in different types of facilities was of a potentially very different nature. While some indicators may be less specific to one type of facility (e.g. financial resources), others would unavoidably be more so (e.g. beds), owing to the different nature of facilities at different levels of care. However, without a detailed look at these less generic indicators across different levels of care, one would lose the ability to appreciate fully the various details and dimensions of the balance between them.

Fourth, the temporal dimension also influences both the mixture of resources and functions and the composition of facilities themselves. While developing the indicators systematically, the study also tried to reflect the issue of the balance between hospitals and primary care providers comprehensively in the historical context of China. When considered appropriate, the study used multiple indicators for one sub-domain, as well as multiple ways of measuring one indicator occasionally, though this did not solve all issues. However, simply making the measurement more flexible and complicated could not solve all the issues. While a marginal analysis might generate useful information for short-term policy modification and resource planning, it would not allow the analysis of overall structure and functional balance needed to understand an entire level of care. As the margins changed all the time, a narrow focus on marginal areas would fragment the analysis and make it difficult if not impossible to appreciate the long-term evolution of the health delivery system. A further problem about measuring the balance between the primary care providers and hospitals over a long time was the lack of continuous reporting of data, particularly in the earlier periods.

Fifth, the diversity of China means that any measurement at national level has little to say about local situations. The huge variations between coastal and western landlocked areas and between urban and rural areas would be inevitably missed in measurement at the national level.

The next sub-section concerns the methodological issues related to the data used in this chapter.

4.3.2 Data

All data used in this chapter came from official sources and were at national level. Data sources consisted of *Health Statistics information in China 1949-1988*, the statistical sections of *China Health Yearbooks* (published annually from 1983 to 2008—when the collection of the series ended), *China Health Statistical Yearbooks* (published annually from 2003 to 2012), and *China Health and Family Planning Statistical Yearbooks* (published in 2013 and 2014). All these sources were published by the Ministry of Health (National Health and Family Planning Commission). As the data were officially collected and reported, the way of reporting seemed to have been influenced by the historical changes of policies (and the emphases of policies during particular periods). The changes affected the availability of data and influenced the selection of indicators as well as reporting of their values. Two issues were particularly of concern.

One issue concerned the changes in the items of data reported. This sub-study covered 64 years of dynamic changes in China's health system. The Ministry of Health (and the later National Health and Family Planning Commission) did not collect and report all indicators for all years. For instance, financial information of health facilities was rarely reported (and in an inconsistent way when reported at all) before 2004. When necessary, substitute indicators for earlier periods were used along with necessary cut-offs between periods to make comparisons meaningful.

Another issue concerned the selection of types of facilities for reporting under primary care facilities. This issue was related to the pluralistic nature of health facilities in China mentioned above, as well as changes in categorization and reporting. As Section 4.2 has mentioned, the official term for primary care providers did not appear in national statistical reports until well into the first decade of the 21st Century. Therefore, the study here reconstructed the total number of primary care facilities based on reports about a range of providers that were retrospectively identified as primary care providers. As not all items of statistics were available for all types of facilities in all years, both the total of primary care facilities and their breakdown provide useful information and were included as much as possible. As the total information in earlier years were constructed by summing up available types of primary care facilities, a break sign (“//”) is added to the series with changed composition of types of facilities and indicate where the changes of composition take place. These issues are revisited in the interpretation of results in Section 4.4.

A third issue was that indicators need to reflect the development of the quality of health system resources. For example, the indicators of quality of human resources needed to reflect the significant improvement of educational level between 1949 and 2013.

The original data came in various forms: paper-based books, electronic books, on-line documents, scanned images, and Excel spreadsheets. Data were entered into a Microsoft 2013 Excel document and checked by the author with assistance from two Chinese master students in public health at Peking University. When data were available for the indicators for the specific types of facilities, they were directly harvested. However, as reporting about certain categories of facilities changed, original data of separated items were aggregated. Simulated values were used on one occasion, which will be noted in the results section.

4.3.3 Justification for indicators and their operational definitions

Table 4-1 provides the list of all the indicators selected. The table also shows the sources and years for which data were available.

Table 4-1. Indicators of structural and functional balance between primary and hospital care

Domain	Sub-domain	Indicator	Sources of data	Available years
Structure	Human resources	1.1 Number of staff	40YR, CHY (1990-2002), CHSY (2003-2014)	1949-2013
		1.2 Qualification of health practitioners	40YR, CHY (1990-2002), CHSY (2003-2014)	1952-1988, 2002-2013
		1.3 Average annual income of staff	40YR, CHY (1990-2002), CHSY (2003-2014)	2005-2013
	Assets, equipment and infrastructure	2.1 Value of assets	CHSY (2005-2014)	2004-2013
		2.2 Number of beds	40YR, CHSY (2003-2014)	1949-2013
		2.3 Number of valuable devices	CHSY (2004-2014)	2003-2013
		2.4 Building space	CHY 1991, CHSY (2004-2014)	1989, 2002-2013
Function	Service utilisation	3.1 Number of visits	40YR, CHSY (2003-2014)	1981-2013
		3.2 Number of admissions	40YR, CHSY (2003-2014)	1981-2013
		3.3 Visits to admissions ratio	40YR, CHSY (2003-2014)	1981-2013
	Financing	4.1 Revenues	CHY (1997-2004), CHSY (2004-2014)	2004-2013

Notes: 40YR = Health Statistics Information in China (1949-1988); CHY = Chinese Health Yearbook; CHSY = Chinese Health Statistics Yearbook

1) Indicators for human resources

Human resources are a critical factor within the structural domain and featured prominently in previous studies (see Chapter 2). Under the sub-domain of “human resources”, indicators were selected to reflect the quantity and quality of staff.

- Indicator 1.1: Number of staff

Measuring quantity of staff involves the issue of the pluralistic nature of human

resources. While both total number of staff and number of selected types of staff (particularly doctors) provide useful information, data availability for the latter was more limited. Therefore, I decided to use the total number of staff for all types of facilities between 1949 and 2013. Availability of data for certain types of facilities was also restricted (see Figure 4-3). For total staff number of all hospitals, data were available only for the period between 1984 and 2013 and the sum of staff numbers of hospitals above the level of county under the government health sector and county hospitals were used for the years before 1984, implying underestimation of the number of hospital staff. It should be noted that each staff member was counted as one without distinguishing the proportion of time they spend working as health practitioners, as such data were not available. In particular, village health workers (barefoot doctors and health personnel) were supposed to participate in collective agricultural work besides health work. Hence, this measure could give an exaggerated impression regarding certain groups of practitioners who normally worked part time (particularly barefoot doctors).

- Indicator 1.2: Qualification of health practitioners

Measuring quality of staff involves again the issue of the plural nature of human resources, and there are also different ways to measure the quality of health practitioners. First, it was possible to compare both health practitioners of a specific type and health practitioners as a whole. Since comparison of the qualification of staff across the board could be difficult to interpret, it seemed reasonable to confine the analysis to doctors only, as they were the key decision makers in clinical care. Second, it is possible to compare the level of qualification using different criteria. There were two types of qualifications, namely professional qualifications and academic qualifications.

Two systems of professional qualification existed during the past decades in China,

namely, a system of technical qualification (*zhicheng* 职称) and a system of licensing (*zhizhao* 执照). A policy document issued in 1956 clarified health professionals into three main grades of technical qualification. Correspondingly, three professional titles were given for each grade of practitioners: *shi* (师, roughly translated as a “professional”) corresponded to high grade practitioners, *shi* (士, roughly translated as a “practitioner”) corresponded to middle grade practitioners, and *yuan* (员, roughly translated as “personnel”) corresponded to primary (i.e. low) grade practitioners (Ministry of Health, 1956b). The grading system reflected professional competence, generally a combination of academic achievement and work experience. Hence, the proportion of “professionals” was a good indicator of the quality of doctors. This regulation was revised in 1986, when the qualification system changed. From then on, the high grade corresponded to chief (or deputy chief) professionals, the middle grade corresponded to executive professionals, and the primary grade corresponded to (general) professionals and practitioners (Ministry of Health, 1986b). The change probably reflected the improvement in training of health practitioners in the country. However, it also meant that the proportion of “professionals” became less useful as an indicator of quality of staff, because both could be with a lower level of training.

The percentage of professional doctors among all practitioners was thus a useful measurement of quality for the periods between the 1950s and 1980s. However, statistics on human resources were generally reported as either in urban or rural areas (i.e. within urban districts or rural counties), and were not distinguishable by hospitals and primary care facilities except for within rural areas only. As a result, two alternative methods were used for these years. The first method used urban facilities to represent hospitals, and rural facilities to represent primary care facilities. The rationale was that urban hospitals were the predominant health facilities in urban areas in terms of human resources, while primary care facilities dominated the rural areas. The other alternative method looked only at rural hospitals (county hospitals) and rural

primary care facilities. For both methods, data were available for the years between 1952 and 1988. Except for a negligible fraction of professional doctors practising integrated Chinese and Western medicine in the 1980s, professional doctors practising Western medicine were the only professional practitioners (the total of practitioners could be practising Chinese medicine, Western medicine or integrated Chinese and Western medicine). Hence this measurement reflected the quality of doctors mainly from the perspective of biomedicine.

An alternative way to measure professional qualification was related to the licences each level of doctor held. Various types of licences allowed different ranges (types and locations) of services. Available data on licences were restricted to more recent years (2002-2013). Above the level of village, doctors were required after 1992 to have a licence or assistant licence in order to practice (Ministry of Health, 1993).

Besides professional qualifications, there were academic qualifications, i.e. the level of educational qualification (e.g. secondary schools or university degrees). Data regarding educational level were also not available until recently (for selected years between 2002 and 2013). Therefore, the qualification of health practitioners between 1989 and 2013 was operationalized as the percentage of professional grade doctors among all practitioners; percentage of practitioners with full practicing licenses; and percentage of practitioners with university degrees. Unfortunately, data about both percentage of professional doctors and percentage of practitioners with university degrees were unavailable for the years between 1989 and 2001.

- Indicator 1.3: Average annual income of staff

Earnings of primary care professionals as compared with specialists has been used as an indicator for the orientation of a health system towards primary care by Starfield (1994). Such data, however, were not routinely collected. Instead, national statistics available to the study reported the annual spending on staff (combining salary, bonus,

etc.) in various types of health facility between 2004 and 2013. This information could be used together with the information about the number of staff in each type of health facility to calculate average income. The income was adjusted for inflation using the price level of 2000 (see Appendix Table 1).

2) Indicators for assets, equipment and infrastructure

Assets not only are an important indicator of accumulated inputs but also represent the physical capacity of health facilities to provide services. The value of assets provides a representation of the overall non-human capital of a particular type of facility. Such factors have been particularly important in the historical and contemporary balance between primary and hospital care in China. As both primary care facilities and hospitals provided inpatient services, inpatient beds were important assets. Involvement in hospitalization-based services has been considered an important opportunity to improve doctors' clinical experience and competence (Honigsbaum, 1979). Similarly, equipment and infrastructure are important conditions for doctors and other health practitioners to expand clinical service capacity. As facilities generally did not share devices and infrastructure with each other, an important aspect of technical capacity of facilities was the possession of valuable devices (including diagnostic and other devices). As building space is an important measurement of the infrastructure of facilities, and the only one for which data were available, it was selected as an indicator for infrastructure.

These indicators of non-human capital arguably favour hospitals disproportionately and imbalance in such factors is expected. Nevertheless, including non-human capital serves to balance the bias towards primary care when only human resources are concerned. Hence, under the sub-domain of "assets, equipment and infrastructure", four indicators were chosen regarding the value of assets, number of beds, number of valuable devices, and building space. Except for the number of beds,

the data under this sub-domain were available only for the recent decade.

- Indicator 2.1: Value of assets

Data were available for the years between 2004 and 2013 for hospitals and primary care facilities (with a break down consisting of township health centres, community health centres and stations, sub-district health centres and private clinics). The value of assets was adjusted for inflation using the price level of 2000.

- Indicator 2.2: Number of beds

The number of beds were reported for hospitals, township health centres, and community health centres and stations. While data for hospitals were available from 1949 onwards, data for township health centres and community health centres and stations were available only for the period after these types of facility became popular (i.e. from 1959 and from 2002, respectively).

- Indicator 2.3: Number of valuable devices

The total value of all devices above 10,000 *yuan* (price unadjusted) was used as the indicator of possession of high-value devices. Data were available for the period between 2003 and 2013 for hospitals, primary care facilities as a whole, and township health centres. An alternative measure was the number of all devices valued above 10 thousand *yuan* each. However, data for this covered only the years between 2007 and 2013, and the value of each device for inflation could not be adjusted. Therefore, the number of devices with value above 10,000 *yuan* at the price level of corresponding years was used.

- Indicator 2.4: Building space

The working definition for the indicator is the total construction space of all buildings (including rented space) under each type of facility. Data were reported for

hospitals, township health centres, village clinics and community health centres (stations), sub-district health centres, and private clinics, first in 1988 and then generally continuously between 2003 and 2013.

3) Indicators for service utilisation

The sub-domain of service utilisation consisted of three indicators, namely, number of visits, number of admissions, and visits to admissions ratio. Patient visits (consisting overwhelmingly of ambulatory patient visits, but also including emergency care, home visit, health check-up, and health consulting) were a key indicator of services utilized by patients in both hospitals and primary care facilities. Changing proportions and absolute number of visits reflected the change in the functional role of a particular level/type of facilities. Inpatient services had been a significant function of primary care facilities in China as mentioned above. Township health centres, particularly, had been a significant provider of inpatient services. It should be expected though that the nature and severity of inpatient services would be different between hospitals and primary care facilities, but such data were not available. Therefore, the numbers regarding service utilisation were not adjusted for changes in population, as different population groups (particularly rural vs urban residents) appeared to be using different types of facility at varied frequencies.

- Indicator 3.1: Number of visits

Annual patient visits were used to measure number of visits. Data availability for different types of facility varied. While the number of visits was reported for hospitals as early as in 1952, the number of visits to township health centres was not available before 1981. Comparison between 1952 and 1980 was not meaningful and therefore omitted. Visits to village health clinics and to community health centres and stations were reported from only 2003/2004 onwards.

- Indicator 3.2: Number of admissions

The operational definition for the indicator of inpatient services was the number of hospitalisations at hospitals and township health centres from 1981 to 2013. Data for the years before 1981 were dropped from the analysis, because admissions to hospitals only were reported.

- Indicator 3.3: Visits to admissions ratio

The visits to admissions ratio measured the number of ambulatory visits per inpatient for each type of facility. This indicator reflects the overall trend regarding balance of outpatient and inpatient care in hospitals and primary care facilities between 1981 and 2013. Among primary care facilities, data were available only for township health centres between 1981 and 2003. Data for other primary care facilities were included only between 2003 and 2013.

4) Indicators for financing

Chapter 2 suggested that resource allocation was an important measure of the functional balance that has featured in previous studies. Here total revenue of facilities was used as an indicator to reflect the resources utilised for patient services. Revenue also complemented indicators on quantity of services as a better reflection of service function in terms of intensity of services from a financial perspective.

- Indicator 4.1: total revenue

Total revenue measures the annual total amount of revenue of health facilities. Data for total revenue by type of facility were reported for 2004 and 2013.

4.4 Results

4.4.1 Human resources

The study selected and analysed three indicators to reflect the balance between

hospitals and primary care facilities in terms of human resources (see Figure 4-3).

1) Indicator 1.1: Number of staff

As mentioned in Section 4.3, the situation of human resources was complicated over the period from 1949 to 2013, and influenced by inconsistent reporting of data. The analysis here is segmented by 1984/1985, after which data become more available. Over the first three decades or so, the number of staff at hospitals (represented by hospitals under the government health sector, for which data were available) grew steadily from 0.1 million in 1952 to 1.97 million in 1984. Among primary care facilities, the number of staff at township health centres amounted to 0.39 million in 1958 and then decreased to 0.25 million in 1965. When data about the staff of township health centres became available once again in 1972, their number had jumped up by three times, to 0.75 million. Data for staff number of village clinics became available from 1970 with the rise of barefoot doctors. The number of village clinic staff (mainly barefoot doctors around this time) fluctuated around 5 million before decreasing by half to 2.4 million in 1984. Due to the overwhelming number of barefoot doctors, the trend of primary care staff generally followed the trend of village clinic staff from 1970 to 1984.

Since 1984, both hospitals and primary care facilities have seen significant development in the number of staff. The total number of hospital staff increased significantly from 2 million in 1984 to 5.4 million in 2013, by 2.7 times. Amongst these three decades, the growth was slowest in the second decade (1994-2003) and fastest in the last decade (2004-2013). The number of primary care staff rose from a similar starting point at 2.2 million in 1984 to 3.5 million in 2013, by 1.6 times. Growth of total primary care staff number was generally less significant compared to hospitals. Rapid growth of total primary care staff numbers happened only between 2003 and 2010 and with smaller magnitude compared to the growth of hospitals during these

years.

The breakdown of primary care facilities showed a generally stagnant situation of staff numbers at township health centres and private clinics, when compared to hospitals. Staff numbers grew from 1 million in 1984 to 1.2 million in 2013 in township health centres, and remained stable at 0.4 million in private clinics between 2003 and 2013 (data were not available before 2003). The staff numbers of village clinics reduced by half almost overnight, from 2.4 million in 1984 to 1.3 million in 1985. The decline reflected a reclassification of barefoot doctors and health personnel. In 1985, barefoot doctors were required to take an examination, about half who passed were certified as village doctors. The other half became the post-1985 village health personnel. Former (i.e. pre-1985) village health workers were no longer included in the statistics after 1985, though it was not clear how many of them continued to work on village health activities. The number of village clinic staff reduced further to 1 million in 2013. The only types of facility that saw significant growth in staff numbers were the nascent community health centres and stations whose number of staff grew from 0.1 million in 2004 to 0.5 million in 2013, by 5 times. Overall, staff were increasingly concentrated in hospitals, with the proportion of staff in hospitals rising from 48% in 1985 to 60% in 2013.

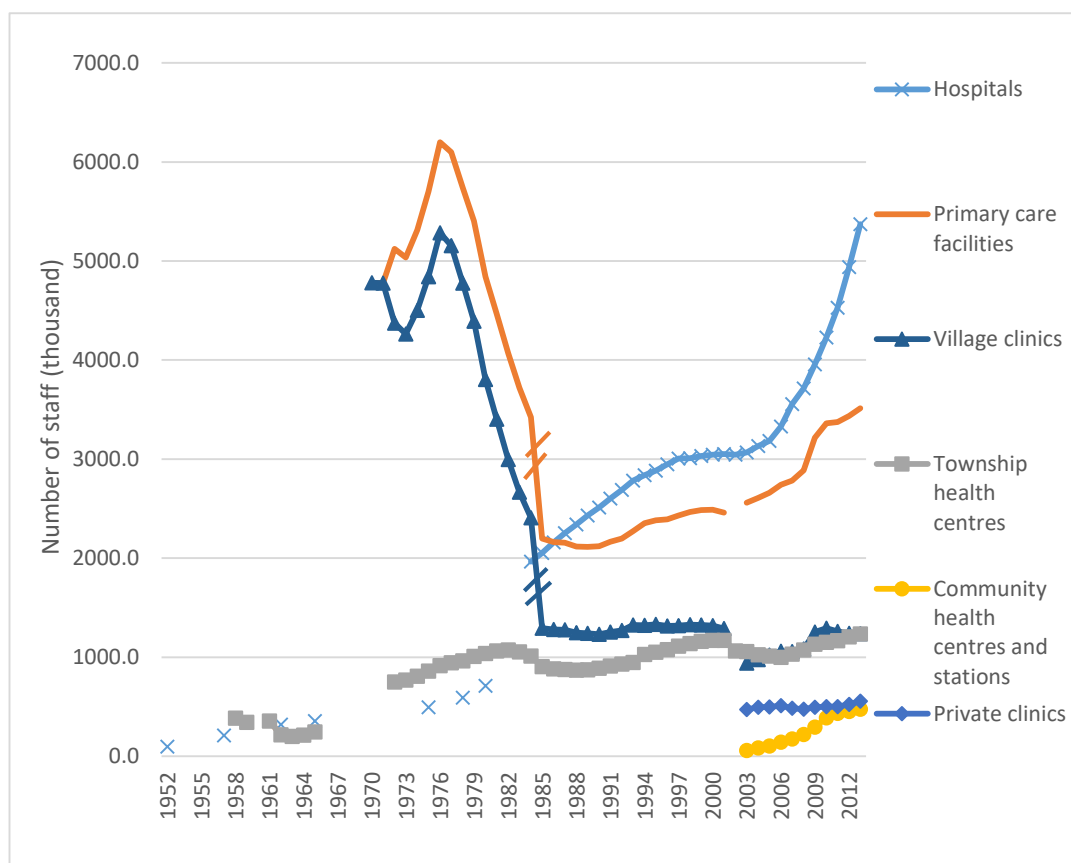


Figure 4-3. Number of staff at hospitals and primary care facilities (1952-2013)

Data sources: (Ministry of Health, 1985a, 1986a, 1987, 1988, 1989, 1990a, 1991, 1992, 1993, 1994a, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012; National Health and Family Planning Commission, 2013, 2014)

Note:

- a) Number of staff at all hospitals in the years between 1952 and 1985 were not available. For these years, sum of numbers for hospitals above the level of county under the government health sector and county hospitals is shown. Hence, hospital staff under industrial and other sectors were not included. The sharp increase of hospital staff in 1985 reflected the inclusion of staff of these hospitals.
- b) Number of staff at all primary care facilities was not available for years before 2011. For these years, sum of numbers for all types of primary care facilities with available data is shown.

2) Indicator 1.2: Qualification of health practitioners

The trend in the qualification of health practitioners revealed long-lasting gaps between hospitals and primary care facilities, which were still widening up to 2013. As mentioned in Section 4.3, data about doctors' qualifications by level of care (or types of health facility) were not available until the 2000s. The sub-study used two alternative measurements for this period. First, the proportion of professional doctors

between urban areas and rural areas was compared, and the number for urban areas was used to represent the hospital sector, while the number for rural areas was used to represent the primary care sector. Then, the proportion of professional doctors between rural hospitals (i.e. county hospitals) and rural primary care facilities was compared.

Figure 4-4 presents professional doctors as a proportion of all practitioners. Professional doctors accounted for 39% of all practitioners in urban facilities and 6% in rural facilities in 1952. The percentages remained at 41% in urban areas and 7% in rural areas by 1962. The next decade saw the percentage increase in urban areas to 50% in 1972, and more rapidly in rural areas rising to 25% in 1972. By 1985, just before the qualification system changed in 1986, percentage of professionals reached 55% in urban areas and 30% in rural areas. The final year saw a jump in both urban and rural areas, however, it probably reflected the change of qualification system more than the qualification of staff.

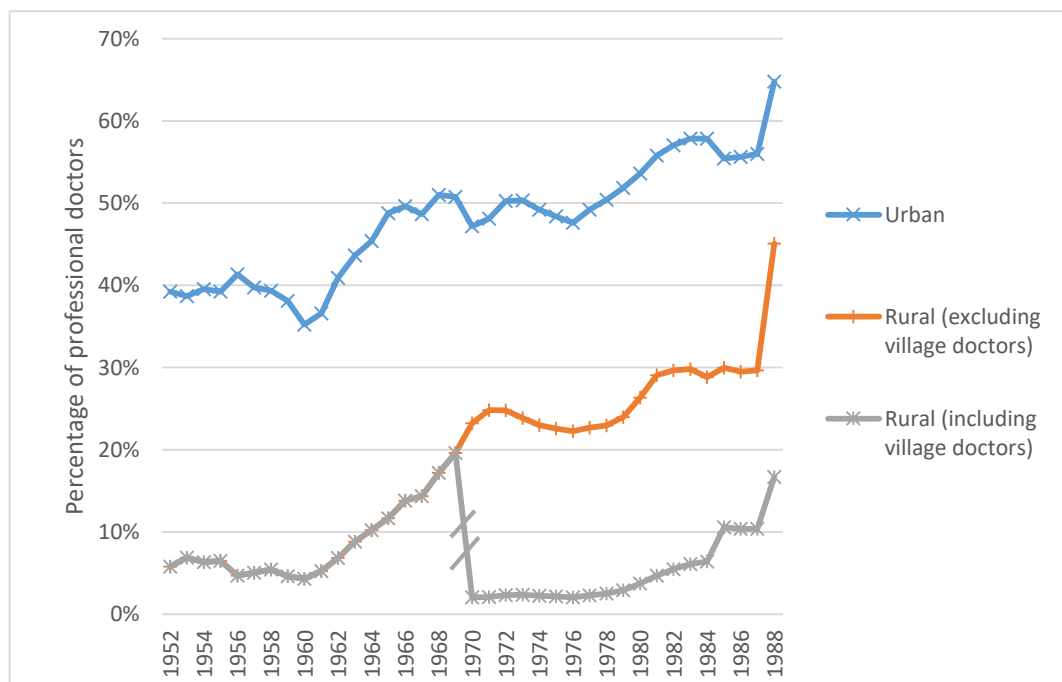


Figure 4-4. Percentage of professional doctors in all practitioners in urban and rural health facilities (1952-1988)

Data sources: (Ministry of Health, 1989)

Note: Data for village doctors (or barefoot doctors before 1985) were not available for the period before 1970.

A similar situation is observed when only rural hospitals and primary care facilities were assessed (see Figure 4-5). There was a fluctuation in the percentage of professionals in county hospitals during the decade between 1952 and 1962, which was likely caused by the increased allocation of middle grade (practitioner-level) practitioners before 1958, and increased allocation of high grade (professional-level) practitioners after 1958, in county hospitals. Overall, the percentage of professionals in county hospital doctors increased slightly from 48% in 1952 to 52% in 1978, and further increased to around 60% in 1986. The percentage of professional doctors at primary care facilities (mainly township health centres) increased from around 5% between 1952 and 1965 significantly over the decade between 1966 and 1975 to 18% and slightly increased to 23% by 1986. When barefoot/village doctors were included, percentage of professionals among doctors in primary care facilities dropped to below 10% between 1975 and 1986, varying between 2% and 7%. In other words, the emergence of barefoot doctors significantly diluted the percentage of professionals among all rural doctors by about 16 percentage points. Overall, the gap between the proportions of professionals among doctors in county hospitals and rural primary care facilities over the three decades or so varied between 25 and 34 or 50 percentage points, depending on whether village doctors were excluded or not.

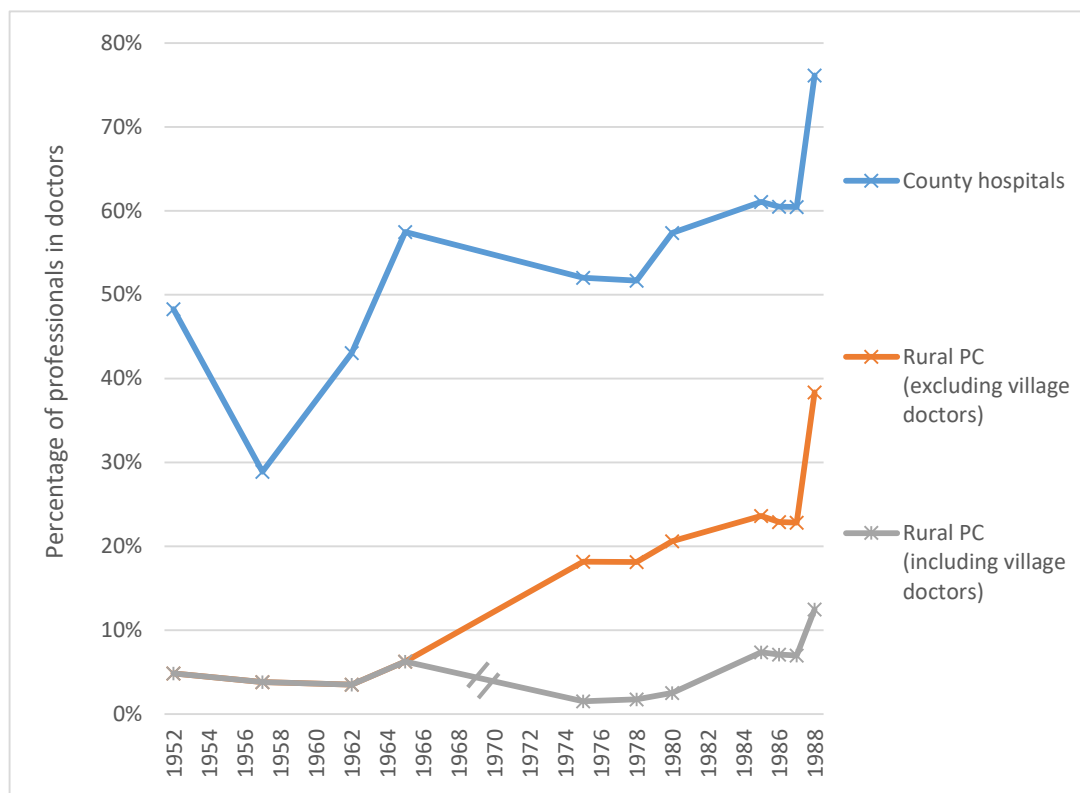


Figure 4-5. Percentage of professional doctors in all doctors in rural primary care facilities and county hospitals (selected years between 1952 and 1988)

Data sources: (Ministry of Health, 1989)

Note: Data for village doctors (or barefoot doctors before 1985) were not available for the period before 1970. Percentage of professionals in doctors in rural primary care facilities was calculated as the total number of primary care professionals (i.e. total rural professionals minus county hospital professionals) divided by the total number of primary care doctors (i.e. total rural doctors, with or without village doctors, minus county hospital doctors)

For the period between 2002 and 2013, the proportion of doctors with a full licence was used. The results are shown in Figure 4-6. Hospitals saw their percentage of fully licenced doctors rose steadily from 88% in 2002 to 93% in 2013. The percentage of fully licenced doctors at primary care facilities excluding village clinics increased steadily from 69% in 2002 to 74% in 2013. The percentage of fully licenced doctors at primary care facilities including village clinics remained stagnant over the period, and reduced slightly from 35% in 2002 to 34% in 2013. Between 2002 and 2013, the percentage rose from 78% to 80% at community health centres and stations, and from 77% to 84% at private clinics. Other facilities saw the percentage of fully licenced

doctors slightly dropping over the same period: sub-district health centres, from 67% to 66%; township health centres, from 61% to 58%; and village clinics, from 7% to 5%.

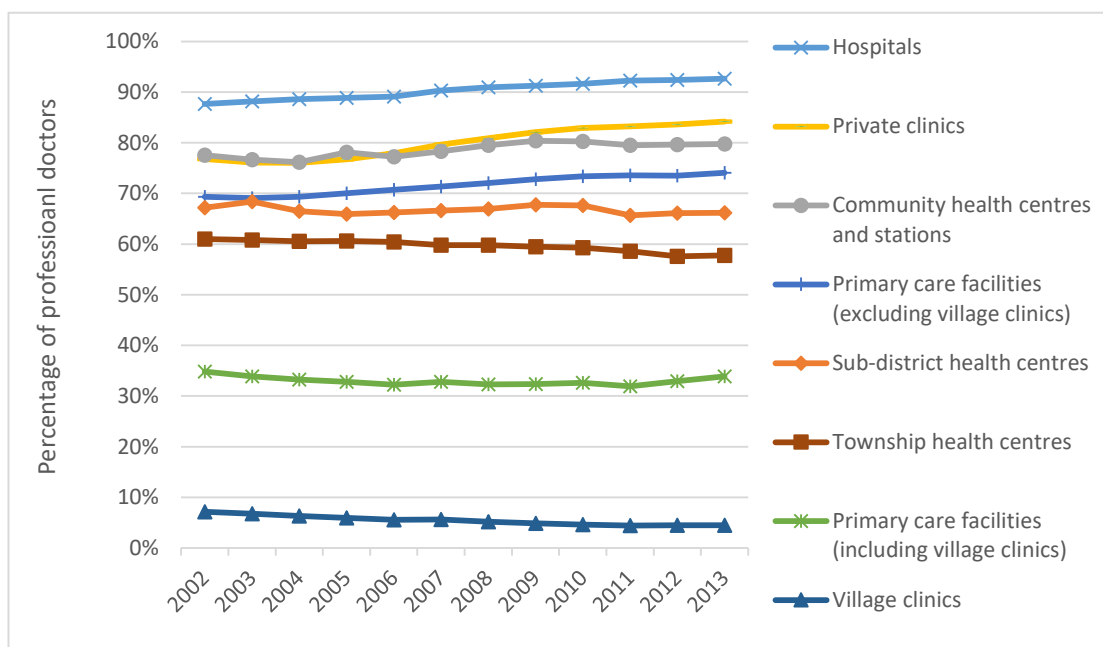


Figure 4-6. Percentage of fully licenced doctors in all doctors at various types of health facilities (2002-2013)

Data sources: (Ministry of Health, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012; National Health and Family Planning Commission, 2013, 2014)

Note:

- Data for village clinics for the years between 2002 and 2009 were projected based on data for the years between 2010 and 2013 for village clinics, assuming a continuous trend.
- Values for primary care facilities were calculated both including and excluding the village clinics.

As mentioned earlier, a large proportion of doctors in China did not have a college degree. After 2002, an increasing proportion of doctors had full medical training, but many practitioners still had only 3-year associate degrees, or graduated from even shorter training programmes. As shown in Figure 4-7, the percentage of doctors with university degrees was 43% in hospitals in 2002. Over the years it grew rapidly, and reached 66% by 2013. In contrast, staff members with university degrees accounted for a very small share of township health centre staff, from 3% in 2002 to 9% in 2013. The staff quality of community health centres improved, as reflected in the increase in

the proportion of staff members with a college degree from 32% to 37% between 2010 and 2013, though this still lagged far behind that of hospitals.⁶

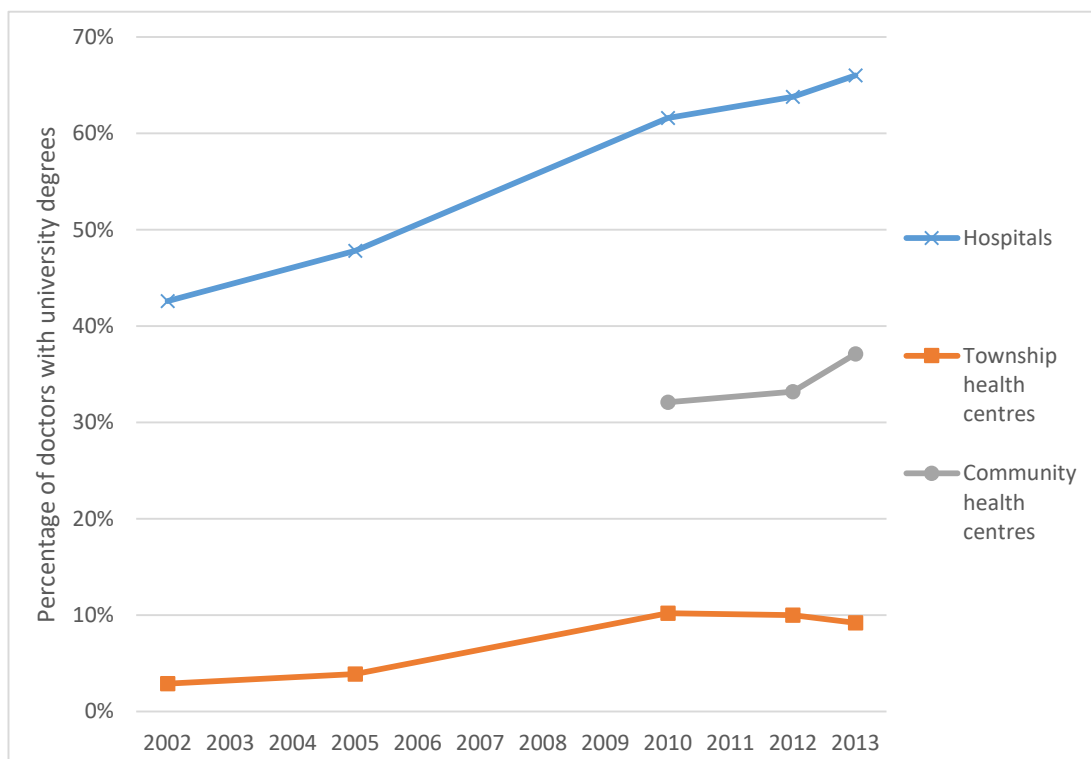


Figure 4-7. Percentage of health professionals with a university degree in selected years between 2002 and 2013

Data sources: (Ministry of Health, 2005, 2009, 2011; National Health and Family Planning Commission, 2013, 2014)

Note: Data are not available for community health centres in 2002 and 2005, and not available for all years for community health stations.

When Figure 4-4, Figure 4-5, Figure 4-6 and Figure 4-7 are examined together, it is clear that the qualifications of doctors in both hospitals and primary care facilities have improved since 1949. However, the stratification of quality of doctors at different levels of care, i.e. good quality doctors were associated primarily with hospitals, has

⁶ The gap between hospitals and primary care facilities in terms of doctors with a university degree was wide, as shown in Appendix Figure 5-1, and the gap in terms of the proportions of college graduates between doctors at hospitals and township health centres particularly increased, from 39.7 to 56.8 percentage points.

been stable over the past six decades.

3) Indicator 1.3: Average annual income of staff

The real average annual income per staff member showed a similar stratification between 2004 and 2013. The average annual staff income (Figure 4-8) rose from 26,500 *yuan* in 2004 to 71,000 *yuan* in 2013 in hospitals, by 2.7 times; and from 10,700 *yuan* in 2004 to 19,800 *yuan* in 2013 in primary care facilities, by 1.9 times. The gap particularly widened after around 2010. By the year 2013, the average staff income in hospitals was about 3.6 times that in primary care facilities, as compared to 2.5 times in 2004. Among primary care facilities, community health centres and stations, sub-district health centres and township health centres tended to do better than other primary care facilities throughout the period. From 2004 to 2013, the average annual staff income at community health centres, for example, increased from 29.5 thousand *yuan* to 36.2 thousand *yuan*; at the bottom of the chart is the average annual staff income at village clinics, which increased from 4.8 thousand to 6.6 thousand *yuan*. Despite some spikes in the income of staff at primary care facilities between 2007 and 2009 caused by no obvious reasons, the general trend was stable from 2004 to 2013.

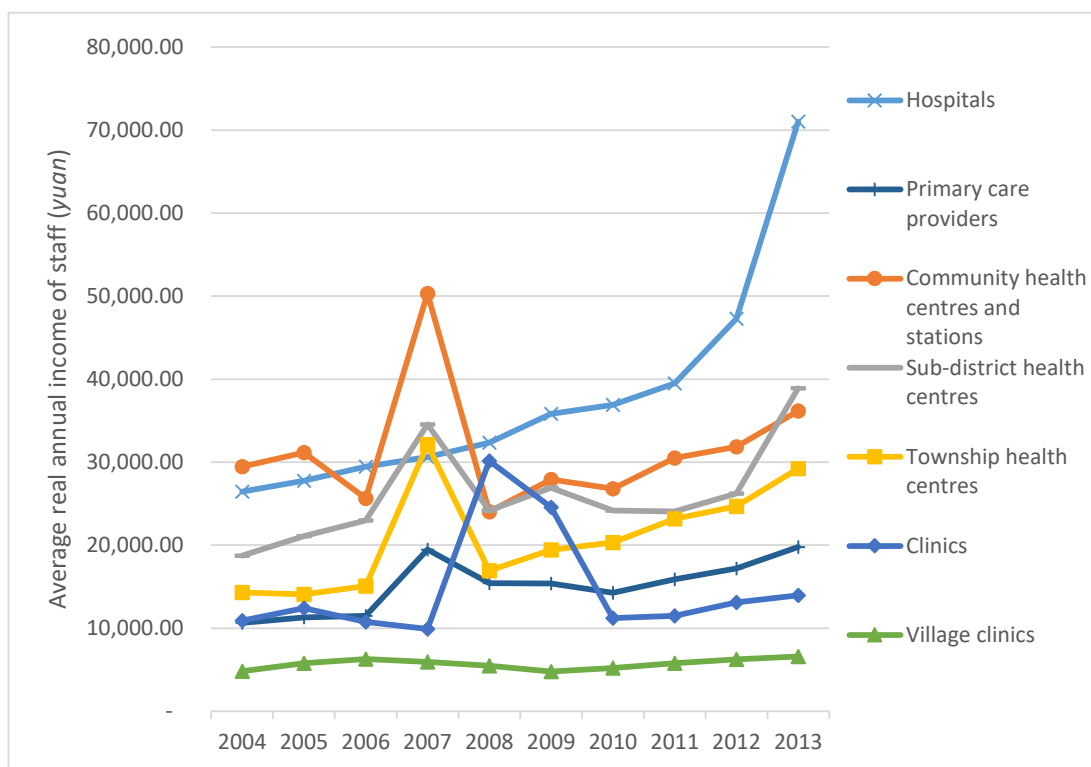


Figure 4-8. Average annual income of staff at various health facilities (2004-2013)

Sources: Author's calculations: average annual income of staff=(annual spending on staff)/(number of staff in the same year). Original data were collected from the China Health (and Family Planning) Statistical Yearbooks (2005-2014) (Ministry of Health, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012; National Health and Family Planning Commission, 2013, 2014).

Note: Values were adjusted for inflation using price deflators calculated based on the price level of the year 2000. The price deflators were calculated from gross domestic product at constant prices and nominal gross domestic product reported by National Bureau of Statistics of China in *China Statistical Yearbook 2015*.

4.4.2 Assets, equipment and infrastructure

Indicators for assets, equipment and infrastructure measured the disparity between the two sectors in terms of non-human capital. They complemented the indicators on human resources in measuring the structural balance. Four indicators under this category are reported here: value of assets, number of beds, number of valuable devices, and building space.

1) Indicator 2.1: Value of assets

As reflected in Figure 4-9, almost all types of health facility saw steady and

significant growth in terms of value of assets. The real total value of hospital assets increased rapidly, from 493 billion *yuan* in 2004 to 1.2 trillion *yuan* in 2013, by almost 2.4 times. Primary care facilities had total assets valued at 68 billion *yuan* in 2004, which increased to 149 billion *yuan* in 2013, by 2.2 times. Over the decade from 2004 to 2013, the value of assets of hospitals was stable and around 88% of all assets of all facilities. Among primary care facilities, community health centres and stations saw their value increased fastest, by 8 times, while township health centres had the predominant share of all primary care assets, varying between 57% and 81%. Private clinics and sub-district health clinics saw some fluctuations over the years, their total share of value of assets among primary care facilities increased from 13% in 2004 to 20% in 2007 and fell to 10% by 2011 (after which the value of assets of private clinics was not available).

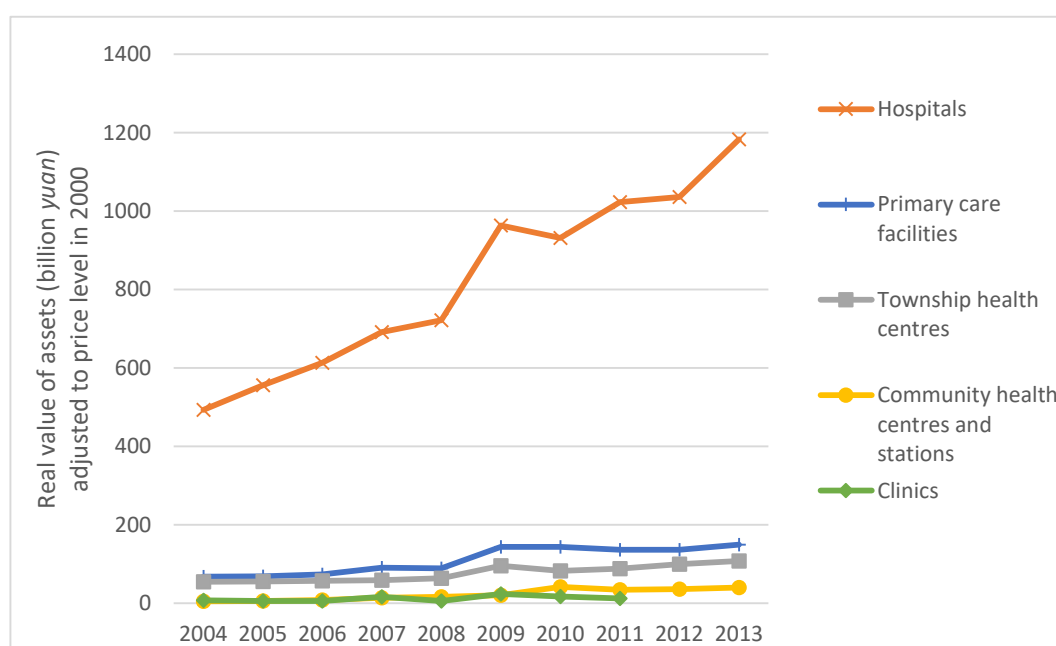


Figure 4-9. Value of assets of selected facilities (1 billion *yuan*) (2004-2013)

Data sources: (Ministry of Health, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012; National Health and Family Planning Commission, 2013, 2014)

Note:

a) Numbers for sub-district health centres are negligible compared to other types of facility and thus are not shown in the graph, other than being included in the total quantity owned by primary care facilities.

b) The numbers for primary care facilities were summed from all types of primary care facility with

available data from 2004 and 2009, after which the numbers reported by the MOH were used.

c) Data for clinics in 2012 and 2013 were not available.

d) Values were adjusted for inflation using price deflators calculated based on the price level of the year 2000. The price deflators were calculated from gross domestic product at constant prices and nominal gross domestic product reported by National Bureau of Statistics of China in *China Statistical Yearbook 2015*.

2) Indicator 2.2: Number of beds

Figure 4-10 shows the trend regarding the number of beds in hospitals and primary care facilities between 1949 and 2013. Both levels saw a steady growth in the number of beds before 1980. Hospital beds rose from about 80,000 in 1949 to 1.2 million in 1980. The number of beds at primary care facilities started at around 30,000 in 1959 following the emergence of township health centres (known as then commune health centres) and reached 0.8 million in 1980. In 1980, while hospitals controlled 61% of all beds, primary care facilities controlled 39%.

The years around 1980 appear to be a turning point when the trend diverged. The number of beds in hospitals maintained rapid growth during the 1980s. The rise slowed down only during the 1990s and beds reached 2.2 million by 2002. In the last decade, the number of beds in hospitals has grown faster and amounted to 4.6 million in 2013. By contrast, the number of beds in primary care facilities kept dropping slowly after 1980 and reached a minimum in 2002 and 2003 at 0.7 million, before doubling in recent years and reaching 1.4 million by 2013. The overall trend of beds in primary care facilities basically represented that of township health centres, though in the last decade (2002-2013) the emergence of community health centres and stations added 0.2 million beds by 2013. In general, the gap between hospitals and township health centres regarding inpatient beds has continued to widen over the past 35 years. By 2013, hospitals controlled 77% of all beds, and primary care facilities controlled 23%.

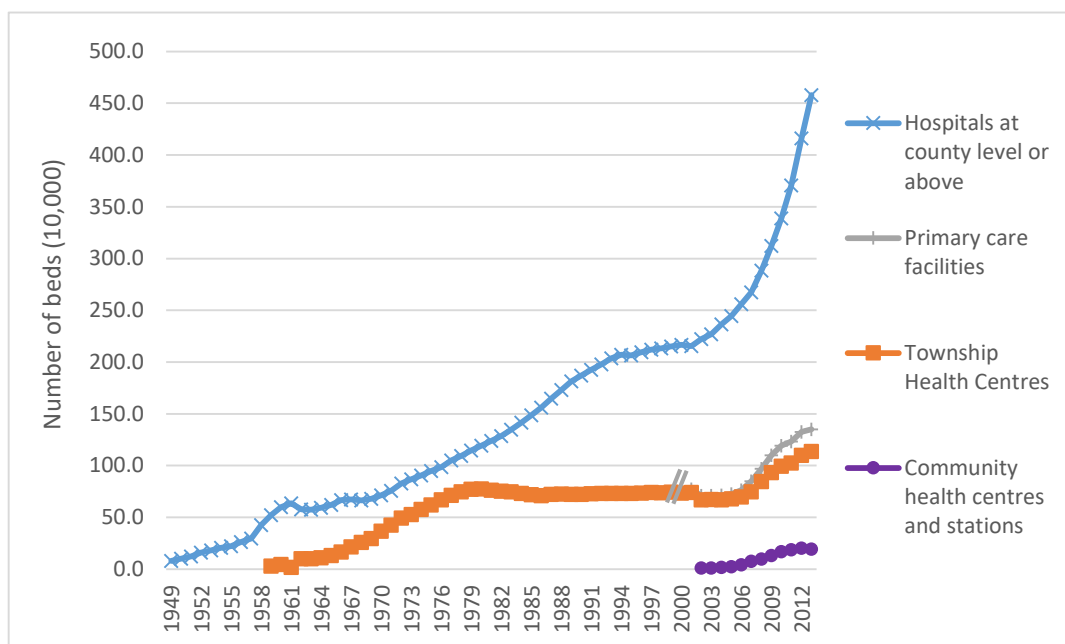


Figure 4-10. Number of beds in hospitals and township health centres (1949-2013)

Source: (Ministry of Health, 1989, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012; National Health and Family Planning Commission, 2013, 2014)

Note: Other types of primary care facilities (sub-district health clinics and clinics) had negligible number of beds and are omitted in the figure.

3) Indicator 2.3: Number of valuable devices

Figure 4-11 shows the trend in number of valuable devices. The number of devices valued above 10,000 *yuan* increased more than 3 times in hospitals, from 1 million in 2003 to 3.2 million in 2013. During the same period, the number increased by 2.4 times, from 0.2 million to 0.5 million in primary care facilities. Although both hospitals and primary care facilities saw rapid growth in numbers of valuable devices, the gap between the two sectors was widening.

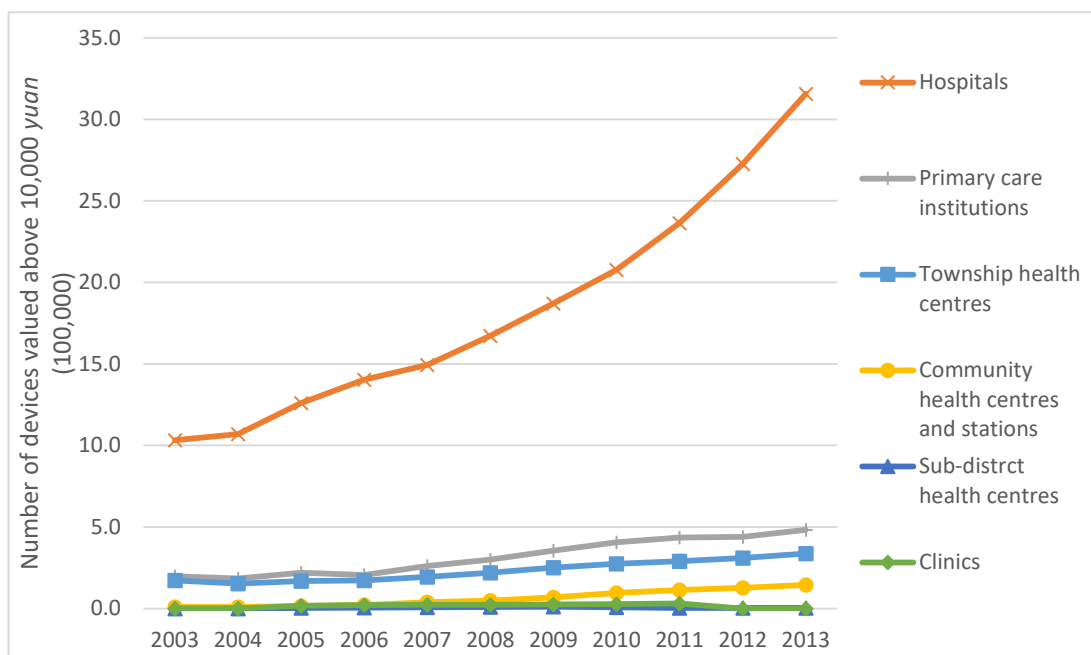


Figure 4-11. Number of valuable devices at hospitals and primary care facilities (2003-2013)

Data sources: (Ministry of Health, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012; National Health and Family Planning Commission, 2013, 2014)

Note:

- a) Primary care facilities here consist of township health centres, village clinics, community health service institutions and other private clinics.
- b) Value of devices not adjusted for inflation.

4) Indicator 2.4: Building space

Total building space saw tremendous growth in both hospitals and primary care facilities from 1988 to 2013. The most rapid development for both levels happened between 1988 and 2003, when total building space of hospitals grew from 82.3 million square metres to 227.1 million square metres, by 2.8 times, and building space of primary care facilities (primarily township health centres) increased from 8.9 million square metres to 88.3 million square metres, by about 10 times. Primary care facilities represented a growing proportion of total building space during this period, from 9.8% to 28% (a proportion that would be much greater if village clinics were included).

During the last two decades, the building space of township health centres was

rather stable. However, with the emergence of community health centres and the inclusion of data from village clinics, the gap between the building space of hospitals and that of primary care facilities continued to shrink. By 2010, the gap was at its narrowest, when primary care facilities accounted for 38.8% of all building space. The proportion declined to 32.9% in 2013.

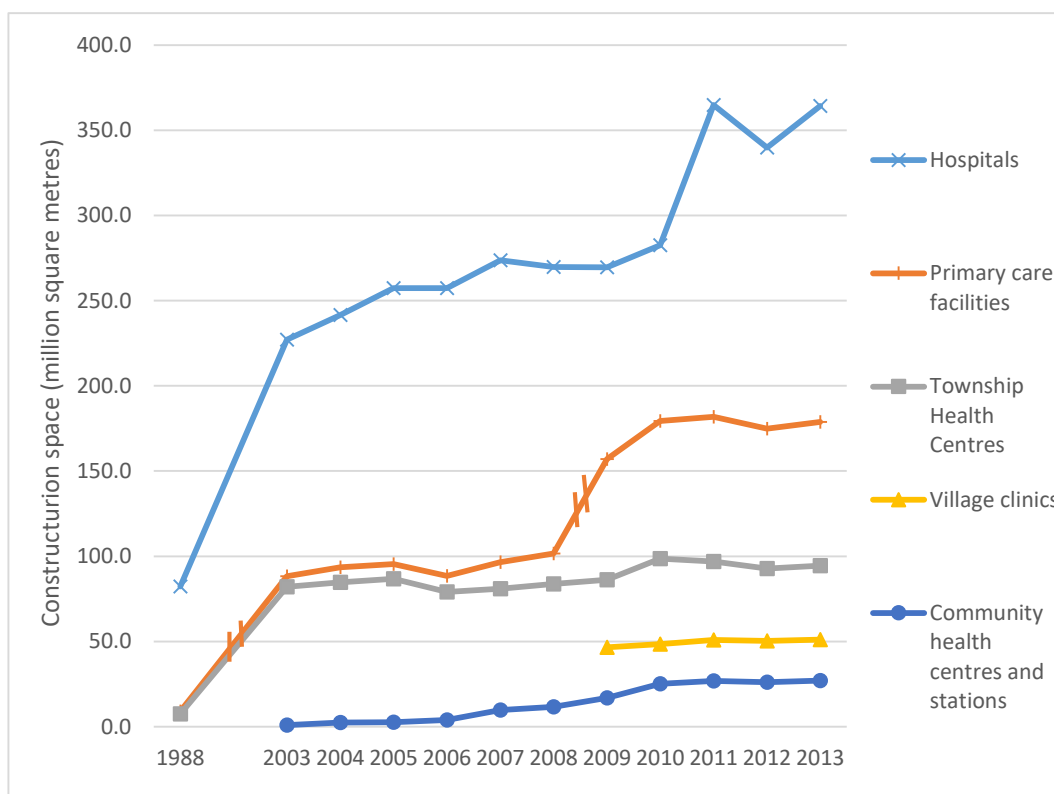


Figure 4-12. Building spaces of hospitals and primary care facilities (1988 and 2003-2013)

Source: (Ministry of Health, 1991, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012; National Health and Family Planning Commission, 2013, 2014)

Note:

- No data were available for the period between 1989 and 2002.
- No data were available for village clinics in years 2003-2008.
- Compared to the types of facility listed in the figure, the space of sub-district health centres and ambulatory clinics was negligible and hence not reported here (though included in the sum of primary care facilities).

4.4.3 Service utilisation

Patient utilisation of services constitutes an important measure of the functional

balance. The study analysed the number of patient visits, the number of admissions, and the ratio of patient visits to admissions.

1) Indicator 3.1: Number of visits

In terms of patient visits, the hospitals seemed to be playing an ever-larger role over the last three decades. The results are presented in Figure 4-13. Data were available from 1981 to 2013. The first two decades (1981-2003) showed divergence between hospitals and township health centres. The number of visits to hospitals increased from 1.1 billion to 1.3 billion by 2004. By contrast, number of visits to township health centres continuously declined, from 1.4 billion in 1981, and reached a minimum (0.68 billion) in 2004. Unfortunately, data were generally not available for other primary care facilities during these two decades.

The years 2003 and 2004 were a turning point after which all types of facility saw rapid increase of patient volume. Visits to hospitals started to take off and more than doubled, reaching 2.7 billion by 2013. Total visits to primary care facilities also increased, from 2.6 billion in 2004 to 4.3 billion in 2013. Among primary care facilities, visits to township health centres grew from 0.7 billion in 2004 to 1 billion in 2013; visits to village clinics grew from 1.2 billion in 2004 to 2 billion in 2013; and visits to community health centres and stations also grew from an initial 0.1 billion to 0.7 billion. However, the proportion of visits to primary care facilities among all facilities decreased, from 66.4% to 61.2%.

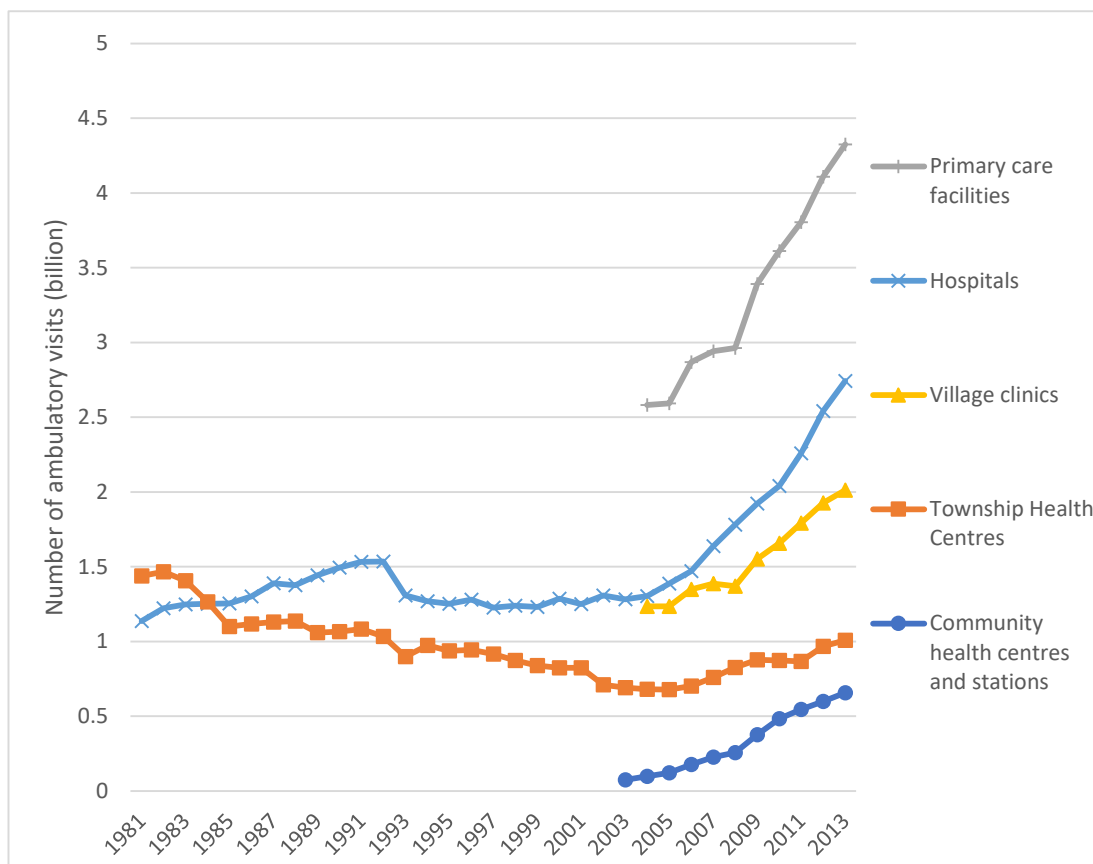


Figure 4-13. Number of patient visits to hospitals and primary care facilities (1981-2013)

Data sources: (Ministry of Health, 1989, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012; National Health and Family Planning Commission, 2013, 2014)

Note:

- 1) Patient visits consisted mainly of outpatient visits and emergency care, but also included physical check-ups and health consultations.
- 2) Data are not available for primary care facilities (except township health centres) before 2004.
- 3) Primary care facilities here include township health centres, village clinics, community health service institutions and other private clinics.
- 4) Sub-district health centres and clinics are not shown in the graph as their numbers were negligible compared to other facilities.

2) Indicator 3.2: Number of admissions

The number of inpatient admissions reflected a similar trend to patient visits, and can be described in two segments divided by 2003 (see Figure 4-14). Between 1981 and 2003, the number of admissions in hospitals almost doubled, from 23.5 million to 43.8 million. In contrast, the number of admissions in township health centres, which

started at a similar level to that of admissions in hospitals, declined by 24.1% from 21.2 million to 16.1 million. Over the last decade, however, admissions rose significantly for both types of facility. The number of admissions in hospitals increased by 3.2 times from 2003 to 2013 and reached 39.4 million. The number of admissions in township health centres increased by 2.4 times from 2003 to 2009 and reached 38.1 million. After a little fluctuation, the number climbed back to 39.4 million in 2013. This was in sharp contrast to the steady and rapid growth in hospitals. Admissions to other types of facilities were negligible compared to these two types of facility.

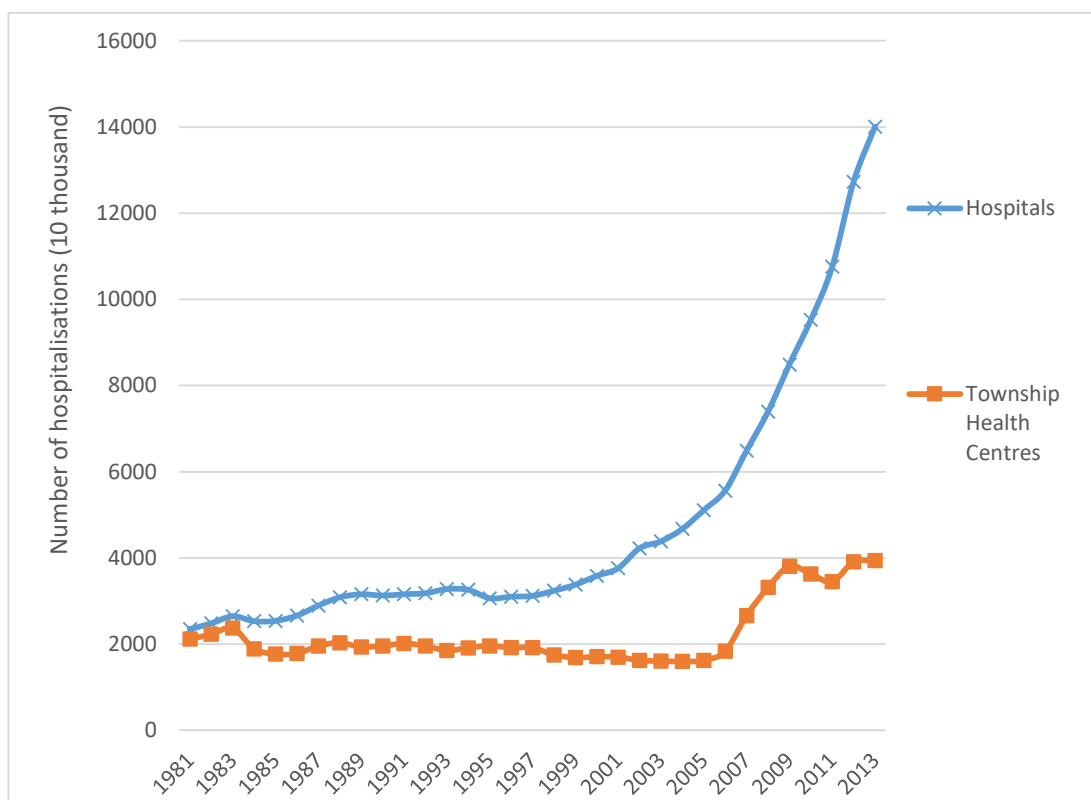


Figure 4-14. Number of admissions at hospitals and township health centres (1981-2013)

Data sources: (Ministry of Health, 1989, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012; National Health and Family Planning Commission, 2013, 2014)

Note:

- Information for township health centres before 1981 was not available, so the figure compares only the period between 1981 and 2013.
- Compared to the township health centres, the number of inpatients at community health centres was negligible and thus is not reported here.

3) Indicator 3.3: Visits to admissions ratio

Data were available only for hospitals and township health centres for the first two decades since 1986, while data for other primary care facilities were available for the years from 2004 to 2013. As Figure 4-15 shows, the visits to admissions ratio reduced consistently over the past three decades in both township health centres, from 67.7 to 25.6, and hospitals, from 48.4 to 19.6, mainly due to the rapid increase in inpatient services, which favoured hospitals. As township health centres were the only group of primary care facilities providing substantial inpatient services, this underestimated the visits to admissions ratio at primary care providers. When other primary care facilities are included from 2004, the ratio became much larger for primary care facilities. However, the general trend of a decreasing visits to admissions ratio still held, and it fell from 159.9 in 2004 to 101.5 in 2013. Overall, there appeared to be a continuous phenomenon of an increasingly hospitalized style of medical care across both primary care facilities and hospitals over the past three decades.

The graph also reflects the similarity of township health centres to hospitals in general. Indeed, the visits to admissions ratio in township health centres closely followed that of hospitals throughout the period for which the data were collected, and particularly around 1991 and after 2007.

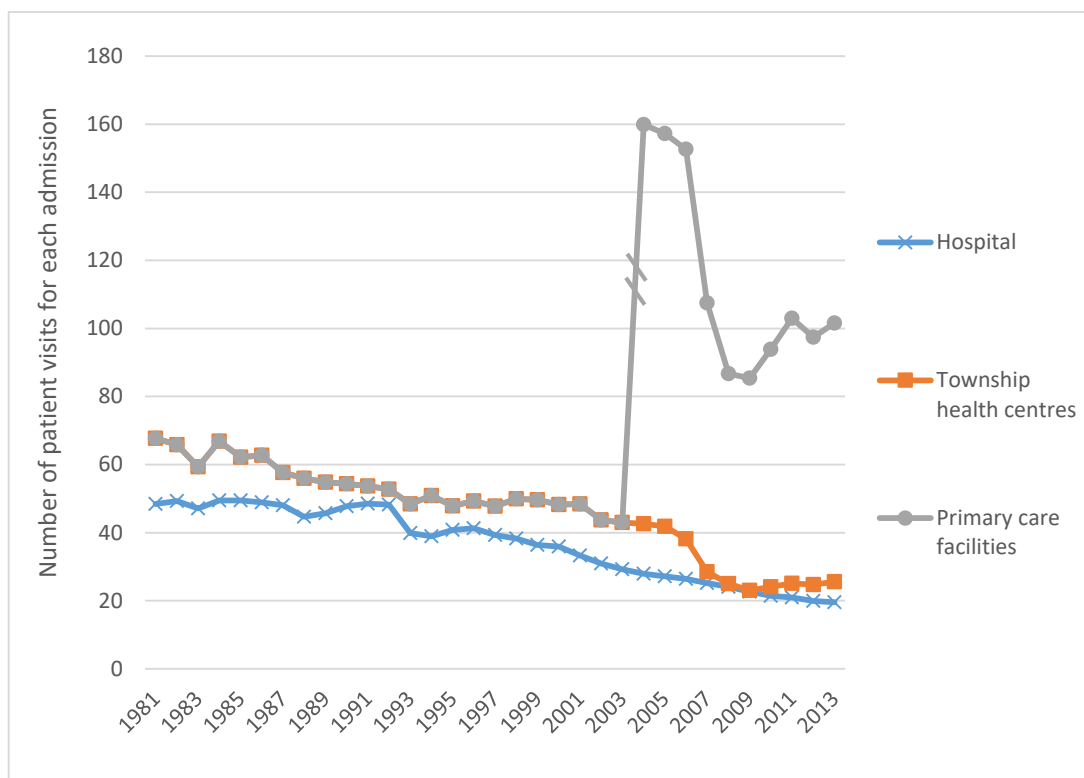


Figure 4-15. Visits to admissions ratio in hospitals and primary care facilities (1981-2013)

Data sources: Original data were collected from national health statistical reports (Ministry of Health, 1989, 2004; National Health and Family Planning Commission, 2014).

Note:

- Visits to admissions ratio=annual patient visits ÷ annual inpatient admissions;
- Prior to 2004, only data for township health centres were available; from 2004, data for all other primary care facilities were included.

4.4.4 Financing

1) Indicator 4.1: Revenues

All types of facility saw significant growth in revenue over the last decade, while total revenue continued to be concentrated in hospitals. Between 2004 and 2013, revenue of hospitals increased rapidly, from 340.7 billion *yuan* to 1,039.2 billion *yuan*, by 3 times, while primary care facilities saw revenue increased from 103.9 billion *yuan* to 206.8 billion *yuan*, by 2 times. Hospitals accounted for 83.4% of total revenue of health facilities in 2013, compared to 76.6% in 2004. Within primary care facilities, community health centres saw the greatest growth, by 6.4 times, while township health

centres maintained the lead in the share of overall revenue of primary care facilities, which increased from 41% in 2004 to 52% in 2013. Despite a temporary upturn of township health centres in 2007, the trend of primary care facilities has been generally stable.

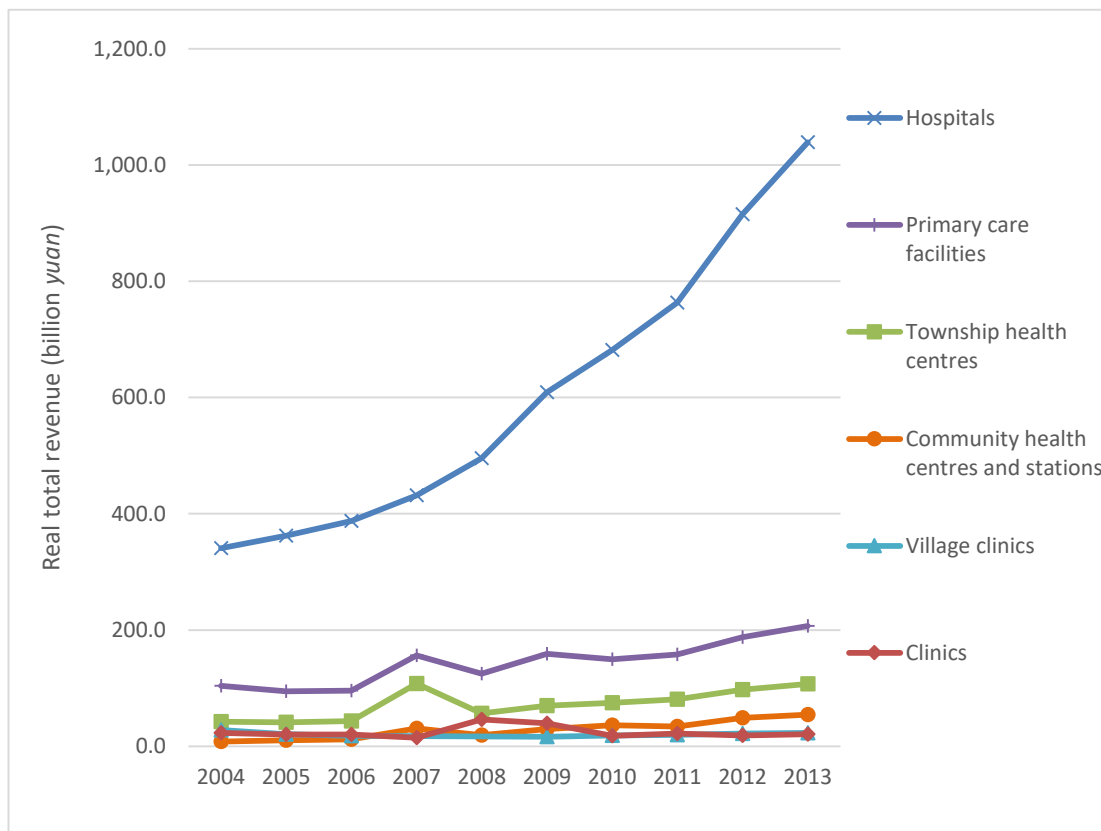


Figure 4-16. Total revenue of hospitals and primary care facilities (billion yuan) (2004-2013)

Data sources: (Ministry of Health, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012; National Bureau of Statistics of China, 2015; National Health and Family Planning Commission, 2013, 2014)

Note:

- Information about village clinics was not available for the years 2007 and 2008.
- Sub-district health centre revenue was negligible compared to other types of facilities and thus not reported.
- Revenue was adjusted for inflation using the price level of the year 2000. The inflation rates were calculated from gross domestic product at constant prices and nominal gross domestic product reported by National Bureau of Statistics of China in *China Statistical Yearbook 2015*.

4.5 Discussion

The chapter has developed and used a system of metrics with 11 indicators to

describe and analyse the trend of the balance between hospitals and primary care providers from 1949 to 2013. This section discusses the strengths and limitations of the methods and the findings of the study.

4.5.1 Strengths and limitations

The chapter has attempted to analyse the trends of the balance between hospitals and primary care providers between 1949 and 2013. In order to address a lack of comprehensive measurement of the balance identified in the literature, the sub-study has built a system of metrics based on previously used indicators found in the literature which measured the two categories of structural and functional balance between the two sectors.

For the structural domain, the metrics included not only indicators related to the previous attention regarding human resources in the literature but also new indicators to cover non-human capital, to provide a more comprehensive view of the structural balance. Under human resources, the metrics included both quantity and quality measurements, following Starfield (1994) and adapting her comparison between primary care doctors and specialists to the comparison between primary care providers and hospitals. Under non-human capital, the sub-study added value of assets, number of beds, number of valuable devices, and infrastructure (building space).

For the functional domain, the sub-study used the previous categories identified regarding service utilisation and financing. Under service utilisation, the metrics adapted previous attempts to measure distribution of service utilisation found in the literature (particularly White et al. (1961) and Starfield (1994)), to include numbers of visits and admissions, as well as the visits to admission ratio, in order to reflect the reality in China that primary care providers and hospitals have been providing a significantly overlapping range of services in terms of both inpatient and ambulatory care. Under financing, the metrics followed Mills (1990) and some more recent studies

and compared the total revenues of primary care providers and hospitals.

Retrospectively constructing the historical trends of the structural and functional balance between hospital and primary care for China has major inherent challenges. Section 4.3 has raised issues related to both the nature of the subject and the nature of the data. The sub-study involves various aspects of structural and functional balance between the two groups of various types of facility, over a dynamic period of history. The study tried to develop a comprehensive system of metrics with a degree of flexibility. However, it is not possible to capture all dimensions of the structural and functional balance between hospitals and primary care providers using a limited number of indicators. A key limitation of the metrics is the lack of indicators that reflect the value and quality of care. Such data were not available.

There were also issues about comparability between hospitals and primary care providers. This perhaps makes interpretation of some aspects of balance more difficult than others. Several indicators are biased towards hospitals, i.e. beds, expensive equipment, space, which are not directly vital to primary health care. The visits and admissions to hospitals and primary care facilities probably were of different illness severity, but again the information could not be captured in this analysis. The chapter has attempted to compare the two sectors as holistic entities rather than looking at the margins with more obvious overlaps.

Data availability issues existed for almost all the indicators used in the study, but to differing extents. The functional aspects and certain types of primary care facilities were more likely to have been under-represented, particularly in earlier periods. The original data were collected and reported by the Ministry of Health (National Health and Family Planning Commission since 2013), which generally compiled data from provincial annual statistics. They shared the strengths and weaknesses of officially collected and reported data. The data were checked multiple times by multiple layers

of public servants, but might have been manipulated for their own interests (e.g. potential over-reporting to show achievements or under-reporting to justify policy changes).

4.5.2 Discussion about findings

Table 4-2 pulls together the results to show the changes in all indicators across hospitals and primary care facilities over the years with available data. The table also presents the ratio of changes between hospitals and primary care facilities over three specific periods (see rows with “H:PC”), namely from 1952 to 1985, from 1985 to 2013, and from 2003 to 2013. Besides accommodating varied availability of data across different indicators and years, these periods were chosen to allow comparison of trends across roughly the first three decades by the early 1980s, roughly the second three decades from then until 2013, and the most recent decade. The punctuation around the early 1980s was based on both the above analysis but also the analysis in Chapter 6 where the history after 1949 was segmented by the year 1978. The last decade was also highlighted to have a sense of changes more recently. The remainder of this sub-section draws from both the results section and Table 4-2 to discuss structural and functional balance between hospitals and primary care providers from 1949 to 2013.

Table 4-2. Comparison across indicators over time

Indicator		1985:1952 ^a	2013:1985	2013:2003
1.1 Number of staff	H ^b	9.72 (1985:1957)	2.62	1.75
	PC ^c	5.70 (1985:1958)	1.60	1.37
	H:PC^d	1.71	1.64	1.28
1.2 Qualification of health practitioners^e	H	16 p.p. ^f		23 p.p. (2013:2002)
	PC	6 p.p. ^f		6 p.p. (2013:2002)
	H:PC	2.67		3.71
1.3 Average annual income of staff	H			2.68 (2013:2004)
	PC			1.86 (2013:2004)
	H:PC			1.45
2.1 Value of assets	H			2.40 (2013:2004)
	PC			2.20 (2013:2004)
	H:PC			1.09
2.2 Number of beds	H	2.85 (1985:1959)	3.08	2.02
	PC	23.84 (1985:1959)	1.87	1.90
	H:PC	0.12	1.64	1.06
2.3 Number of valuable devices	H			3.06
	PC			2.43
	H:PC			1.26
2.4 Building space	H		4.43 (2013:1988)	1.60
	PC		20.04 (2013:1988)	2.03
	H:PC		0.22	0.79
3.1 Number of visits	H		3.03 (2013:1988)	1.83 (2013:2004)
	PC		0.92 (2013:1988)	1.68 (2013:2004)
	H:PC		3.31	1.09
3.2 Number of admissions	H		5.52	3.20
	PC		2.22	2.45
	H:PC		2.48	1.30
3.3 Visits to admissions ratio	H		0.40	0.70 (2013:2004)
	PC		0.41 ^g	0.64 (2013:2004)
	H:PC		1.10	1.10
4.1 Revenues	H			3.05 (2013:2004)
	PC			1.99 (2013:2004)
	H:PC			1.53

Data sources: author's calculations based on original data from *Health Statistics information in China 1949-1988*, the statistical sections of *China Health Yearbooks* (1983-2008), *China Health Statistical Yearbooks* (2003-2012), and *China Health and Family Planning Statistical Yearbooks* (2013 and 2014).

Note:

a) For each period, the ratio is between the value of a certain level of care in the year at the end of the period and the value of the same level of care in the year at the start of the period.

b) H: hospitals.

c) PC: primary care facilities.

d) H:PC stands for the ratio between the magnitude of changes of hospitals and primary care facilities for a certain indicator over a specified period.

e) Qualification of health practitioners refers to percentage of high quality doctors (professionals for 1985:1952, doctors with university degrees for 2013:2003) in all doctors by types of facility. Change in percentage points (p.p.) is measured here.

f) Urban value was used for hospitals, while rural value was used for primary care facilities.

g) The value for township health centres was used here for primary care providers.

Within the structural domain, data about human resources covered most periods from 1949 to 2013. The growth of hospitals in terms of both quantity and quality was more stable over the years, while there was greater fluctuation in primary care facility. Despite a drastic emergence of a vast number of barefoot doctors during the 1970s, the number of staff at hospitals increased faster overall than at primary care facilities (i.e. 1.71 times that of primary care facilities from 1952 to 1985 and 1.64 times from 1985 to 2013). The number of staff at hospitals overtook that at primary care providers in 1987. Afterwards, the gap widened, and hospitals had 1.5 times the number of staff of primary care providers in 2013.

Hospitals maintained superiority regarding quality of practitioners over primary care providers, with clear stratification and a widening gap across levels of care over the period studied. Three processes seemed to be associated with hospital dominance in quality of staff. First, the township health centres, which emerged in the late 1950s, consistently had less qualified practitioners compared to hospitals over time. Second, the recruitment of de-professionalized barefoot doctors during the 1970s further lowered the level of overall qualification of primary care staff. Third, in the last decade (2003-2013), hospitals developed much more rapidly, and increasingly absorbed doctors with university degrees. This recent development was also encouraged by the continuously widening gap between the salaries of staff at hospitals and primary care facilities.

In terms of physical assets, data were available only for more recent decades. The lead of hospitals over primary care facilities was more dramatic. The analysis demonstrated at first a parallel and then diverging development between the two sectors. Primary care facilities (solely represented here by township health centres) followed the pattern of hospitals in accumulating a substantial number of beds. The catching up was conspicuously reflected in the much faster growth of beds in primary

care than in hospitals from 1959 to 1985 (H:PC=0.12). In the last three decades, the increase in number of beds at primary care facilities slowed down and fell behind hospitals. Over the last decade, hospitals accumulated assets both rapidly and steadily, and hospitals' lead over primary care facilities widened. However, primary care providers saw their building space developing much faster than that of hospitals in the last three decades. Both primary care facilities and hospitals saw growth in terms of indicators on assets (number of beds from 1952 to 2013, and value of assets from 2003 to 2013, and building space from 1985 to 2013) which was which was faster than growth in human resources. This relatively slower development in human resources perhaps reflected a lack of human-centeredness critical for primary care development.

Among functional indicators, available data were again limited to the last three decades for service utilisation. During the years between 1981 and 2002, there was at hospitals an increasing number of inpatients *vis-a-vis* generally stable ambulatory visits, pushing down the visits to admissions ratio of hospitals. Meanwhile, at township health centres, a rapidly decreasing number of ambulatory visits and a slowly decreasing number of inpatients led to a decline of the visits to admissions ratio there. The last decade saw substantial growth of both visits and admissions to both hospitals and primary care facilities; however, hospitals were playing an increasing role in both respects. While both hospitals and primary care facilities were becoming more oriented towards inpatient care, hospitals accounted for almost 2 in 5 ambulatory visits (up from about 1 in 3 in 2004) and 3 in 4 inpatients (up from 1 in 2 in 1981) by 2013. As an increasing proportion of patients flowed to hospitals, funding was also distributed predominantly to hospitals, increasing from 76.6% in 2004 to 83.4% in 2013.

Overall, the dominance of hospitals on structural and functional dimensions seemed to be an enduring issue from 1949 to 2013. Hospitals were staffed with a large number of practitioners (overtaking primary care providers since the 1980s) who were much more likely to be well qualified and better paid, as compared with primary care

providers. Hospitals also had stronger physical assets, including beds and devices, than primary care providers. Functionally, the share of hospitals in both ambulatory and inpatient care was substantial. Hospitals have also been taking an increasingly large share in total spending on facilities. The increasing dominance of hospitals was obvious after 1978, associated with two decades of deterioration of primary care providers (particularly township health centres) in ambulatory care provision and the recent boom of hospitals in both inpatient and ambulatory care.

The metrics obscure the reality that hospitals and primary care providers were not easily discernable. Hospitals have been a key provider in ambulatory care throughout the last three decades when data were available, and were increasingly so. Primary care providers as a whole seemed to have been mimicking the practice of hospitals, particularly in terms of the development of hospital beds and provision of inpatient care in township health centres. There was an almost parallel increase in focus on inpatient care in both hospitals and primary care facilities. Meanwhile, primary care also developed a large sector of de-professionalized staff at the village level. What was formed appeared to be less like two levels of differentiated specialist vs generalist care and more like a two-tier delivery system with a similarly hybrid structure of inpatient and ambulatory care with a stratified quality of practitioners. With average income of hospital staff increasing significantly faster than that of primary care staff, hospitals were more likely to be able to hire and retain better-qualified staff. At the end of the period studied, hospitals appeared to be enjoying an environment, including premium staff quality and pay, superior assets, increasing share of ambulatory care, and orientation towards inpatient care, that fundamentally gave them an edge over primary care facilities.

4.6 Conclusion

This chapter has addressed the objective to study the historical trend of balance

between primary and hospital care in China by developing a system of metrics consisting of 11 indicators regarding balance of care adapted to the Chinese context, and using it to collect and analyse data. Despite limited and inconsistent availability of data, the analysis has shown that the bias of balance of health delivery towards hospitals was continuous over the past six decades. Over time, hospitals were staffed with a larger number of practitioners both better qualified and paid, and equipped with a greater value of physical capital, as compared with primary care providers. Functionally, the share of hospitals has been disproportional in ambulatory care, and dominant in total spending on facilities.

Apparent from the analysis was that the hospitals and primary care providers were not easily distinguishable. While hospitals have been an increasingly substantial provider in ambulatory care, primary care providers delivered both ambulatory and inpatient care though with relatively under-professionalized staff. The definition of primary care facilities used here followed the definition of relatively recent government policy documents. However, as will be clear from the coming chapters, this is not just a matter of different definitions. Explaining the historical roots of such a strange situation is central to understanding of contemporary difficulties in strengthening primary care in China. This is the task of the next two chapters.

Chapter 5 Historical roots of hospital dominance (1835-1949)

“It is difficult to convey to a person who has not witnessed the scenes of the hospital, a just idea of them. He needs to be present on a day for receiving new patients, and behold respectable women and children assembling at the door the preceding evening, and sitting all night in the streets, that they might be in time to obtain an early ticket, so as to be treated the same day. He needs to behold in the morning the long line of sedans extending far in each direction; see the mandarins, with their attendant footmen, horsemen, and standard-bearers; observe the dense mass in the room below – parents lifting their children at arm’s length above the crowd, lest they should be suffocated or injured; stand by during the examination and are registered and prescribed for; in urgent cases being admitted at once, while others are directed to come in five or ten days, according to the ability to attend to them.” (Parker, 1842, p. 22)

5.1 Introduction

The scene of the Canton Ophthalmologist Hospital (better known as Canton Hospital) that Parker founded in 1835 would seem oddly familiar to ordinary Chinese today, if the sedans were replaced with cars, and attendant footmen and other companions were replaced with family dependents. Even more striking is that such a form of health care provider set up by foreigners two hundred years ago stayed and became a significant part of ordinary lives in China today, despite the fact that health policy making had long been under the control of Chinese nationals and that several waves of effort to shift the balance towards primary care. This chapter and the next seek to explain how hospital-centralism came into being as a set of institutions, and how it became resilient over periods of major changes in China since 1835. The rationale for the *longue durée* came from the hypothesis of the theory of path

dependence—that earlier events at critical junctures had huge influence over establishing long-term feedback that became difficult to challenge, and individual events were influenced by the long-term processes (Pierson, 2004). The two chapters are punctuated by 1949, the year when the People’s Republic of China was founded. The methods used for the two chapters are explained in this chapter, and an overall conclusion of the two periods is given in Chapter 6.

The two chapters use an analytical framework based on historical institutionalism set out in Section 5.2 along with other methodological details. Section 5.3 further provides the overall historical context necessary to understand the later analysis. Before starting the main story, Section 5.4 explains the development of medical services and health-related philanthropy from the late Ming (1572-1644) to before China’s engagement with Western medicine via the Canton Hospital in 1835. Section 5.5 describes development from 1835 to 1928, the year the Ministry of Health was established. Section 5.6 describes the rise of a state medicine agenda and its impact from 1928 to 1949.

5.2 Methods

Based on the theoretical lens of path dependence and the approach of critical juncture analysis (see Chapter 2: sub-section 2.6.3), the sub-study developed the logical steps of a path dependence analysis for understanding the historical co-evolution of hospitals and primary care providers in China. After laying out the analytical steps, the section will provide a brief account of data collection.

5.2.1 Analytical framework

Path dependence theory has been shown useful for understanding complex and long-term health systems changes. Sub-section 2.6.3 reviewed the theories related to historical institutionalism and path dependence and summarized four types of path

dependence model (see Figure 2-3). An iterative process of analysis of historical data (documents, archives, secondary sources, etc.) has led to the following model that fit with the theoretical prediction of path dependence in health policy and systems analysis (see Figure 5-1). The model starts with the pre-juncture where contingency of paths was possible. At a critical juncture, events take place and shape the developmental path. The critical juncture is followed by a post-juncture period of reinforcing. Then at a point of time (conjuncture) when multiple factors gathered, external forces opened up an opportunity for major changes (“policy episodes” according to Tuohy (1999)). After the conjuncture, policies normalized and the systems previously created will exhibit elements of their resilience to some extent.

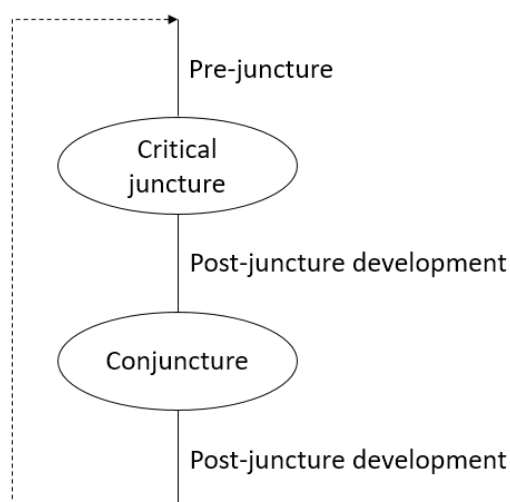


Figure 5-1. A model of dynamic path dependence

5.2.2 Protocol

Table 5-1 lists the analytical steps developed based on this framework. The protocol was modified based on an approach for critical juncture analysis (Capoccia, 2015). The first two steps clarified the focus of the study: the institutions and timeframes of interests. Step 3 and step 4 provided the context regarding wider history

and political structure. Step 5 was the main step of data gathering: preliminary identification of main actors and extraction of what they had done regarding the policy and system dynamics during the periods of interest.

The theoretical analysis started with Step 6. Step 6 and 7 identified candidates for critical junctures and conjunctures, and then selected the most consequential critical junctures and conjunctures. Step 8 reconstructed the temporal sequence of key historical processes using the sequence template of pre-critical-juncture, critical juncture, post-critical-juncture path dependence, conjuncture, and post-conjuncture development. The components regarding counterfactual analysis was dropped due to the lack of comparative cases. The significance of events and processes was judged by their historical consequences.

Table 5-1. Analytical steps for path dependence analysis

Step	Content
1	Clarify the unit of analysis
2	Clarify chronological scope and periodization
3	Extraction of historical background
4	Extraction of political structure
5	Preliminary identification of main actors and extraction of their actions and interactions during the periods of interest
6	Identify “candidate” critical junctures and select for analysis the most consequential critical junctures
7	Identify “candidate” conjunctures and select for analysis the most consequential conjunctures
8	Reconstruct the key historical processes following their temporal sequence

Source: adapted from (Capoccia, 2015) with substantial modification

5.2.3 Unit of analysis: institutional complex behind the coevolution of primary care and hospital sectors

For historical institutionalism, the unit of analysis is usually a policy or institution or a number of inter-related institutions. The two history chapters studied the institutional complex related to the balance between primary and hospital care, which involves a set of interrelated institutions associated with hospitals and primary care

providers. Informed by a categorization of health system building blocks (World Health Organization, 2007) (i.e. service delivery, health workforce, health information, medical technologies, health financing, and leadership and governance), and the major institutions governing decision making in health proposed by Tuohy (1999) (namely, hierarchy, market and collegiate), the thesis proposed four categories of institutions related to hospitals and primary care facilities (and therefore the balance between the two): professional, organizational, financial and governance institutions.

First, professional institutions referred to institutional arrangements related to human resources and the medical education policies which determined not only the human capital of health facilities, but also what levels of practitioners were associated with the facilities. Professional institutions are particularly prone to path dependence as health care is a highly specialized field and professional training comprises specific assets not easily transferrable to other occupations. Besides, the professional institutions also involve construction of identities and trust, which takes time to build.

Second, organizational institutions referred to institutional arrangements related to the organizational boundaries between hospitals and primary care providers. Whether or not hospitals contained large outpatient departments, whether the primary care facilities provided significant hospitalization functions, whether the primary care facilities provided primarily public health services, and whether there were functional referral arrangements between hospitals and primary care providers, were the key factors of concern here.

Third, financial institutions referred to institutional arrangements related to the sources and mechanisms for paying hospitals and primary care facilities. Through different historical periods, the sources of finance varied greatly, ranging from private philanthropy to public finance and private out of pocket payment, or through kickback or mark-up from drug sales, etc.

Fourth, governance institutions referred to institutional arrangements that determine the relationship between facilities and the actors to whom they are responsible. The actors could be government, private individuals, organizations, collectives, or a group of multiple players, etc. These institutions determined the facilities' access to political power and therefore how powerful or vulnerable they were politically.

The four categories of institutions characterised the institutional underpinning of the accumulation of resources in hospital and primary care sectors. Two points should be made before proceeding. First, focusing on these institutions does not mean ruling out the importance of other institutions (e.g. pharmaceuticals, technologies, information, etc.). Other relevant institutions were considered as attached to and dependent on the above-mentioned four institutions. For example, the availability of certain technologies at a certain level of care was considered as influencing coevolution of hospital and primary care sectors through enhancing the professional influence of such level of facility. Second, within the four categories of institutions, there would likely have been variations within a particular sector (for example, hospitals at different levels might face different policies). Given the diversity in China, any generic analysis is problematic. However, the focus on sectors was determined by the thesis's interests in the comparison between hospitals and primary care providers as two sectors, and how they converged or diverged from each other.

5.2.4 Chronological scope and periodization

Part II covers more than a century and a half in the history of the health delivery system in China. Periodization was crucial in making sense of such a long history, for allowing both an in-depth discussion of the most significant processes within each temporal segment, and the ability to make sense of the stages of the long-term process. The historical study originally divided the chronological scope into three main periods,

based on two important time points: the founding of the People’s Republic of China in 1949, and the Reform and Opening-up in 1978.

5.2.5 Issues about collection and critical reading of texts

There were methodological issues related to the nature of documents (see Table 5.2). A basic principle for documentary analysis is that documents may not be truthful records of what happened. Instead, documents are the central carriers of human-produced “documentary realities” (Atkinson & Coffey, 2004). Critical reading of the sources collected was conducted with attention to the following four aspects identified.

Table 5-2. Key issues of documentary analysis

• Authenticity: soundness and authorship
• Credibility: sincerity and accuracy
• Representativeness: survival and availability
• Meaning: literal and interpretative understanding

Source: based on (Scott, 1990: Chapter 1)

First, I considered issues related to the authenticity of documents. Validity of documents is a crucial concern in documentary analysis (Green & Thorogood, 2004: Chapter 7), Scott (1990: Chapter 1) distinguished authenticity and credibility under validity. There are two main ways to judge the authenticity of documents, namely, soundness and authorship. Soundness refers to whether the document is original or a good copy of the original document. Authorship concerns whether the documents was really produced by the person or organisation who claimed it. As most of the primary sources came from official archives or public libraries or archives with general soundness and clear authorship, authenticity of documents was not a concern for this thesis.

Second, I paid attention to the credibility of sources. Credibility means whether the account is distorted, consciously or unconsciously. To judge credibility, one has to check for sincerity and accuracy. Sincerity refers to whether the document’s producer

actually thought what was recorded to be true. This judgement usually involves an appreciation of the underlying motivation of the author—why the author chose to produce the document. Assessing accuracy is important as sincere authors could have documented things inaccurately due to limits of observation and documentation, and the proximity of the document producer to the event. Both “primary sources” (defined as reports by people involved in the events) and “secondary sources” (defined as reports by people who did not participate in the events but relied on second-hand information) can have issues of inaccuracy. Whenever possible, primary sources were collected and compared with secondary sources. The motivation of the document producers was also considered, while the documents were interpreted.

Third, the sub-study attempted to achieve representativeness of documents (Green et al., 2004: Chapter 7; Scott, 1990: Chapter 1). There might be issues of survival and availability regarding representativeness. Therefore, the sample of documents might be biased. Important historical events might not be documented, or might be lost during storage. This study concerned complex processes across a prolonged period of time, a vast country as well as numerous actors. Statistical representativeness was not achievable. The study addressed this issue by pooling a large range of national and local data. For each period, I tried to include data for the most consequential actors. The main issue in this regard was the dominance of the perspectives of those controlling power: government authorities, professional elites, and non-governmental organizations (particularly missionary and philanthropic groups). Following the practice of “heuristic” of early German historians (Scott, 1990: Chapter 1), I scanned through the full lists of the catalogues of archives to identify the underlying logic of the construction of archival documents to identify relevant documents that appeared to be representative of the materials. In the process of searching two local archives, it appeared to me that the archives mainly kept the official documents that were produced by the specific agencies. These documents included policy documents, annual work-

plans, annual work summaries, special statistical reports, budgets, summary and speech scripts and transcripts of participants of important conferences. They generally represented the key external bureaucratic processes, with relatively few useful materials regarding daily internal document-based communications, which potentially could have contained more credible information about important events.

Fourth, the meanings of documents were understood and interpreted in their context. The documents can be seen as “social facts” produced by a claimed authorship for an intended readership, with both the production and reception taking place within a social/cultural context (Atkinson et al., 2004). There are two levels of understanding of meaning. A basic level concerns literal understanding. Issues regarding this include the change of “terms” and dating (how dates were represented). Interpretative understanding concerns the definition of concepts and genre (style of the document: public documents or memoir). The sub-study contained concepts which had varied meanings over time and used a broad range of different genres of documents, and took such issues into consideration during the analysis.

More importantly, interpretation is a process of appreciating the meaning that the document producers (historic contemporaries or later writers) relied on to understand their experiences, the meaning that influenced the creation of the documents, as well as the meaning at the receiving end--the “three moments of text” according to Scott (1990: Chapter 1). The meaning involved is crucial in reconstructing the social world beyond the immediate documents. Researchers need a sufficient knowledge of the wider worlds of the historical actors to be able to carry out “explanatory understanding”. Therefore, I not only immersed myself in the various sources, but also went through a series of general historical books on the period, events and people studied. Attention was also given to the partiality of some primary sources, which were produced not by dispassionate researchers but by policy actors with a specific agenda and perspective.

5.2.6 Sources

The study used a range of different types of sources. The first were history books on history of medicine in China, which provided a general timeline to follow. Three particularly important longitudinal books included *the History of Chinese Medicine* (Wong & Wu, 1936)-- a chronicle of medicine-related events in China from ancient times to the present period, the *General History of Medicine in China* (Deng & Cheng, 2000) and the *Chinese Medical Encyclopaedia* (Chen & Xu, 1984). The *History of Chinese Medicine* was a case in point where I maintained caution while using sources, as one of its authors, Wu Liande was a passionate reformer and considered the founder of modern public health and scientific medicine in China. The motivation of co-authoring the book was likely to link Chinese history with his contemporary medical reform initiative, arguing that there was an inherent tradition of state-sponsored health care lost in history and to be revived with new knowledge. The works of other historians and social scientists who studied history on more specific periods or topics were also included. A wide range of contextual books was also consulted to allow interpretation of actors and events in their historical context.

Second, official documentations of history, statistics and compiled policy documents provided by the Ministry of Health (National Family Planning and Health Commission since 2013) were included. These consisted of the series of the *China Health Yearbooks* (published annually from 1983) which provided an official account of information on key events of that year, with the first volume (1983) including an overview of the years between 1949 and 1982. Policies related to rural health were drawn from a compiled book that contained policies from 1949 to 2010. The advantage of official chronicles and compiled documents was that they contained systematically collected key information on what happened each year. Therefore, such information provided a continuous timeline to make sense of the key events and development of key indicators, as well as key speeches by leaders of health authorities and/or state

leaders whose job division included health affairs. Three provincial health gazettes from eastern, central and western China (Shandong, Jiangxi, Guangxi) were selected to complement the national yearbooks (Guangxi Provincial Health Gazette Compiling Committee, 1998; Jiangxi Provincial Health Gazette Compiling Committee, 1997; Shandong Provincial Health Historical Gazette Compiling Committee, 1992). Local Health Gazettes, usually compiled by local health bureaux, provided a local official account of history.

Third, although national archives are inaccessible for post-1949 periods (Gao, 2010), local archives kept a rich source of detailed historical documents. For this study, I accessed two local archives: the Beijing Municipal Archive and the Pinggu (a district of Beijing and a rural county before 2001) District Archive. At the Beijing Municipal Archive, an online electronic catalogue was available. I searched all items containing the title of *yiliao* (medicine) and/or *weisheng* (health) and screened the resulting list for potentially relevant sources, and identified and took notes of relevant sources. Some of the documents were scanned and stored on an intranet system, and could be read directly on the computers at the archive. Others were not scanned, and I submitted a list of titles I would like to read each time. As there was a daily cap of 10 files to copy, I only made copies when necessary, and took notes of the other relevant ones. At the Pinggu District Archive, archival documents were kept for district-level agencies and access was limited by requiring official permission (in the form of an official letter of reference) from the agencies which generated the documents. There was no electronic system to search documents within the catalogue of the local health bureau. Therefore, I manually screened the catalogue books to identify potentially relevant documents. Then I asked the archive administrators to help retrieve them, read and made copies of the relevant ones. Archival materials were also searched on and acquired from a second hand book website (www.kongfz.com) which contained materials regarding the Maoist period.

Fourth, anthologies, biographies and memoirs of important actors and people who had experienced the history were collected and analysed.

Fifth, journals and newspapers were searched. Missionary publications, including the *Chinese Repository* and the *China Medical Missionary Journal*, were the earliest periodicals used in this study. For the early half of the 20th Century, the *Chinese Medical Journal* and the *National Medical Journal of China* provided longitudinally rich sources of information, though these journals generally reflected the viewpoints of elite Western medicine doctors (both Chinese and foreign). Newspapers and journals provided a rich volume of both primary and secondary sources of information that could be searched through electronic bibliographic databases (e.g. China National Knowledge Infrastructure). Journals also contained research and reflection on contemporary or historical events. The newspaper *Renmin Ribao* (People's Daily), the mouthpiece of the Central Committee of the Communist Party of China, preserved rich information regarding the political elites' perspective on health issues.

Finally, three propagandist paintings were also included in the analysis, to reflect the way the forces behind the policies sought to (re)construct their intended rhetoric.

5.3 Overall historical context (1835-1949)

China was a country closed to the outside world by the Late Ming Dynasty and the early Qing Dynasty (see Appendix Table 2 for a standard periodization of Chinese history). Through the First Opium War (1839-1842) and ensuing invasions, Britain and other foreign powers forced China to sign a number of unequal treaties, which opened ports and concessions, mainly in coastal regions (Fairbank, 1978: Chapter 4). The pro-modernization Self-Strengthening Movement was organized between 1861 and 1895 with limited success, ending with a humiliating defeat by Japan in the First Sino-Japanese War (1894-1895) (Fairbank, 1978: Chapter 10). Another wave of pro-modernization movement took place in 1898 and quickly failed, only to prompt the

Boxers' Rebellion against foreigners, which ended with the occupation of Peking (Beijing) by the Eight-Nation Alliance and the payment of a huge indemnity fund. The Qing Dynasty, and along with it the millennial cycle of imperial history, was ended by the 1911 Revolution that resulted in the establishment of the Republic of China. However, the country soon fell into fragmentation during what was known as the Warlord Era (1916-1928), and control of the country was divided among its military cliques (Fairbank, 1978: Chapter 6).

After unification through the North Expedition (1926-1928), a national administration controlled by the Chinese Nationalist Party (founded in 1912) was formed in 1928. With about a decade of peace and development, the country again fell into war due to the Japanese invasion in 1937 (Fairbank & Feuerwerker, 1986: Chapter 3). While the Communist Party of China (CPC) and Nationalists had joined up in the 1920s, the CPC was excluded from the Nationalists and purged, driven away from its base area in Jiangxi in 1927, and eventually to embark on the Long March (1934-1936) and resettle in the North West with Yan'an as its capital (Fairbank et al., 1986: Chapter 4). The country won the Chinese People's War of Resistance Against Japanese Aggression in 1945, and then a civil war broke out between the Nationalists and the Communists. Four years later, the Nationalist Government fled to Taiwan, while Chairman Mao declared the founding of the People's Republic of China at *Tian'anmen*.

Before the end of imperial history in 1911, the imperial court was a high conservative national state actor. The imperial state was very limited in both tax extraction and its public services, and the limited resources it extracted were spent primarily on hydraulics, as well as military defence and rebel suppression (Balazs, Wright, & Wright, 1966). After 1928, a national administration was set up in Nanjing. A nation-wide system of local governments was also established, with counties functioning as self-sufficient and significantly autonomized governments responsible for establishing and delivering local health services.

Traditionally, China had a large and overwhelmingly rural population. Research suggested that traditional China was already divided into an urban privileged population and a rural population (Kuhn, 2002). Industrialization was slow before 1949 and foreign trade benefited the treaty ports, which were highly concentrated in coastal areas. Most people were living in rural areas and in dire conditions, particularly during the 1930s and the 1940s, with a vast number of preventable deaths (Štampar, 1936). Foreigners and foreign organisations played an important role in China's modernization from the 1840s to the 1940s, as demonstrated below in relation to health.

5.4 Organisation of health services in traditional China (1572-1835)

In Western societies, charitable organizations, religious organizations, and medical practitioners, played central roles in the rise of modern health systems (Risse, 1999: Chapter 5). What China had developed by the period of encounter with Western medicine became an important default for policy actors. Examining this default therefore provides important clues to understanding the initial set up of modern medicine in China. This section traces the organisation of health services in traditional Chinese society before the introduction of modern medicine. Particularly, the section attends to the development of institutional care (including but also going beyond medical care) and the status of medical practitioners in the society.

Institutional care had a very long history in China and was documented as existing widely before the late Ming period (from 1572-1644). One of the essential teachings of Confucius mentioned an ideal society where “all widowers, widows, orphans, the childless, the handicapped are taken care of” (Anonymous, circa 200 B.C.). As early as the Han Dynasty (202 B.C. to 220), there was a record of hospitals built for epidemics (Fan, 1951). In the Epoch of Division of North and South (420-589), “house for people with the six types of ailments” (*liujiguan*) were built by the state (Zhang, 2006, p. 79). During the Tang Dynasty (618-907), such institutions were supported and

popularized by Buddhist activities and named “house of the sad field” (*beitianfang*) or “house for caring of the sick” (*yangbingfang*). They provided places for the sick and the destitute to stay (Leung, 1987). Then they were taken over by the imperial state when Buddhist monasteries⁷ were purged between 842 and 846 (Chen, 1991).

The state sponsorship of infirmaries continued in the ensuing Song Dynasty (960-1279), when the name of these institutions was changed into “wards for peace and relief” (*anjifang*). Besides hospitals were other types of state-sponsored welfare and medical institutions which developed during subsequent dynasties. Charity pharmacies started to distribute free medical formulas: e.g. “pharmaceutical bureaus for people’s benefits” (*huimin yaoju*) became popular in Song. The government also distributed free medicines through public pharmacies (the Imperial Medical Service (*taiyiju*)) and regulated drug prices (Taiping Prescriptions for the Benefit of the People from the Office of Pharmacy (*Taiping huimin hejiju fang*)) in the Song Dynasty. There were also medical agencies providing medical training, medical relief and professional regulation: e.g. “official medical promotion bureau” (*guanyi tijusi*), “universal aid promotion bureau” (*guangji tijusi*) and “medical bureau” (*yixue*) set up in the Yuan Dynasty. During Yuan, the state also widely built “Temples of the Three Sages” (*sanhuangmiao*) to promote the worship of the Yellow Emperor as patrons of medicine (Leung, 1987).

From the Ming Dynasty, state-sponsored organisations of health activities generally fell into decay. Compensating to the declining role of the state, local communities started to organize various kinds of benevolent societies with official endorsement since the Ming Dynasty. Some scholars set up private funds with donations from “societies for sharing goodness (*tongshanhui*)”, and created the so-

⁷ These infirmaries were located in monasteries, parallel with Japanese and Western societies where prototype hospitals were set up within religious premises.

called “universal relief halls (*pujitang*)” (Leung, 1987). Based on Leung’s study of local gazettes from the Yangtze River delta (a relatively rich area) before the Opium War (1840), almost every county seemed to have a benevolent society (Leung, 1987). In the late Qing, an expanding base of lower ranking Confucian scholars became the key sponsors behind local benevolent societies, and the network of benevolent societies was decentralized to township level (Liang, 2012).

The medical function of benevolent societies was very limited. As Lockhart (1861), a prominent medical missionary, observed in Shanghai, there was a hospital/alms-house for the old and infirm and a foundling hospital, with little medical function. The only “establishment for gratuitous medical relief” (*She-e-kung-keuh* [sic]) was created after the establishment of a missionary hospital in 1845. Wong et al. (1936, pp. 137-138) also commented that the Chinese prototype of a hospital “was more of the nature of a benevolent institution where there was very little medical attention or trained nursing”. Study of local gazettes showed that these organizations might have some basic medical function during times of epidemics, including segregation of the seriously ill, as well as government subsidized medicines and provision of doctors. However, the nature of these efforts was more disaster or epidemic relief than routine medical care delivery (Leung, 1987). This low level of medicalisation (from both a modern perspective and a historical perspective before the late Ming) probably reflected the preference of their sponsors. Benevolent societies became a vehicle for the vast number of low ranking scholars to reinforce their values and acquire social influence, and also suited the government’s need of societal control with minimal state responsibility (Liang, 2001). In her extensive study of local gazettes, Leung (1987) found only one dispensary with a sick ward, which she suggested was quite an outlier.

In traditional China, there was generally no clear boundary between medical professionals and lay medical practitioners, and the social status of doctors was

generally low. In the Song Dynasty, almost none of the doctors in Fuzhou prefecture, Jiangxi were from notable families (Hymes, 1987). In the Jin Dynasty, a central level Imperial Medical College (*taiyiuan*, also translated as the College of Royal Physicians) and local level medical bureaux (*yixue*) were responsible to promote both professional qualifications and popularization of officially verified medical literature. However, except for strict training for doctors for the imperial court, implementation of professional training was plagued with fraud and corruption.

One piece of evidence for the low social status of doctors lay in the fate of the Temples of the Three Sages of the Yuan Dynasty mentioned above. Among the three sages was the Yellow Emperor, who was worshipped as the ancestor of the Han people (the majority ethnic group). To worship the Yellow Emperor as a key patron of medicine appeared to elevate the status of doctors. However, due to the low prestige of medical practitioners, this move of the Mongol rulers was considered a demeaning act to the Han people's worship of the three sages, and was quickly abolished in the Ming Dynasty (Zhang, 2004). The kind of professional societies like guilds that formed a monopoly of medical training and disqualified lay people (Ramsey, 1992; Wear, 1992) did not develop in the Ming and Qing Dynasties.

By the Ming Dynasty, the medical bureaux had become basically inactive and their medical officers tended to be poorly trained and county-level medical officers were not salaried. In addition, the Ming Dynasty saw rapid popularization of medical literature, which led to a further decline of the social status of medical doctors. Medicine was an alternative of lower status for scholars who had no prospects as officials or landlords (Leung, 1987). Official medical practitioners tended to be poorly trained and with no prospects unless elected into the Imperial Medical College. A local official proposed that reputable doctors should be trained to revitalise medical bureaux to supervise the practice and training of local practitioners to elevate standards (Liang, 2012, p. 46), suggesting the weakness of existing professional institutions. To the

missionaries coming later, a Chinese doctor suffered from lack of trust by their patients, and except in rare cases (e.g. notable families), “never becomes a family physician” and “knows quite well that unless his prescription has given immediate relief a second and perhaps third physician will be called in after him, almost as soon as his medicine has been sampled” (Balme, 1921, p. 28). Along with learned doctors, there were a broad range of relief providers including demonological healers, Buddhist monks, itinerant herb vendors, priests, mediums and illiterate healing women (Andrews, 2014; Chapter 2).

Opportunities for potential revolution of medical knowledge were missed. In the 1600s, Jesuit missionaries brought with them several medical titles including Andreas Vesalius’ treatise on human anatomy. From 1622 to 1643, in the last years of the Ming, Terrenz translated a short text on anatomy into Chinese with the title of “Western explanations of the human body (*Tai-si jen-shen sho-kai*)” based on the European anatomy book of *Anatomica corporis virilis et muliebris historica* (1597). The translation was not disseminated. The Jesuits won the support of the open-minded Kangxi Emperor for Western medicine by curing the Emperor’s malaria with cinchona bark as early as 1693. In 1698, the court asked the Jesuits to translate two anatomy texts into Manchurian. Five years later, the translation ended up being shelved in the imperial library, as its materials were considered inappropriate for publishing (Fu, 2011a). In the *History of Chinese Medicine*, Wong et al. (1936) lamented that the failure of the Jesuits to “grasp the opportunity offered to establish proper hospitals and medical schools” was a loss of a valuable historical time window for the popularization of Western medicine in China. While this remark was apparently made from a perspective similar to teleological history (i.e. assuming events were moving (or not moving) towards a desirable end-point), it did reflect the difficulty for Western medicine to develop based on intellectual enlightenment in the conservative climate of the late Ming and the Qing.

The section reveals a situation where official interests in medical affairs withered over time, despite what appeared to be a long previous tradition of health protection and benevolent activities for the weak and sick. The conservative state also refrained from publishing unorthodox new medical knowledge from the West. By the 1830s, the sponsors of the benevolent societies were mainly lower-ranking Confucian scholars with official endorsement, whose main interests were to conserve traditional values and rituals for their own values and status, while medical services if provided at all by these societies was more of the nature of rituals. Medical care was rarely hospital-based in traditional Chinese society. The under-development of medical professional boundaries, popularization of medicine, and the low social status of doctors undermined the role of medical practitioners in influencing medical organisation and development, perhaps except for the Imperial Medical College. All these not only made indigenous adoption of new medical knowledge and practice difficult by the late Qing, but also became the context for the newcomers, who would need to carve out a new arena for their medical practice.

5.5 Introduction of Western medicine (1835-1928)

The state and Chinese practitioners of medicine were generally uninterested in Western medicine around 1835, and the introduction of modern medicine was taken on by foreigners, particularly Protestant medical missionaries. In consequence, the agenda of these foreigners at this specific historical period of early engagement with China largely decided the initial set up of providers of Western medicine in China, and thus the formation of the early institutions behind the coevolution of hospital and primary care sectors. Also important was how domestic actors, particularly the state, reacted to the activity of missionaries. In the later years (1901-1928), philanthropic and other groups also joined this movement of Western medicine, and responded to and redefined these initial structures. Therefore, understanding the perspectives and influences of these late-comers is also important for analysing the historical dynamics.

5.5.1 Introduction of mission hospitals (1835-1844)

The years between 1835 and 1844, i.e. from the establishment of the Canton Hospital to the signing of the Treaty of Wangxia after the First Opium War, seemed to be a critical juncture where the seeds of hospital centrism were sown. The initial driving force behind the medical mission appeared to be a strong desire to establish civil engagement, which was a particular concern in the years approaching the outbreak of the First Opium War (1840-1842). By the early 1830s, a Protestant missionary foothold was established in Guangzhou (or Canton, the single port allowed for commercial trading before the Opium War), led by Bridgman of the American Board of Commissioners for Foreign Missions. Dr. Colledge, once a ship's surgeon for the British East India Company, operated a successful ophthalmological hospital in Macau (*The Chinese Repository*, 1834, p. 289). Colledge was convinced that eye diseases were prevalent, however, the indigenous doctors had no measure to deal with them, and surgery was not available to even the rich in China (Wu, 2000, pp. 301-310). In a summary of the initial years of medical missionaries in China, Lockhart (1861, p. 113), another leading figure of medical mission noted:

“The practitioners of surgery, or those who attend to external [sic] diseases, are of a lower grade than those who treat internal affections, chiefly because surgical manipulations are little understood. None but the most trifling operations are attempted, and, in fact, small attention is paid to this branch of practice, because the relief afforded is so insignificant.”

Encouraged by the example of Colledge, Bridgman proposed the idea of a medical mission and sent Peter Parker to China. Parker established the Canton Ophthalmology Hospital in 1835. The focus on operations on eyes (including cataract surgery) proved a right decision. Parker's report of the cases he treated during a quarter of 1836 showed that the hospital had treated 323 cases with eye diseases, 32 cases with ear diseases, and 66 other cases (Parker, 1836). Hospitals were important as a venue to demonstrate

the technical advantage of Western medicine, particularly in surgery.

More importantly for the foreign powers, the success provided a channel of civil interaction. The American Board of Commissioners for Foreign Missions commended that “Dr Parker opened the gates of China with a lancet when European cannon could not heave a single bar” (Balme, 1921, p. 44). Colledge, Parker and Bridgman proposed the “Medical Missionary Society in China” to promote the medical missionary activities; they explained that the particular “good effects” of the medical mission were “to bring about a more social and friendly intercourse between them [Chinese] and foreigners, as well as to diffuse the arts and sciences of Europe and America, and in the end to introduce the Gospel of our Savior in place of the pitiable superstitions by which their minds are now governed” (Colledge, Parker, & Bridgman, 1835). This proposal prompted the establishment of the Society in 1838 (Lazich, 2006). After the Opium War (1840-1842) which further opened China to Western powers, Peter Parker called for “the maintenance of the Hospitals already established, and for the founding of others at every accessible and eligible part of China” (Parker, 1842, p. 4). The medical missionary enterprise expanded further, particularly after 1844, when the signing of the unequal Treaty of Wanghia allowed Americans to open hospitals, among other institutions, in five ports in China.

The medical success of hospitals should not be exaggerated. Chinese were not used to hospitalizations and were quite suspicious of the motivation of medical missionaries, and the initial patients were not really hospitalized. After all, pre-20th-Century Western surgery would have been both painful and dangerous, as this was the period before the invention of antiseptic and popularization of anaesthesia (Ellis, 2002). Indeed, the documents for the early years of the Canton Hospital suggested there were barely any record of hospitalizations (Liang, 2013).

5.5.2 Further development of mission hospitals (1844-1901)

The initiation of this model of hospital-based medicine appeared to set in motion a post-juncture self-reinforcing process. As Renshaw (2003) pointed out, few medical missionaries had worked at hospitals before coming to China, nor were hospitals the dominant form of medical practice back in their home countries. The missionaries' passion for building hospitals and the way hospital care was provided were rooted in both their missionary purposes and the need to "accommodate" Chinese.

As the medical missionaries tried to expand their niche, they assessed different ways of providing care. Hospitalizations provided lengthy interaction between patients and missionary staff. As the ultimate agenda of the medical mission was to win over converts, the lengthy engagement of hospitalizations seemed ideal for preaching. For example, it was a world-wide practice among medical missions to extend hospital stay as much as possible to increase the length of religious interactions between providers and patients (Hardiman, 2006). The recommended protocol for organizing outpatient care at that time (Balme, 1921, pp. 72-74) also encouraged missionaries to mingle with patients, get to know them one by one, win their confidence, preach to and follow up with them in the hospitals. In other words, the outpatient department was a platform important for the hospitals to reach out to a wider range of potential believers.

Table 5-3 summarizes the analysis of the pros and cons of different delivery options in the report of the China Centenary Missionary Conference (Conference Committee, 1907, pp. 256-257), which convincingly explained the reasons behind the preference for a hospital-based medical missionary. From a missionary perspective, hospitals provided the most effective means of religious engagement as they treated the most severe patients, for whom religious work was thought to be more influential. The spiritual work was considered to achieve maximal impact effective in hospitals even if medical intervention failed. It provided a place for continuous interaction,

which was usually missing in an ambulatory setting. Indeed, the hospitals were so much valued that dispensaries should be built mainly “to feed hospital(s)”. These missionaries “universally agreed that the hospital is the centre, and that all other work must, to a large extent, lead to or proceed from it” (Conference Committee, 1907, pp. 256-257).

Table 5-3. Comparison of the three types of medical missionary work

	Itinerant (i.e. Mobile)	Dispensary/ outpatient work	Hospital
Advantage	Extended reach, in many cases the only way to approach some populations (particularly home-bound women)	Deeper impression on beneficiaries.	As diseases treated at hospitals were more severe than in ambulatory settings, hence medical mission's influence on the minds would be “much greater”. Every patient, “whether a cure is effected or not”, experienced close and lasting contact with staff and the atmosphere of the hospitals. When properly staffed, it could minimize the risk of breakdown of doctors, and provide continuity in case of holidays or sickness, etc.
Disadvantage	“very expensive”, not satisfactory professionally because of inability to follow up	Ineffective if patients did not return.	
Values	Indispensable in the early stages	The greatest use was “to feed the hospital”. It gave “much better results” than itinerant work.	“most satisfactory and fruitful part of the work”
Implementation	In many places, workload from hospital and other station work made it impossible to conduct itinerant work.		A small number of well-equipped and supported hospitals would be more effective than several separated ones.

Source: author’s summary of (Conference Committee, 1907, pp. 256-257)

This hospital-centred model of care provision had enduring impacts on the

coevolution of hospitals and primary care providers in China by providing a new venue, a new organization, as well as a new profession. First, in terms of professional institutions, a professional link formed between Western medicine doctors and mission hospitals. To address the tension between professional and missionary activities, separation of their roles was encouraged (Lockhart, 1861, p. 119), which was the first step of developing a medical professional identity. The doctors were also tied up with hospitals. Doctors were not able to spend much time paying visits to patients at home (who paid much more generously) due to the workload at hospitals, as reflected later in the report of the China Medical Commission of the Rockefeller Foundation (1914, pp. 74-77).

Second, in terms of organizational institutions, these hospitals integrated a large proportion of outpatient care. Outpatient work was frequently described as a reason that kept medical missionaries busy. The ambulatory department was both the main access point for patients to seek care at hospitals (Renshaw, 2003), and the dominant access for hospital-based medical missionaries to interact with local communities. The fusion of Western medicine doctors and open-access outpatient departments of hospitals formed a structure that has persisted all the way to today. Later sections will visit the other two pillars of the institutional complex underpinning the coevolution of hospitals and primary care providers.

During the most part of the period covered in this sub-section, the state barely sponsored any Western medical education. Indeed, an effort to introduce Western medicine at national level was nipped in the bud in the latter half of the 19th Century. The Imperial Tungwen College, also known as the School of Combined Learning, was founded in 1861 during the Self-strengthening Movement (1861-1895), as an imperial school in Beijing under the Ministry of Foreign Affairs to introduce Western science and technology. William Alexander Parsons Martin, a missionary and translator was recruited as the President of the Imperial Tungwen College from 1869 to 1895. Martin

(1896, p. 320) wrote one year after leaving the presidency on a failed attempt to organize Western medicine education at the college, revealing the political resistance from vested interests within the imperial bureaucracy.

“Many years ago we formed a medical class, which was placed under Dr. Dudgeon, of the London Mission, who was and continues to be the best-known practitioner in the northern capital... The Yamen [meaning the imperial court] gave him [referring to Dr. Dudgeon], as I proposed, the title of professor, and invited him to lecture, but refused to permit our students to receive clinical instruction at the mission hospital. Ten years were thus lost, the lectures amounting to nothing more than the communication of ideas such as ought to form a part of a liberal education. A change of ministry occurring, I again proposed that the class should receive practical instruction at the hospital. The new ministers consented, but they declined to expand the class into a medical school for fear of encroaching on the domain of the Tai-i-Yuen [the Imperial Medical College], an effete college of medicine which has charge of the emperor's health and is supposed to possess a monopoly of medical science.”

Based on Martin's account, the opposition of conservative bureaucrats managing the Imperial Tungwen College and the royal doctors in the Imperial Medical College repeatedly frustrated his attempt to establish systematic Western medical training. In parallel, medicine was also neglected when the court sent students to learn science and technology in Western countries during the Self-strengthening Movement (Li, 1990, p. 59). Before 1900, the government established only one medical school, in Tianjin, under the pro-reform Viceroy Li Hongzhang in 1881--the Viceroy's Hospital Medical School led by an English medical missionary Dr. Mackenzie. Yet by 1888 when Dr. Mackenzie died of smallpox and the school was closed due to the ending of funding (Ho, 2012), there were only nineteen graduates, all of whom “did not get posts commensurate with their training” (Hume, 1922). Based on documents written by

medical missionaries examined by Liang (2013), benevolent societies in China seemed to be inactive during this period, except for only a very small fraction mimicking the practice model of hospitals, with most still providing just Chinese medicine. More than a decade later, the China Medical Commission of the Rockefeller Foundation (1914, p. 8) reported that “[o]utside the mission hospitals and outside the treaty ports there were very few practitioners in China who have had any training at all in Western medicine and almost none who have been adequately trained”.

The success of medical missionaries was limited, in both medical and missionary terms. They won few converts (Yang, 2006, pp. 1-44), and did not manage to establish fully the technological advantage of Western medicine (Balme, 1921, pp. 60-81). As Parker noticed, patients went only to Chinese doctors when they sensed acute illnesses, and tended not to visit a Western medicine hospital unless they needed surgery (Liang, 2013). This was supported by Chen Zhiqian’s memoir, which mentioned that people in the 1910s generally believed Western medicine to be good at surgical cases, while traditional Chinese medicine was more trusted by the Chinese people when it came to non-surgical conditions (Chen, 1989, p. 20). While helping the Rockefeller Foundation design its strategy for health philanthropy in China, Welch commented that “Chinese appreciate Western surgery, but do not seem to consider Western drugs superior to their own” (Bullock, 1980a, p. 39). The China Medical Commission of the Rockefeller Foundation (1914, pp. 74-77) also pointed out that “[a]lmost all the foreign physicians have been interested primarily in surgery”, but even surgery was limited to the “simple and obvious” cases due to distrust of patients and potential risk of hospitalization “not safe in performing any operation which might end fatally”, and abdominal surgery was rarely performed.

The Canton Hospital still relied on eye surgery to attract patients in 1894—60 years after its establishment (Liang, 2013). Liang (2013) studied Chinese and Western medicine hospitals in Guangdong and Hong Kong in the late Qing period (1840—1911)

and showed that Chinese benevolent societies and medical providers had started to adopt the model of hospital-based care (with Chinese or Western medicine). However, in a hospital that used both Chinese and Western medicine, which the China Medical Commission of the Rockefeller Foundation (1914, p. 75) visited, patients exhibited clear preference for native medicine over Western medicine when given both options.

As springboards of Western civil engagement with China, mission hospitals received support from religious organisations back home. In terms of governance and financing, the mission hospitals were primarily accountable to their home churches, though a significant proportion of funding would have come from local donations and paying patients. Hence, hospitals by 1901 were above all spiritual enterprises, focusing on saving souls, while medical services as an act of body-healing (curative by nature) were a secondary purpose. Despite a small reach and the limited medical function that to an extent resembled the native benevolent societies, these mission hospitals established an important transition in the Chinese medical world by not only building a new type of medical organisation but also embedding a new medical profession in hospital-centred facilities.

5.5.3 Consolidation of hospital-based professionalization (1901-1928)

The years between 1901 and 1928 saw further post-juncture strengthening of hospital-related institutions. As the country was in continuous chaos, e.g. the Boxer Rebellion (1899-1901, also known as the Boxer Uprising), the Xinhai Revolution (1911), and the Warlord Era (1916-1928), there was barely any national coordination of medical development by the government. Progress in establishment of health services under the state was limited to either local initiative or selected sectors (e.g. railways and customs). A number of medical schools under missionary organisations were established, reinforced by new sources of money, particularly brought first by the indemnity for missionary hospitals after the Boxer Rebellion, and then by the

establishment of the China Medical Board by the Rockefeller Foundation (Deng et al., 2000; Wong et al., 1936). Emboldened by the newly found resources, and pushed by the rapid medical development in Western countries and China's engagement with Japan, medical professionals started to demand modernization of hospitals and became increasingly assertive of professional values. As this sub-section argues, the arrival of the Rockefeller Foundation amongst a number of emerging trends emboldened a movement towards scientific medicine and further consolidated hospital-based professionalization of modern medicine.

The unusual timing of the entry of the Foundation deserves attention. Around its entry to China, the Foundation was leading a far-reaching makeover of medical education in the United States (US), its home country, focusing on raising standards in a perceived situation of over-supply of doctors. In 1910, Abraham Flexner conducted a survey of medical education in US and Canada. The result was a report that classified American medical schools into three groups: the group that came out on the top were of a level similar to Johns Hopkins; the second group were substandard schools which required substantial reinforcement; the last group were of poor quality and deserved closure or consolidation (Flexner, 1910). Along with the following enactment of state licensing laws for medical schools and the carrot of funding support, in which the Foundation played a critical role, the Flexner Report triggered a fundamental medical reform through the US based on the Johns Hopkins standard. In the following quarter of a century, more than half of medical schools in America were closed (Cooke et al., 2006).

The Foundation transplanted the same vision of high standard medical education to China. In 1914, the Rockefeller Foundation sent the China Medical Commission to investigate the situation of health and medicine in China and to give recommendations regarding the Foundation's activities in China. They visited 17 medical schools and 97 hospitals, which they believed covered the most important and representative ones in

China. The conclusion was to focus on high standard medical education, including building two medical schools (one in Beijing and one in Shanghai), and providing assistance to medical schools in Guangzhou and Changsha. In terms of hospitals, the Foundation provided aid to mission hospitals to hire additional doctors and improve diagnostics and other services, in the hope that these would prepare them for clinical training of medical students (China Medical Commission of the Rockefeller Foundation, 1914, pp. 92-98). An illustration of the high standard adopted was the overspending on building palatial premises for the Peking Union Medical College (PUMC). PUMC opened in 1921, and its cost was among the main reasons why the medical college in Shanghai was never built. Besides luxurious facilities, the Foundation was also hiring world-class professors with very generous salaries and recruiting students who could afford a high tuition fee (Bullock, 1980a: Chapter 4).

Such an elitist strategy could appear problematic (and was criticized later on) as the situations of US and China could not have been more different. The fact that the Foundation's strategy was warmly received by some leading officials and medical reformers tells as much about the change going on in China as about the Foundation's preference. While there was considered an oversupply of doctors in the US (1 doctor for every 568 people), in 1910, there were fewer than 500 trained doctors in China, and only between 45 and 55 were Chinese doctors trained in Western medicine serving more than 400 million Chinese (China Medical Commission of the Rockefeller Foundation, 1914, p. 8). Despite the apparently opposite situations of accumulation of Western medical doctors in China and US, the Foundation's strategy resonated with the Chinese *zeitgeist*. The "May Fourth Movement" took place in 1919 and evolved into the New Cultural Movement, upholding science and democracy as key approaches to modernization while traditional values were condemned. With idealism in the air, Chinese medicine came under fierce criticism. In 1922, the leading leftist writer Lu Xun published the popular book *Outcry* (*Nahan* in Chinese), in the preface harshly

criticizing Chinese medicine doctors as “either unwitting or deliberate charlatans”. Scientific medicine at the high standard promoted by PUMC was accepted by Chinese officials, some of whom praised the Foundation for establishing scientific medicine (Bullock, 1980a: Chapter 1). Besides it appeared also a difficult time for developing other aspects of medicine, the Foundation dropped an idea of large scale public health effort, following the advice of the China Medical Commission of the Rockefeller Foundation (1914) which considered such effort not feasible in a fragmented nation.

More importantly, the vision, and money of the Foundation reinforced the pressure towards modernization of missionary hospitals, which was already accumulating among medical missionaries. The reflection of a Canadian medical missionary Dr. McClure captured this momentum:

“The big problem in the early days of medicine in China really was how scientific you would be. This was an ethical problem with the individual doctor and it was always a struggle. For instance, every time you raised your standards, you raised the cost and therefore the price to the patient. As medicine came into China, you first of all had antiseptic medicine, because that was what was being practised in Canada at the time. The sterilized sheet, the drapes, and that had not yet come into Canada. Then as antiseptic medicine was accepted in Canada and rapidly followed by aseptic medicine, with sterile techniques and scrubbed hands and sterilized drapes, you had this tremendous struggle. Should the mission hospital adapt and enter the new field that it made available? The new field of abdominal surgery and that sort of thing. At each of these stages of progress, some doctors opted out. They said, ‘We came here to preach the gospel and use medicine as a vehicle: it’s not necessary for us to use a scientific medicine.’ Very few did opt out, some merely dragged their feet.” (Austin, 1986, p. 168)

The Chinese Medical Missionary Association (CMMA)—an association covering 80% of all missionary hospitals in China, produced a report, seemingly to expose the poor conditions of these hospitals. Just to list some indications of poor conditions from the report, only 25% of hospitals had good sanitary arrangement, 34% did not have a pressure steriliser, 62% did not have night nursing, and 34% did not even have a nurse. This situation was then reported through a book published in London seeking more

assistance in terms of financial and human resources (Balme, 1921, pp. 82-106). Balme's report, published in the same year as the opening of PUMC, was likely to have been emboldened by the example of PUMC, and demonstrated a shift of strategy from focusing on "quantitative aspect" (e.g. services and winning converts) towards focusing on "qualitative aspect" (e.g. equipment and teaching). Balme expressed in the report the anxiety of medical missionaries and warning on sponsors:

"mission hospitals are now no longer the only hospitals known to the educated classes in China... [I]f... the mission hospital is not to become discredited in the eyes of the educated Chinese from a professional standpoint..., it must attempt to keep pace with modern developments—in other words it must become, in fact as well as in name, a Christian hospital [sic]".

The Foundation's annual report in 1922 showed that it had been supporting a number of missionary hospitals (Rockefeller Foundation, 1922) (see Figure 5-2). John Stanley's study of Weixian in Hebei also showed that the coming of the Foundation brought both new funding and new pressure on missionary groups to raise the scientific standard of mission hospitals (Stanley, 2006). Medical graduates from prestigious universities also set up locally accommodated copies of their alma maters—"Harvard-in-China, Yale-in-China, and Oberlin-in-China". For example, graduates of the University of Toronto established, built and staffed the medical college of the West China Union University (Austin, 1986, pp. 169-170). Such governance institutions alienated the universities from local society and reasserted professional values.

The missionary conference also allowed hospitals to charge patients money (Renshaw, 2003). According to the China Medical Commission's report mentioned above, the mission hospitals had most of their non-staff expenditures covered by local sources by 1914. The high degree of financial self-reliance entailed the concentration of mission hospitals in urban and coastal areas where there was a rise of middle class and richer people. This was reinforced by the Foundation's focus of financial assistance to areas which were perceived as strategically important by the 1914 report

(China Medical Commission of the Rockefeller Foundation, 1914, pp. 41-52). In other words, the overall accumulation of Western medicine was slow and highly inequitable geographically.

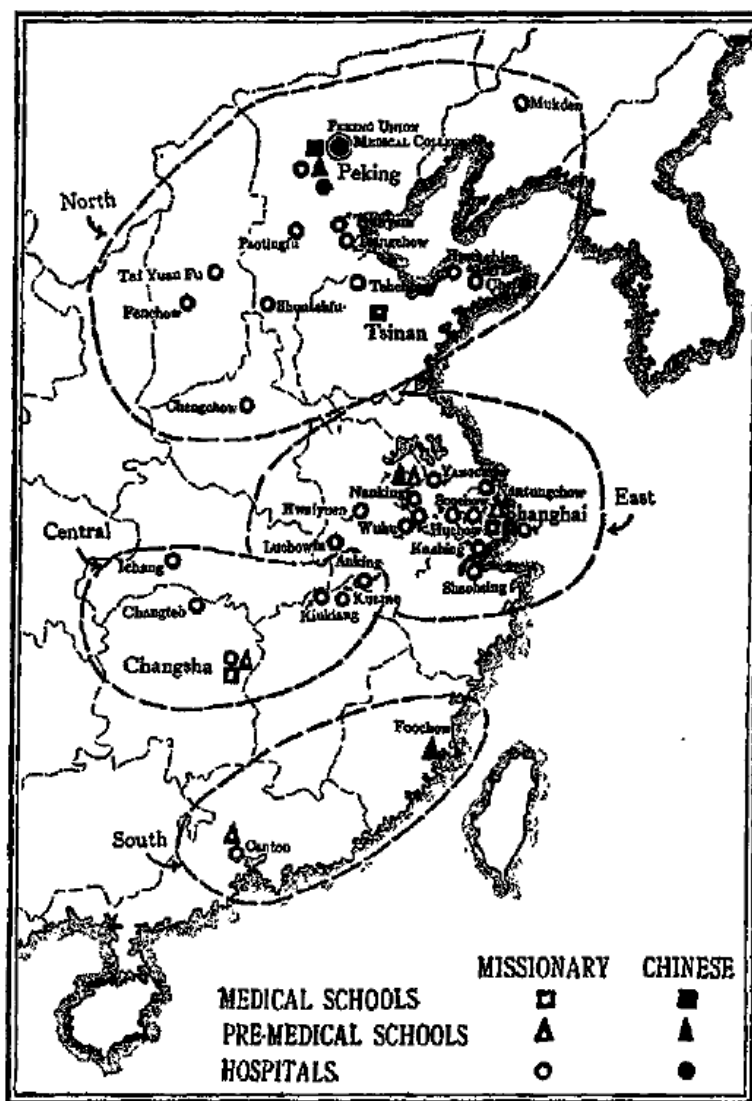


Figure 5-2. Location of institutions receiving aid from the China Medical Board

Source: (Rockefeller Foundation, 1922)

Note: the China Medical Board was a main funding agency providing financial assistance in medical and health affairs in China from the Rockefeller Foundation from 1914, and later became independent in 1928.

The consolidation of high standard teaching hospitals resulted in a concentration of relatively well-trained Western medicine doctors in hospitals. Building upon the

establishment of mission hospitals, this was a second step forward in hospital-based professionalization. At least from the perspective of Western medicine advocates, the central role of hospitals was obvious (Yin, 2013, p. 186). An article in a weekly popular Chinese medical journal sponsored by a student association of PUMC promoted the advantages of hospitals in both quality of care (attention to patients' conditions, quarantine against infections, etc) and professional development (development of specialities, use of sophisticated technologies) (Anonymous, 1928). As the attention of medical modernization was focusing on mission hospitals in selected areas, medical care for the vast majority of Chinese was provided, if at all, by Chinese medicine practitioners of various sorts (China Medical Commission of the Rockefeller Foundation, 1914, pp. 6-8).

The initial introduction of mission hospitals between 1835 and 1844 seemed to have set in motion a dynamic process that led to the step-wise development of hospital-centrism. First emerging as a venue to showcase technological advantage and establish civil engagement, hospitals then became as a venue to preach, and later as a venue to train and employ modern medical professionals. By 1928, a hospital-based profession of bio-medicine had truly emerged, with all four institutional pillars distinct from the primary care sector based on Chinese medicine. As to primary care, there was no evidence suggesting much indigenous change. Therefore, it was a period of radical contrast, with islands of modern hospitals in an ocean of Chinese medicine practitioners.

5.6 Two tracks of health services development (1928-1949)

The state has so far been playing a minor role in medical services, arguably limited by the state's capacity. By contrast, the period between 1928 and 1949 saw the state entering the centre of the policy stage. With this transformation of the state's role in health, understanding how the state actors came to see the importance of primary care

and the strengths and limitations of state capacity in introducing changes became crucial to analysis of the coevolution of the two sectors. Using the path dependence analysis framework, this section focuses on whether (and in what way) this transformation led to a conjuncture of major change in coevolution of the two sectors of hospitals and primary care providers, and after this conjuncture, how such policy changes were actually implemented or modified, and whether (and in what way) the process led to a more balanced development between primary care providers and hospitals.

5.6.1 The emerging agenda of state medicine (1928-1934)

The period between 1928 and 1934 constituted a conjuncture and witnessed a major turn in the pervading agenda for medical development from focusing on high standard scientific medicine to a new vision of providing modern medicine to the people. Critical in explaining this change seemed to be four interrelated factors: the establishment of state-level health authority, the active role of key reform advocates, the contributions of international exchanges, and the development of local experiences. In 1928, the Nationalist Party of China (also known as Kuo-min Tang) put an end to the chaotic Warlord Era and unified China. Around the time of the establishment of the Ministry of Health, there was tension between graduates from an Anglo-American style of medical education and graduates from a German-Japanese style of medical education. In what appeared to be an effort to build consensus, Grant pushed for China's engagement with the League of Nations Health Organisation (LNHO) (Bullock, 1980b: Chapter 6).

The LNHO also offered important support for China to reclaim its sovereignty in quarantine services previously controlled by the British, and engagement with the LNHO was particularly valued by the Nationalist Administration (Yip, 1992). There was clear enthusiasm also from the side of LNHO, as its engagement with China

greatly enhanced its status in helping design a national health system (Borowy, 2009). The technical assistance from LNHO buttressed a domestic call for state medicine with international consensus and experiences. It should be noted that this was the heyday of social medicine. Štampar--a pioneer of social medicine--was sent to China, among other technical advisers, to help China establish its own health system (Grmek, 1966, p. 16). Štampar's experience in designing the Yugoslavian system was also introduced to China via LNHO (Borowy, 2009; Zylberman, 2004). The importance of international experience, and particularly the example of Yugoslavia, in the advancement of the agenda of state medicine was reflected in an article by Wu (1931). Arguing that the growth of public hospitals was replacing privately practising general doctors in the West, and acknowledging the limited number of qualified doctors (12,000 for a population of 450 million) in China, Wu said that

*“a more economical and perhaps more effective way of meeting this deficiency lies in the establishment of state and municipal clinics and health centres for the people, such as have been instituted in Jugo-Slavia [Yugoslavia] ... Only trained experts in limited numbers need then be employed, each of whom can treat a large number of cases with more satisfactory results than if the same had been treated in private consulting rooms”.*⁸

Lending crucial empirical support for the plan were several local pilots, in particular Chen Zhiqian's experiment in Ding County. Chen was a favourite student of Grant at PUMC, and had been an activist for state medicine (Bullock, 1980a: Chapter 7). In 1932, Chen took charge of the Department of Rural Health of the Mass Education Movement in Ding County, an ordinary poor agricultural county. Chen designed the rural health programme to address the villagers' primary health problems which were mainly infectious diseases, e.g. smallpox, trachoma, dysentery, tetanus,

⁸ Health centres during this period were actually more than small hospitals.

typhoid, etc., and to satisfy the low cost needed to win over the villagers who spent only 30 cents on average (based on an earlier survey in Ding County) on traditional medicine (Ch'en, 1933). While other local experimentations were organized, Chen's experiment in Ding County was influential in turning the agenda of medical development towards rural areas (Lei, 2014: Chapter 10).

Overall, the structure of the Ding County health programme included vertical health professional supervision and horizontal multi-sectoral community reconstruction (Grant, 1963 (1934)). The concept of community was critical in the Ding County model, which was represented best in the central role of village health personnel in the health programme. The village health personnel were selected from graduates of a local school organized by the Mass Education Movement and became village health workers after a short training course. The training lasted for about ten days and covered registration of births and deaths, small-pox vaccination and first-aid (Ch'en, 1933). Villagers organized through the school provided the supervision. Chen's experiences showed that the minimal training was enough to address the most basic health needs and provide the most basic functions (disinfection, vaccination, etc.), while taking into consideration economic sustainability. Chen also acknowledged repeatedly that the expansion of health services in Ding County was built on organized local community efforts and "depends on the effectiveness of not only health programme alone, but of the all-round reconstruction programme, raising the general intellectual and economic level of the people" (Ch'en, 1933, 1934).

The other concept was that of a well-coordinated organization of health services (Ch'en, 1934). Upwards from the core formed by village health workers, were the sub-district health stations with doctors graduated from provincial medical schools, whose "first duty" was to train and supervise village health workers regarding both curative and preventive medicine, besides attending daily clinics (Ch'en, 1933). Chen put great emphasis on the supervisory role, and even criticized that the Yugoslavian practice was

inadequate in that there was no obvious professional supervision of village-level health workers (Chen, 1989); he argued that the training of medical school graduates needed to be reformed to enable them to work in rural areas. At the top of the technical ladder was the county health centre, which combined a hospital with fifty beds, a laboratory as well as administrative roles, and which coordinated and supplemented the activities of sub-district health stations. Originally, the hospital only accepted patients referred from sub-stations (Ch'en, 1933). Later about 50% of the inpatients were admitted through referral (Ch'en, 1937b).

With these converging factors, state medicine soon became a national policy guideline in 1934, as the Public Health Technical Conference proposed implementation of state medicine (Jin, 1946): “For a unit of 5-10,000 population, there should be a rural sub-station to take care of simple medical and health work, and for every five or ten such sub-stations, there should be a district health station to render more fundamental health and medical services to the people. In a *hsien* (county), there should be a health centre which includes a hospital, a simple laboratory and an administrative organ for the supervision of medical and public health work under its jurisdiction.” (Wong et al., 1936, p. 801).

Overall, the introduction of the plan for state medicine constituted a major turn not only in terms of health policy but also in the broad movement of medical development. First, state medicine involved the state’s responsibility, particularly for the previously ignored rural population. Second, attention shifted from high standard (Western) medical professionals to a more equalized distribution of a workforce sustainably trained and recruited. Third, there was consensus around the combination of preventive care and curative care. Within this was an assumption that low-cost preventive care should be prioritized. Fourth, village-level autonomy and reorganization was considered to be needed to support, deliver and supervise services. Fifth, a technical hierarchy would be needed for both supervision and technology

transfer. This change in the prevailing mind-set was well reflected in an editorial of the *Chinese Medical Journal* (Zhi, 1928), which proposed priorities for the newly founded Ministry of Health. In that article, medical reformers expressed regrets about the late adoption of Western medicine as compared to Japan. However, they were also abandoning the Japanese Westernization approach of medical modernization—considering it unfit for rural China—in favour of the social medicine approach exemplified by Yugoslavia.

By the 1930s, mission hospitals had also shown much progress. Table 5-4 compares the findings of Balme's report with those of Snell's report. Obviously, mission hospitals were still generally very small and poorly resourced by 1933-1934—hospitals on average had only 4.9 physicians, with 64.7 beds on average. The technological progress of hospitals was also obvious in comparison with 1919, with a much improved situation regarding equipment (e.g. 50% had an X-ray in 1933-1934, as compared to 13% in 1919) (Wong et al., 1936, p. 815).

Table 5-4. Comparison between results of mission hospital surveys 1919-1933/1934

	1919	1933-34
Number of hospitals	189	214
Total value (Mexican silver dollars, \$) *	646,780	43,467,121
Average value per hospital (\$)	36,916	214,465
Total bed numbers	11,900	16,930
Total value of equipment (\$)	850,761	6,582,443
Percentage of hospitals with...		
Pure water supply	8%	42~50%
Good sanitary arrangement	25%	43%
Regular bath of hospitalized patients	<50%	81%
Laundry	57%	80%
Food supply control	50%	94.50%
Standard hospital beds	<45%	80%
Adequate bedding	<63%	90%
Percentage of hospitals without...		
Clothing	58%	27%
Pressure sterilizer	34%	9%

Screening against flies and mosquitoes	37%	9%
Laboratory of any kind	31%	4%
Bacteriological incubator	82%	44%
X-ray plant	87%	50%
Nurse	34%	5%
Night nursing	62%	<16%

Source: adapted from (Wong et al., 1936, p. 815)

Note: * At that time, the Mexican silver dollar was a popular currency used in China.

5.6.2 Splitting into two tracks (1935-1949)

The momentum of state medicine continued to grow after the 1934 proposal for state medicine, however the implementation of the plan faced several challenges. This sub-section covers the post-juncture development of the state medicine programme, the challenges the programme faced, and its adaptation and results. As the sub-section argues, development of primary care level eventually emphasised epidemic control rather than general medical development, as epidemic control was a more pressing issue. The evolution of primary care formed a stark contrast with the increasingly secured institutionalization of hospitals as the place for professional medicine. Towards the end of this sub-section, a brief summary of the communist health programmes is also included.

In 1937, the Chinese Medical Association organized a high-profile conference to discuss state medicine, and publicly endorsed the programme (Wu, 1937). In the issue of the *Chinese Medical Journal* that came out of this conference, Wu traced the history of state medicine to ancient China and argued that state-sponsored health services had existed and endured a long period of time, and that state medicine was a lost native tradition to be revived with new knowledge (Wu, 1937). This appeared to be an attempt to establish state medicine as part of a growing patriotic narrative at that time.

The state medicine programme was continued, despite tight resource constraints and other harsh social conditions (Štampar, 1936). The total government (i.e.

combining central, provincial, municipal and county-level) budget for health was 10.8 million *yuan* which represented only 0.7% of the total government budget, or 0.027 *yuan* per capita (based on *Reports of Minister of Finance and Director-General of National Health Administration for 1936* (Sze, 1943)). In 1940, during the War of Resistance against Japanese Aggression, the Nationalist Party Congress approved the proposal for state medicine, which became the fundamental policy in health governance during the Republic of China (Jin, 1946). At least at one point (1942), the county programme benefited from a county tax (Watt, 2013). The country's top leader, Chairman Jiang (1943) of the National Military Council of the Nationalist Government of the Republic of China, published *China's Destiny* during the Second World War, promising the biggest ever investment in the state medicine programme. After the war ended in 1945, Jin Baoshan, then director of the National Health Administration, called again for universal implementation of state medicine (Jin, 1946). In 1947, implementation of state medicine was written into the Constitution of the Republic of China (National Congress, 1947).

The implementation of the national version of the Ding County health programme was limited and deviated from the original model. The emphasis on “community” development in the Ding County model was fundamentally hindered by unwillingness of the government to bring in social reform. Although advocated as a most important option for social reform with great potential to boost community development by medical reformers (League of Nations Health Organization, 1937), land reform was shunned by the Nationalists (while actively pushed by the Communists). Rural reconstruction was also restricted by the Nationalists and stopped by the invasion of the Japanese in 1937 (Hayford, 1990). Without the strong community that underpinned the Ding County model, recruitment and supervision of village health workers in the implementation of state medicine at village level encountered severe difficulty. Implementers of the state medicine programme resorted to discussion about replacing

village health workers (Lei 2014: Chapter 10).

The “system” plank of the Ding County model was held back by the lack of work force, particularly at the middle level, which seemed to reflect the difficulties in changing professional institutions. First, there was a lack of accumulated quantity of Western medicine practitioners. In 1935, China had 5000 physicians, as many as there were in Yugoslavia, which had only 14.5 million people in contrast to China’s 400 million. In addition, eighty percent of Yugoslavian doctors lived in rural areas (Watt, 2013, p. 62). Second, the Chinese state lacked control over private practitioners, unlike in Yugoslavia, where most doctors were state-employed. According to the estimation of the *Report of Director-General of National Health Administration for 1940*, more than half of the 12,000⁹ Western style doctors in China were in private practice. 75% of private practitioners were concentrated in the main ports of the six coastal provinces, which created “extraterritorial difficulties of control” (Sze, 1943, pp. 18-19).

Third, the reform of medical education was not coordinated with the state medicine programme. Bullock (1980a: Chapter 4) described the resistance at PUMC to call for revision of its elitist curriculum and educational arrangements. There were also various opinions on the criteria for training for the village health stations (Ch'en, 1937b). Not only was the number of students trained small, the training was also not suitable for the pressing health needs of the Chinese population. As Chen pointed out, critical skills like birth control, diagnosis of prevalent diseases (e.g. trachoma), vaccination (e.g. for small pox) were not mastered, nor was the admission and education geared towards a rural career (Chen, 1935). Establishment of medical education in the Central Field Health Station, a national level demonstration project and training organisation for of health delivery, also focused primarily on preventive

⁹ There was a nation-wide reform of medical education in the 1930s, which probably explained the increased number of staff.

services and epidemic control (Dyer, 1936). Only during the war was the idea of using medical schools to train state medicine officers realized. In 1941, the Nationalist Government put all medical graduates under its command (Gao, 2012).

Finally and perhaps most importantly, the champions of state medicine were generally reluctant to accept Chinese medicine practitioners. As the Western medicine practitioners attempted to recruit the state to ban Chinese medicine (other than the *materia medica*), the reaction of Chinese medicine practitioners led to increasing struggle between the two camps (Lei, 2014: Chapter 10). The scepticism about Chinese medicine practitioners was expressed by Oldt (1937). The system in Ding county developed by Chen was in a way aimed at dislodging Chinese medicine with a low-cost version of scientific medicine combining public health and medical care (Chen, 1935).

Due to these challenges, the county health systems and hospitals split into two tracks of development. The county health systems benefited from the popularization of a nation-wide reform of government structure, which made each county government a highly autonomous local administration. The programme of establishing county health systems continued during wartime. After the systemic Japanese invasion in 1937, Chen left Ding County but continued to organize county health systems in Sichuan (Watt, 2013, p. Chapter 6). The coverage of county systems expanded from a handful of demonstration areas plus Jiangxi to more areas around the country. The number of county health centres grew quickly from 181 in 1937, to 691 in 1940, 978 in 1945, and 1440 in 1947 (Compiling Committee, 1984; Lucas, 1980, p. 156). The county health centres were isolated, however, as there were only 436 sub-district health stations under the county health centres by 1947, and almost all were of poor quality (Compiling Committee, 1984, pp. 9-16; Jin, 1946). The War of Resistance, despite bringing huge destructions to rural health programmes, drove a more even distribution as the capital was moved to inland China from coastal Nanjing (Watt, 2013) (see Table

5-5). Amid the social turmoil (including war and hyperinflation) during the later period of the War of Resistance, the leadership of county health centres deteriorated and administrators frequently resorted to “commercial behaviours” (Watt, 2013). With weak ability of hospitals to provide technical supervision to primary care facilities, there was a tendency for provincial and county hospitals (health centres) to serve only the provincial capitals or county seat (Lim & Chen, 1937; Štampar, 1936). In 1942, the government also shifted its priority from training physicians for county health centres to training village health workers (Watt, 2013). In the end, the county-system became virtually a network of poorly resourced state machinery focusing on prevention of epidemics.

In terms of hospital development, the state medicine movement provided a rationale for an increasing role of the state. According to Sze (1943, p. 16), the 60 government hospitals existing in 1943 had mainly been built after 1927. By 1947, there were 110 provincial hospitals, and 56 municipal hospitals. While the government started building new public hospitals, it engaged (or rather attempted to take over) mission hospitals. Mission hospitals also actively joined the medical education system, “[t]he Hopei Provincial Medical College has had but inadequate hospital facilities up to the present; and has lately entered into an arrangement with the Taylor Hospital of the Presbyterian Mission in Paoting by which groups of surgical students make teaching ward rounds at stated times in that hospital” (Hume, 1935). Mission hospitals were generally collaborative with the state, an encouragement for which was the conversion of Chairman Jiang to Christianity (Chatterton, 2010).

The state also showed its eagerness to control hospitals under the name of state medicine during the War of Resistance. There was a retreat of missionaries after the Great Depression (the number of missionaries in general reduced from about 8,000 in 1920 to 4,250 in 1936, and funding generally declined by a third (Deng, 2001)). With the rise of Chinese nationalism, leadership of missionary hospitals was gradually

transferred to Chinese practitioners of Western medicine (Hume, 1935; Spence, 1969, pp. 161-183). In the face of financial difficulty, Hume also pushed for the admission of private paying patients (Hume, 1935). Such localization processes contributed to the government's case to take over mission hospitals. Speaking as a representative of the government for issues related to medical missionaries, Shi Siming (Szeming Sze) outlined a plan to take over all mission hospitals and make them either provincial hospitals or county health centres. Shi argued that the government was able to enforce the plan, as it controlled the ability to withhold or withdraw registration, and both the sources of financing and the managerial roles of mission hospitals had been largely shifted from foreign missionaries to Chinese (Sze, 1943, pp. 16-19). The cooperation between the mission hospitals and the Nationalists only foregrounded a full state take-over once the Communists assumed power.

Analysis of distribution of health facilities also revealed the institutional division between hospitals and primary care providers. Designed as the units of “a basic medical organization” (Ch'en, 1937b), the county health system, when scaled up during the difficult period plagued by social turmoil, became a track of medical development detached from the track of hospitals. A comparison between hospitals and county health centres in provinces with available data between 1935 and 1942 showed drastically different distribution patterns (see Table 5-5). The province with the highest density of hospitals had 14.6 times the number of hospitals enjoyed by the province with the fewest number of hospitals between 1935 and 1936. More drastic differences are apparent, for example, the ratios of physicians, nurses and beds between the best resourced provinces and the least were 24.7:1, 65.5:1 and 24:1 respectively. The distribution of county health centres appeared to be much more equalized—the number of county health centres in the province with the biggest number was 5.5 times that of the province with the smallest number.

**Table 5-5. Comparison of distribution of hospital resources and county health centres
(unit: per million population)**

	Hospitals (1935-1936)	Hospital physicians (1935-1936)	Hospital nurses (1935-1936)	Hospital beds (1935-1936)	County health centres (1942)
Zhejiang (Chekiang)	3.65	14.40	22.00	176.20	3.05
Fujian (Fukien)	2.40	5.70	26.20	183.50	6.40
Guangdong (Kwangtung)	0.74	4.23	10.61	59.97	2.32
Sichuan (Szechewan)	0.36	1.13	3.81	22.62	1.38
Hunan	0.61	2.11	4.46	30.29	2.68
Jiangxi (Kiangsi)	0.45	1.09	4.49	33.21	5.32
Shaanxi (Shensi)	0.50	0.50	0.40	8.30	5.40
Guangxi (Kwangsi)	0.38	0.92	3.69	22.54	7.62
Yunnan	0.25	0.58	0.75	7.33	6.42
Max/min	14.6	24.7	65.5	24	5.5

Source of data: author's extraction from and calculation based on (Liu, 1936, p. 1577) cited in Wong et al. (1936), and (Sze, 1943, pp. 14-15)

Note: * To allow comparison between hospitals and county health centres, only provinces with more than half of counties under Chinese control in 1942 were included in this table.

To better appreciate the nature of the split, it is helpful to study Chen's adapted proposal in 1937. The original Ding County Model constituted a technical hierarchy of hospitals (for inpatients only), health stations (for outpatient and preventive services), and lay health workers (for minor infections and preventive services). The idea seemed to embody a strong belief in the role of professionals in serving the public and also constituted a potential platform for developing general practitioners, perhaps reflecting Chen's own training as a medical professional (shared by other leading medical reformers). As mentioned above, the implementation of this plan was challenged by the difficulty of mobilizing sufficient human resources. An important adaption was proposed by Chen himself in 1937 (Ch'en, 1937a). He questioned the

necessity of hospitals in the county health system for the next two decades, arguing that effectively functioning hospitals were expensive and beyond the means of counties. Chen also pointed out that the health centres were functioning as sub-district health stations in practice (Ch'en, 1937a). Lim et al. (1937) also proposed a redrawing of boundaries such that hospitals should be providing inpatient care only, while health stations provided both outpatient services (comparable to practitioners' offices) and preventive services and public health administration. The "general practitioners for public health" (Ch'en, 1937a) at health stations therefore constituted a bridge between the professional medical sphere and that of mass public health work. However, Ch'en (1937a) further moved to the logic of focusing on de-professionalized health workers, as medical services were not the aim, but a condition for preventive services. In other words, hospitals were no longer considered the most pressing need in Chen's county health system, and Chen's focus of professional development shifted from hospitals to ambulatory health stations. Eventually, this professional bridge was also broken as mentioned earlier in this section. Thus through multiple steps of de-professionalization, an alternative track of local health services staffed by lay personnel was established and separate from the hospitals above the county level.

This split reflected the lack of socioeconomic development, the reluctance of professional medical education to change course, and the challenges caused by coordination difficulties in bringing about institutional change. In reality, the importance of epidemic control and the limitations of development in local society replaced the idealism for comprehensive social medicine. The profound institutional division between hospitals above the county level and county health systems meant that primary care providers (in the form of county health systems) were developed as the antithesis of hospitals. Hospitals continued to be based on relatively high standard professional medical education, providing curative care, enjoying subsidy from either external funding agencies or urban private patients, with the sponsorship of higher-

level government. County health systems were based on scant and usually lay health workers, providing mainly preventive services, severely challenged in fund-raising, and constrained by lack of wider rural development. Hence, the hospitals and the county health systems seem to have embarked on two separate tracks of development.

Parallel to the county health systems, a low-cost model of medical and health work was central in the health work of the Communist Party of China (CPC) in response to the harsh conditions for survival. At the eve of the founding of the Nationalist Administration, Nationalists started to purge CPC members, killing more than 90% of them in 1927 (Lieberthal, 2004, p. 35). The CPC survived and managed to establish base areas mainly in Jiangxi (1927-1934), embarked on the Long March (1934-1935), established new base areas in northwest China around Yan'an (1935-1947) before it became the ruling party of China from 1949. Medical cooperatives involving group financing of dispensaries were organized in Yan'an. In 1946, there were 43 medical cooperative societies in the Shaanxi-Gansu-Ningxia border region (Ouyang, 1984). In contrast to the Nationalists, CPC accepted Chinese medicine, Chairman Mao expressing his attitude towards inclusion of Chinese medicine (Mao, 1991 (1944)) as follows: "New medicine is of course better than old medicine, but if practitioners of new medicine don't care about the people's suffering, don't unite with the one thousand odd old medicine practitioners and old-style veterans and help them make progress, then the new medicine practitioners are in fact helping the witches, and are in fact unconcerned about the enormous deaths of people and livestock." The communist low-cost health work also reinforced the split between hospitals and primary care providers under the Nationalists, and paved the way for future development on the split tracks.

The period between 1928 and 1949 therefore witnessed both disruptive changes and continuities. The previous development of hospitals was consolidated under the Nationalists, financially, professionally, and governance-wise. Organisationally,

hospitals also maintained direct access to ambulatory patients. The emerging track of primary care was separated from this professional sphere, which was almost synonymous with the hospital sector.

5.7 Summary

Over the years between 1835 and 1949, a hospital-biased model of health delivery was established in China. Initially established by medical missionaries during the critical juncture between 1835 and 1844, the hospital bias was reinforced by 1928 by both medical missionaries and foreign philanthropists that saw hospitals as a main vehicle to exert influence. When the country was reunited under the Nationalist Administration in 1928, medical reformers took the initiative and introduced a conjuncture that led to a surge of a state medicine programme by 1934. Extending modern health delivery through an increased attention to primary care level was a central feature of the state medicine programme. However, due to the difficulties involved in challenging continuing hospital growth, the coevolution of hospital and primary care fell into a divided two-sector track, with professionalized hospitals and de-professionalized primary care providers, foregrounding development after 1949.

Chapter 6 Path dependence of hospital-centrism (1949-2013)

6.1 Introduction

This chapter presents the historical coevolution of primary care providers and hospitals in China from 1949 to 2013, analysed using the approach set out in Section 5.2. The following section provides an overview of historical context. Then, Sections 6.3 and 6.4 respectively deal with the history from 1949 to 1978 and from 1978 to 2013. The final sections of the chapter discuss the findings and conclude the whole Part II.

6.2 Overall historical context (1949-2013)

This section provides a brief introduction to historical context necessary for understanding the historical coevolution of primary care providers and hospitals in China between 1949 and 2013. The first decades after the founding of the People's Republic of China witnessed massive socialization programmes, and assistance and lessons from the Soviet Union were heavily relied on, particularly after China broke its relationship with America in the Korean War in 1951. Rural collectivization was rapidly achieved in the 1950s, culminating in the communization of rural society in the Great Leap Forward launched in 1958. In 1966, the leadership of the Chinese government split and Chairman Mao launched the decade-long "Great Proletarian Cultural Revolution" (referred to as "the Cultural Revolution" from here on), which emphasized egalitarian ideology and advocated class struggle as the central policy (MacFarquhar & Fairbank, 1987, 1991). After 1978 when the Reform and Opening up was launched, China turned its focus away from the class struggle to economic development. The economy relied increasingly on market mechanisms, particularly after 1992. From 2003 to 2013, the Hu-Wen administration placed a much higher

emphasis on balanced development (so-called harmonious society and scientific development outlook). The years between 1949 and 2013 witnessed the industrialization of the country, with the period after 1978 seeing continuous rapid economic development, urbanization and technological innovation. The country saw rapid increase of life expectancy in the first three decades, with the disease burden shifting from dominance of infectious diseases to non-communicable diseases after 1978 (Yang, Wang, et al., 2013).

After 1949, the overall structure of the state was generally stable. There were two main lines of leadership. China's official government system had five levels: central, provincial, municipal, county and township. At every level of government, there was a parallel leadership by the governing party—the Communist Party of China, which according to Hu (1998, p. 17) was a *de facto* part of government in the Western sense. Departments of a lower level of government reported vertically to the departments/ministries of a higher level of government on the same functional line, and horizontally to their corresponding level of government. This “matrix” generated difficulties in coordinating policies and was managed through informal leadership, small groups and other mechanisms (Lieberthal, 2004, pp. 234-242). In rare situations like the most tumultuous years between 1966 and 1968, Chairman Mao centralized power into a national system of Revolutionary Committees. After Mao passed away in 1976, decisions were made by collective leadership in general.

Vertically, the central-local inter-governmental relationship in post-1949 China can be described as a hybrid structure between a unitary state and federalism, which some have described as “*de facto* federalism”—“a system of multiple centres of power in which the central and local governments have the broad authority to enact policies of their own choice” (Zheng, 2007, pp. 38–39). Local governments were in charge of a significant proportion of health expenditures and administration and therefore were important in health (Gusmano, 2015). Most health facilities were locally managed, so

their local territorial government played a critical role in implementing central reform policies. Horizontally, while health agencies were the main agency in charge of issues related to health and health facilities, other ministries played an instrumental role in setting policies that influenced the health system profoundly (particularly agencies responsible for finance, social security, planning, etc.) (Bloom & Tang, 2004, p. 14).

The state is not only a structure but also a self-interested actor in the political world, and its interrelationship with societal actors determined the influence of these actors. Professionals had little direct influence in the Chinese political system. As the main professional medical association, the Chinese Medical Association, was annexed by the Ministry of Health shortly after 1949 (Chen, 1989). The Association predominantly consisted of highly qualified doctors, and largely became a tool for the “united front”, a mechanism that the leading Communist party used to connect with non-party actors. In the early 1950s, Russian experts played an important advisory role. A household registration (*hukou*) system divided the country into rural residents (overwhelmingly peasants) and urban residents and effectively excluded the rural population from urban social welfare. People were organized under work units, residence committees or rural communes between 1949 and 1978. This urban-rural divide as will be seen in this chapter had substantial implications for the co-evolution of primary and hospital care sectors. After 1978, market-oriented reform introduced a greater degree of freedom in economic decisions. A national system of labour unions functioned as an instrument for the Communist party to engage urban workers (Unger & Chan, 1995), while there was no similar organisation for rural peasants.

6.3 Development of the two-sector system of service delivery (1949-1978)

This section presents the period between 1949 and 1978. State actors and parastatal local community actors (communes in particular) emerged as the country established a dualist urban-rural system based on a planned economy. As the two

sectors of primary care providers and hospitals evolved in this changing social, political and economic landscape, the question shifted from how the professionals recruited the state to promote health system development to how the state took change of the health system and rearranged policies related to the coevolution of the two sectors.

The years between 1949 and 1953 witnessed urban-based industrialisation reinforcing the previous concentration of modern medicine in urban areas, which consolidated the interest coalition around urban hospitals. From 1953 to 1965, efforts to change hospital dominance seem to encountered enormous challenges. Between 1965 and 1968, major policy changes took place to shift the health development priority from urban areas to rural areas, particularly by moving resources to rural areas, as well as setting up systematic health care programmes there. In the decade that followed (1968-1978), these policies were implemented and adapted, eventually evolving into a two-sector system covering almost everybody in the country.

6.3.1 Launching exclusively urban welfare (1949-1953)

The years between 1949 and 1953, the nascent socialist state introduced a second critical juncture in the direction of hospital centrism. After 1949, public hospitals previously under the Nationalist Administration were immediately taken over by the new Communist regime. The missionaries withdrew particularly after the start of the Korean War in 1950—as mentioned previously, the state medicine programme had paved the road for an eventual full nationalization of hospitals, and the formation of an urban-centred system of hospital dominance. For example, the provincial government of Jiangxi took over the Methodist Nanchang Hospital and seven other mission hospitals in 1951 and another one in 1953 (Jiangxi Provincial Health Gazette Compiling Committee, 1997, p. 221). The party also shifted its focus from rural areas to urban areas (Mao, 1949), and relied on the Soviet Union as a guide for economic

and social development. Following the Soviet practice, the Labour Insurance Scheme (LIS, covering formal employees of factories, mines, railways, etc.) and the Government Employees Insurance Scheme (GIS, covering employees of various levels of government and public institutions, as well as students) were established in 1951 and 1952 respectively (Government Administration Council of the Central People's Government, 1951, 1952a, 1952b).

The reform of the medical education system in 1950 to produce a large number of medical school graduates (Ministry of Health, 1951b) also reinforced the technological dominance of hospitals and the hospital-based professionalization of doctors. The government abolished the previous Western and Japanese systems, in which students underwent general medical training without specialization before graduation. Instead, it adopted the Russian medical education system with early specialization and shortened years of training (Qian, 1988). For instance, the training duration for internal medicine, surgery, paediatrics, obstetrics and gynaecology was reduced from six to five years, which also included one year of internship. Through this reform, students increased their training in their specialties, while “complementary disciplines” were less stressed (Zhang, 1953). The result was rapid expansion of numbers of doctors trained for the specialized needs of hospitals, and the abolition of general medical education. In the year 1951, the number of enrolled students at medical schools (excluding the military medical education system) was 19,770, larger than the accumulated number of medical graduates over the 69 years before 1949 (People's Daily, 1951). In other words, formal medical education became reformed so that it mainly produced graduates suitable for the demand of hospitals. Accompanying this development was the elimination of private practices in urban areas which mainly were generalist in nature (Ministry of Health, 1989).

In the rural areas, the government stressed the prioritization of preventive medicine organized with local resources. Officially, the government prioritized the

restoration and construction of county health centres and *qu* health stations (Xinhua Agency, 1951). The Ministry repeatedly stressed that the priority of county and sub-county health centres/stations was not to provide curative care but to function as a fulcrum of all aspects of health work and to organize a local health workforce to provide preventive-first health work (i.e. providing guidance and training in epidemic prevention and maternal and child health care besides clinical work) (Ministry of Health, 1950, 1951a). This focus was enduring as health campaigns were more difficult in rural areas due to the problems of illiteracy, superstition, poor transportation system, etc (Zhang, 1950). The focus on prevention was also reflected in Zhang Zikuan's report on some rural health experiments organized by the government in the early 1950s (Zhang, 1952, 1955). However, the government refrained from nationalizing practitioners, allowing the continuation of private practice and the formation of group practices (essentially fee-for-service). This contrasted with the pre-1949 state medicine programme which saw private practice as incompatible with public health services. Health committees and associations of health professionals were organized below the level of the county to carry out anti-epidemic and maternal and child health work under the supervision of local government. Chinese medicine practitioners were the dominant curative providers in rural areas.

The resulting apparently urban-biased welfare system was a major turn away from both the previous rural-centric state medicine model, and the CPC's pre-1949 reliance on rural revolution and cooperative medicine approach. It was justified in official propaganda as a short-term phenomenon:

“immediate implementation of publicly funded medical and preventive care system for the entire population is impossible. [We] can only start from the partial, and gradually expand (the coverage) incrementally from now on” (Zhang, 1953).

6.3.2 Self-reinforcing development of hospital dominance (1953-1965)

The initial urban bias quickly saw a post-juncture process of financial self-

reinforcement. While both hospitals and primary care facilities were able to generate extra-budgetary revenue from patient charges, hospitals enjoyed the demand-side subsidy from insurance payment as well as the greater affordability of urbanites—both strong and stable. The bias was reinforced in urban areas due to the prioritization for hospitalizations under GIS and LIS. According to the implementation document for the GIS, “30% [of the fund] is for outpatient medical and pharmaceutical devices and materials, and physical examination; 70% is for inpatient medical and pharmaceutical devices and materials” (Government Administration Council of the Central People's Government, 1952b). Such a reimbursement system favouring hospitalization contributed to the concentration of resources in hospitals. From 1954 to 1957, the bed occupancy rate of hospitals nationwide increased from 69.4% to 80.3% (Ministry of Health, 1989, p. 91).

During the post-Great-Leap-Forward retrenchment period, government provided fully budgeted wages to hospital staff in 1960, and also started to request hospitals to lower charges for hospitalizations more than for outpatient services, in official terms “less reduction for minor illnesses, more reduction for major illnesses” (Ministry of Health & Ministry of Finance, 1960). Hospitals consumed 70% of total health expenditures in the 1960s (He, 2001). Pharmaceutical and materials sales mark-ups (i.e. “bought at wholesale prices from distributors and sold at retail prices to patients”) became an important way to finance hospital services, replacing previous commitments to fund the gap between expenditures and revenues of hospitals (Ministry of Health et al., 1960).

The financial investment skewed towards hospital development resulted in institutional resilience related to powerful societal interests. The most obvious illustration was the government's largely failed attempt to rein in deficits of medical spending. In the 1950s, spending on GIS and LIS surpassed the growth rate of the overall economy (Liu, 2009). In Shandong, affiliated institutions of the provincial

government suffered from an increasing deficit in the GIS: 18% in 1954, 33% in 1955 and 42% in 1956. Similar or worse deficits existed widely in other provinces, such as Zhejiang, Shanxi, Hunan, Hubei and Jiangxi (Wu, 2014). *Jiankangbao* (or *Health News*, a main mouthpiece of the Ministry of Health) in 1957 blamed the urbanites for allegedly excessive use of medical resources: “[y]oung cadres, students and workers... each year on average visit a clinic twelve or thirteen times” (Hillier & Jewell, 1983, p. 81). The CPC’s principal mouthpiece--*Renmin ribao* (People’s Daily), as well as a key journal of the Ministry of Finance, also pointed towards poor fund management, use of GIS and LIS benefits by people not covered (relatives of beneficiaries), incentive of medical professionals to generate income from patient utilisation, and the “egalitarian attitudes” of some beneficiaries (thinking that everybody, sick or not, is entitled to the same amount of expenditures) (Cao, 1955; Huang, 1957; Liu, 1957; Yin, 1955).

In response to the overwhelming expenditures on urban health services, the government took four major actions: lowering the price of pharmaceuticals, setting up regionalized referral networks, establishing sanatoriums, and expanding hospital capacity, however, these efforts were not able to contain growing demand, and ended up tying hospitals even closer to urban users. The government ordered a decrease of pharmaceutical prices in 1957, feeding further increase in utilisation (Liu, 2009). Lessons from the Soviet model of “sectional medical service” (a system of regionalized referral) were introduced first in large cities and then rolled out nationwide. Reports about earlier pilots in Beijing and Shanghai showed that regional planning and referral systems were being organized and with some evidence of effectiveness. For instance, between 1955 and 1956, large hospitals in Beijing saw their average daily outpatient visits per physician reduced from 12.7 to 11.8, while

urban union clinics¹⁰ saw their daily outpatient visits increased from 16.1 to 27.6 (Beijing Municipal Health Bureau, 1955; Wang, 1957). The Ministry of Health (1957b) issued a policy in 1957 to establish a regional health delivery system with a technical hierarchy and referral mechanisms between urban health care facilities. Changes were made through the following measures: 1) moving and establishing new facilities; 2) establishing a technical and professional leadership relationship between health facilities at different levels; 3) taking advantage of the contract between health facilities and LIS plans to require gatekeeping; 4) establishing health care facilities specifically for industry workers; and 5) getting health care closer to the people and integrating preventive and curative care (Wang, 1957).

The effects of the efforts to rein in the surging utilisation seemed limited. The MOH admitted that there were difficulties in shifting care: “the specialisation and quality of grassroots health facilities are highly uneven at present”, and a significant number of urban residents (who did not enjoy formal employment status) were paying out-of-pocket and therefore were hard to influence through reforming insurance payment policies (Ministry of Health, 1957a). The MOH also increased the supply by enlarging hospital-based outpatient departments and introducing simplified beds and family based beds to expand hospital service capacity (Ministry of Health, 1962). This further expanded the hospitals’ role in more general cases that could have been the basis for developing general practice.

A large number of sanatoria, essentially rehabilitation beds, were also established (from 3,900 to 68,860 between 1949 and 1957) so that patients who needed rehabilitation could be transferred away from hospitals to ease the pressure on hospitals (Ministry of Health, 1989, p. 25). Also encouraged by the Soviet experience,

¹⁰ Union clinics were group-owned clinics set up in the early 1950s mainly consisting of previously private practitioners.

independent industrial health delivery systems flourished during this period. Between 1949 and 1957, the number of health staff in industrial and other sectors rose from 25,880 (6.5% of the total number of health staff) to 142,972 (15.1% of the total number of health staff), or about 40% of the number of staff working in the government health sector (353,538) in 1957 (Ministry of Health, 1989). However, the industrial health facilities served workers based in factories, and therefore were beyond the realm of the Ministry of Health. The demand on urban facilities made any significant shift of resources to rural areas in the future difficult, and gradually the intended nation-wide referral system was turned into a fragmented system of health service delivery and risk pooling. The limited achievement in developing urban primary care seemed overwhelmed by the urban demand on hospital services. As a result, the hospital sector became embedded in the urban socio-economy.

The rural health system developed further on a diverged path from that of urban areas. In terms of ownership and responsibility for funding, the MOH constantly stressed the distinction between vertical professional leadership and horizontal local responsibility, throughout its changing policy language regarding rural primary care along with rural reforms (see Table 6-1). The lack of a country-wide development of state medical welfare following the coverage for urban populations, that some official propaganda promised in the early 1950s, was legitimized, as the lack of publicly owned facilities at sub-county levels was said to be compatible with the socialist system until 1965. Take for example the policy in 1957, the agricultural cooperativization movement was in full swing and leading some areas to form medical cooperatives. The MOH encouraged the development of self-reliant health stations based on agricultural cooperatives, but made it clear that while the county health agencies were to assume professional leadership, commune party secretaries were to be responsible administratively and financially for commune health centres. The state was not supposed to take over “united clinics” that were essentially joint private

practices and which charged patient fees for services under government regulation (in the price schedule and location of practices, and participation in public health services), stating the legitimacy of patient charges and private practitioners and discouraging “averagism” (i.e. equal pay for all staff members). It also discouraged annexing of so-called half-peasants-half-doctors within united clinics (Ministry of Health, 1957a).

After 1958, the formation of rural communes (a collective economic and political organization corresponding to townships) provided the economic and organizational basis for commune-funded medical care. After the Qishan County party secretary reported their collective rural health network as the first national model of commune health organization in the People’s Daily (Yi, 1958), the MOH organized a national field conference in 1959 to promote the experience (Ministry of Health, 1960). Structurally, commune health centres were formed as mini-hospitals and replaced private practice. The commune party secretary (or a commissioner of the commune party committee) usually assumed the directorship of commune health centres. Funds for commune health centres were collected through a collective welfare fund (Renmin Ribao, 1960). As the commune health policies embedded rural medical welfare within agricultural production, the naturally unstable revenue due to uncertain yields was passed onto rural primary care financing.

Table 6-1. Policies on ownership of primary care facilities

Year	Policy
1957	“union clinics emerged from the people... are health welfare institutions with socialist nature, the state should not take them over” (Ministry of Health, 1957a)--implying that private or collective provision of care based on user charges was not only legitimate but also good for the rural population
1959	“for the medical expenditures of the people... it is best to mainly rely on individual payment, with appropriate subsidy from the state and communes” (Ministry of Health, 1959)
1960	“collective health and medical schemes are considered preferable” (Ministry of Health, 1960)
1962	“the main form of rural grassroots health organizations should be doctor-owned group practices for a very long period of time”(Ministry of Health, 1962)

1965	“doctors’ group practices are the most numerous and the most problematic, and should gradually move towards commune/brigade ownership” (Ministry of Health, 1965)
1974	“the policy of ‘commune sponsorship with public aid’ applies for collectively-sponsored commune health centres” (Ministry of Health, 1974)

The development of medical welfare was likely to have been affected by the continued focus on epidemic prevention and maternal and child health. In 1956, county health centres went through a separation of their public health functions (as county epidemic control stations) and medical care functions (as county hospitals) (Ministry of Health, 1956a). The constant effort to justify fee-charging private practice on the part of the MOH suggests that commune medical welfare was also driven by the commune leadership. Indeed, medical entitlement was among a pack of tangible benefits provided by the communes, in a way functioning as a reward/compensation for giving up private ownership (Huang, 2000, p. 64). The Great Leap Forward seemed to have unintended effects on medical welfare as well, as increasing rural workloads during the Great Leap Forward and a post-Leap famine contributed to infectious and nutrition related diseases in rural areas and in a way compromised the success of health campaigns (State Council, 1963).

In terms of human resources in rural areas, a multi-tier strategy was adopted to train high, middle and primary grade practitioners as mentioned in Chapter 4. At the county level, health centres/hospitals were mainly staffed with midlevel graduates. These graduates received three-year training in Western medicine (as compared to generally five years for fully-trained doctors, and weeks or months of training for primary level health personnel) and became assistant doctors. The admission required elementary education only (compared to senior middle school education for university entrance and basic literacy for primary level health personnel). At what gradually became the township level, half a million Chinese medicine practitioners were registered, and a large proportion of them were organized as united clinics (Ministry

of Health, 1989). Three-year Western medicine assistant doctors were added to the Chinese medicine practitioners from the Great Leap Forward. For example, between 1958 and 1961, 4800 young graduates (clinical medicine graduates mainly trained in three year programmes) from medical schools were reported to be assigned posts in sub-county health facilities (Xinhua Agency, 1961).

Despite the government's rhetorical emphasis on unification of Chinese medicine and Western medicine, the training model for Western medicine appeared more productive as the number of Western medicine practitioners gradually overtook that of Chinese medicine practitioners. The institutionalization of village level health workers relied on training local peasants as half-peasant-half-doctors. Before the Great Leap Forward, union clinics with local Chinese medicine doctors trained local peasants for basic health work, and training was mainly done through apprenticeship following the tradition of Chinese medicine. The productivity of training was boosted after the construction of commune health clinics, which became the arena for training village staff. Support in training also came from urban hospitals. There formed a very long chain of knowledge transfer all the way from urban hospitals to county hospitals, from county hospitals to commune health centres, and from commune health centres to village health workers (Fang, 2012; Xu, 1960, p. 55). The apprenticeship model and the long training chain, however, limited the speed of developing quality and quantity of human resources. As the county health centre/hospital doctors were mainly three-year medical graduates until 1960, there was also limited capacity for technology transfer (Ministry of Health, 1989). The development of rural human resources was not only lagging behind hospitals, but was also set on a completely different trajectory.

6.3.3 Turning towards rural primary care (1965-1968)

The convergence of major political events between 1965 and 1968 formed a conjuncture to strengthen primary care. In 1965, the "26th June Directive" prompted

the Ministry of Health to shift the focus to rural areas. Starting from 1965, urban hospital-based doctors were organized on a large scale as mobile medical teams to rural areas, with the crucial task of training village-based half-peasants-half-doctors, later known as the barefoot doctors (Commentator, 1965). This marked a drastic shortening of the knowledge transfer pipeline, with classroom-based education using a simplified textbook and clinical shadowing replacing the traditional apprenticeship, and greatly scaled up medical training at the village level (particularly regarding Western medicine) (Editorial Office of the Red Flag Magazine, 1968; Pinggu Health Bureau, 1986a).

Change was halted during the start of the Cultural Revolution, when the country went through turmoil in the state bureaucratic machine, known as “bombarding the headquarters”, introduced by Chairman Mao and his radical followers. The chaos eventually led to a centralized political structure and policy making at the very top) (Lampton, 1977: Chapters 8 and 9), essentially inhibiting the potential “veto points” (urban officials, Ministry of Health and Ministry of Education). This allowed not only a major reform of medical education to train barefoot doctors but also popularization of the Cooperative Medical Scheme (CMS). In 1968, the two main articles respectively promoted CMS, and barefoot doctors and reform of medical education marked the peak of policy change (Editorial Office of the Red Flag Magazine, 1968; (Revolutionary Committee of Leyuan Commune, 1968).

The policy change was not just about substance but also about discourse, representing the ideology of the Cultural Revolution. An illustration of such change was that according to an article promoting CMS in *Renmin ribao* in 1968, endorsed by Chairman Mao, “poor and lower-middle peasants should hold the power of rural medicine” (Revolutionary Committee of Leyuan Commune, 1968), essentially foregrounding not only the reinforcement of a community-based low-cost model of health delivery but also ruling out a state bureaucratic stake in rural health

development. The policy change was both partial and problematic as the quality of policy making seemed affected by the lack of representation of multiple interests and access to information.

6.3.4 Resistance in implementation and development in rhetoric (1968-1978)

Between 1968 and 1978, the most impressive post-juncture changes were perhaps the nation-wide coverage of CMS and barefoot doctors. This was both a break away from previous strategies for rural health development and a continuation of a two-sector approach where hospitals and primary care providers were generally embedded in the urban-rural divide. Also critical for further development was how policies related to the two sectors of primary care providers and hospitals evolved, particularly how solidly the rural primary care providers were institutionalized.

The early formation of an urban-based hospital-centric model of health service delivery meant constraints for policy development when rural health was prioritized in policy. Two sources from the Beijing Health Bureau illustrated the resistance that the policy of sending urban doctors to rural areas was faced with challenges even during the Cultural Revolution. In the two years of 1969 and 1970, 3619 medical staff (30% of total medical staff) were sent to settle down in rural or remote areas through becoming state employees (Beijing Municipal Health Bureau, 1973b). However, the “gap in urban hospital staff increased year-on-year, overload and chaos became a serious phenomenon, affecting the care-seeking of mass people of workers and peasants.” As a response, the document asked for delay in sending down hospital staff until “urban medical staff are appropriately replenished to allow for continuous sending-down of a proportion of medical staff”, suggesting an urban priority over rural areas. The document turned towards the need to address the weakness of rural primary care facilities by increasing training (Beijing Municipal Health Bureau, 1973a). National numbers of hospital doctors and staff also suggested that the ten years

between 1965 and 1975 witnessed generally continuous development of hospitals at the county level or above. The proportion of doctors in urban areas and urban hospitals was constantly higher than in rural areas, and was growing between 1970 and 1976 (Ministry of Health, 1989, pp. 60-70).

In a dynamic process influenced by resistance to change in urban areas, the rural health system (primary care providers in particular) were further developed based on a shift in rhetoric regarding rural self-sufficiency. The rhetorical reconstruction of what “shifting focus to rural areas” meant is reflected in three paintings/posters of rural health workers published between 1962 and 1974, exhibited in Figure 6-1. In the first painting (top left) just before rural health adjustment took place, the rural health worker presented an image of a lonely mobile health worker carefully holding Western medicine (the white medicine box seemed an alien object as the colour would be avoided in Chinese traditional culture as it was linked with funerals) with both her hands. The painting implied that the donkey-riding health worker with a serious look on her face was a deliverer of modern medicine from county or sub-county health facilities to the distant and poor rural areas, isolated by the mountains and enduring hardship.

In the second painting (top right), the smiling health worker walks in the blooming rice field, carrying a medicine box in one hand and a hoe in the other. It highlights the identity of a half-peasant-half-doctor, and suggests that health care was now integrated with and contributing to the front line of agricultural production—the basis of industrialization represented by the high voltage transmission tower in the distant background. The flowers in the painting changed from those of wild grasses to those of a full-blooming peach tree, suggesting a promising spring and pending fruitfulness of the new model of village-based health services.

In the last picture, several barefoot doctors are studying Chinese herbal medicine,

as these medicinal flowers and grasses replaced the centrality of the medicine box. The barefoot doctor programme was established based on the local village community. Professionally, barefoot doctors became the staff of cooperative medicine stations and earned work points as a member of the collectivized village (production brigade). The male peasants smiling along with the barefoot doctor seem to personify three points related to self-sufficiency: first, the shift towards the “treasures” of medical knowledge kept in the poor and lower middle-income peasants; second, the discovery of solutions to the sustainable development of CMS; and third, the firm grasp of the materials for, and the future and the power of, medicine in the hands of the peasants on the fields enclosed by the mountain ridges. The rhetorical reconstruction should be seen as both an integral part of providing an ideological basis for barefoot doctors and CMS, and a propagandist ideological packaging that was in a way blinded to the underlying predicament.

Difficulties of rural primary care were obvious when sources in the local archives were examined. In terms of financing, the scarcity of economic resources in rural areas led to a low degree of accumulation of resources in the primary care sector which was exposed immediately afterwards. The cooperative medical schemes faced serious challenges in achieving self-sufficiency. As sources from the Pinggu County Archive show, in 1972, 51.03% of cooperative medical stations were in deficit. The situation continued and by 1978 still more than a third of cooperative medical stations were in deficit (Pinggu Health Bureau, 1973, 1979). Organizers of CMS emphasized the function of preventive care due to the need to use funds efficiently (Revolutionary Committee of Leyuan Commune, 1968). Contributing to the financial instability seemed to be the fact that the popularization of barefoot doctors and CMS contributed to a rapid popularization of Western medicine.

The lack of sustainability prompted a political campaign to reconstruct the value of local resources--Chinese medicine (also reflected in the last painting in Figure 6-1).

From 1972, with financial constraints and the scarcity of drugs, growing herbal medicine for self-sufficiency, along with training barefoot doctors in acupuncture, was encouraged. For instance, 67 native pharmacies were built in Pinggu County (Pinggu Health Bureau, 1986b, p. 39), and patients were encouraged to use native herbal medicines. Speech scripts of a local conference on the cooperative medical scheme during the Cultural Revolution in the archive of Pinggu suggested on paper that township health centres were complying with the official line described in the poster (Party Branch of Heidouyu Brigade at Huangsongyu Commune, 1977). As the archive was generated by the county (now district) health bureau, it reflected the documented reality that policies were adhered to on paper. However, the oral history conducted by Fang (2012) in a county in Zhejiang suggested that patients did not trust herbal medicines. Fang also reported problems with growing, processing and producing Chinese medicines, as well as patients' growing preference for Western medicines which offered usually a faster fix.

Related to weak financial institutions, rural primary care was established on weak governance institutions. CMS and primary care services were mainly managed at commune (township) and brigade (village) levels, which meant disconnect with the Ministry of Health and medical elites. Professionally, there was a lack of career prospects in medicine for the rural health workforce. Statistics of barefoot doctors in Pinggu County between 1965 and 1975 showed that out of 1652 barefoot doctors, 1182 remained as barefoot doctors by 1975. Only 353 were able to leave the rural areas through the four key approaches of upward social mobility: 71 joined the army, 120 were recruited by factories, 132 were admitted by colleges or schools (generally unrelated to medicine), and 30 became party or government cadres (Pinggu Health Bureau, 1975). In other words, a career as a barefoot doctor meant a dead end for someone aspiring to be a fully qualified medical professional.

The post-juncture development was therefore limited by the rural challenge

associated with the weak institutional underpinnings for primary care and the urban resistance associated with the strong institutional underpinnings for hospital care. What was formed was not a harmonized structure of a system, but a system in which primary care development became an anti-thesis to hospital development. The institutional bias towards hospitals was legitimized by the nature of the propagandist rhetoric on rural health care, and eventually reinforced the divided track of urban hospital care versus rural primary care. By 1978, China had established nation-wide low-cost rural primary care based on barefoot doctors, CMS and three-tier delivery networks, while leaving weak institutional underpinnings for resource accumulation for primary care in rural areas as compared with hospitals in urban areas.



Figure 6-1. Three paintings featuring rural health workers in the 1960s and the 1970s
 Note: Top left: a painting with the title of “Rural Doctors”, published in the People’s Daily in 1962; top right: a poster with the title of “Health worker at the head of the field” published during the rural restoration and construction period (1961-1965); bottom: a poster with title of “Barefoot doctors all over the mountain village, cooperative medicine with new prospects”.
 Sources: (Li & Jin, 1964; Shanghai renmin chubanshe, 1974; Zhang, 1962)

6.4 Adaption and path dependence during reforms (1978-2013)

The period between 1978 and 2013 saw continuous and rapid economic growth,

as well as rapid social transformation that contributed to increasing demand for health care, including the increasing prevalence of non-communicable diseases and population ageing. Critical questions for this period, therefore, include how policies about hospitals and primary care providers were made (or not made) and implemented, how such policies (or the lack of them) channelled increased resources to the two sectors of hospitals and primary care providers to address the increasing demand for health care. This section argues that how the government responded to some short-term challenges immediately after 1978 was of critical importance to the coevolution of hospitals and primary care providers up to 2013.

The years between 1978 and 1985 saw difficulties for financing hospitals and primary care facilities, as well as the response or lack of response of the state to the needs of each sector. From 1985 to 2002, a short-term health financing response to state financial difficulties turned into a lasting divergence in institutions related to the two sectors. Between 2002 and 2009, there was a conjuncture, where the political agenda turned towards constructing universal health coverage, with major health reform plans formulated. The section finally turns to the period after 2009, to investigate the implementation and adaptation of the health reform plans and their consequences for the coevolution of hospitals and primary care providers.

6.4.1 Changes favouring hospitals and technology (1978-1985)

The years between 1978 and 1985 formed a third critical juncture in relation to the coevolution of hospitals and primary care providers. After 1978, the MOH actively restored the technical capacity of hospitals, which were believed to have been harmed during the Cultural Revolution (Bloom et al., 2004, p. 14). The MOH also tried to build hospitals into regional technical centres. In urban areas, a number of advanced medical centres were reinforced as regional technical centres (Henderson, 1989), benefiting from their sponsorship by relatively higher levels of governments. In rural areas, the

MOH launched a programme to strengthen health services in one third of counties. Although the stated goal was to strengthen county-wide health services, the document discussed mainly strengthening county-level facilities, and was followed by an operational policy on strengthening technological development in county hospitals, with no operational policy on primary care facilities (Ministry of Health, 1980a, 1980b). While the initial government support soon faced fiscal challenge (State Council, 1985), the drive to advance medical technology was evident. Technological development of hospitals also benefited from the growing number of medical graduates after the restoration of higher medical schools in 1977 (Anonymous, 1977). The labour insurance scheme for retirees, a scheme which had disintegrated during the Cultural Revolution, was restored (State Council, 1978a, 1978b). These efforts to restore hospital-related institutions immediately after the Cultural Revolution suggested the resilience of the such institutions.

Under the pressure of fiscal difficulties and inflation, the MOH and the Ministry of Finance started to push health facilities step-by-step towards seeking extra-budgetary revenue and increasing efficiency. As early as 1979, Minister Qian started advocating “utilizing economic measures in managing health services” (Ministry of Health, Ministry of Finance, & State General Administration of Labour, 1979), giving public hospitals the right to redistribute residual income from services, though salary was still paid through public finance. One of the MOH’s top priorities was to address the perceived lack of access to hospital admission and the deficits of hospitals (State Council, 1981a, 1985), which were explained as the result of previous waves of price reduction (including the 1960 MOH policy mentioned above), lack of public financial compensation, and rising costs of materials. Accepting that the government was not to pay for the deficits through direct financial assistance, price adjustment became one of the key solutions. In 1981, the State Council approved a double standard fee schedule, allowing hospitals to charge patients with GIS and LIS at cost while keeping the prices

for other urbanites and the rural population unchanged (State Council, 1981a). In 1985, the government again avoided any general adjustment of prices in health care, while allowing price increases in new equipment, new medical procedures, and newly built, renovated and expanded facilities, and it reemphasized the need for separate fee schedules for insured and uninsured patients (Ministry of Health, 1985b; State Council, 1985). The State Council also implemented a management responsibility system for hospital economic operation (State Council, 1985). In effect, hospitals were supposed to cross-subsidize among their own services; non-basic services started to be provided with a profit margin, while basic services were provided below cost (Liu, Liu, & Chen, 2000). Hospitals, like other health facilities, also enjoyed the profit from drug price mark-up. In short, the post-1978 reform addressed the short-term financial needs of hospitals and service demand for hospital services by reinforcing previous policies that enhanced their ability to generate extra-budget revenue and creating high-power incentives to increase service turnover.

In rural areas, the primary care sector faced difficulties. The relatively lower bureaucratic stake in the rural delivery system and the collapse of the Cooperative Medical Scheme (Duckett, 2012) contributed to the difficulties of rural primary care. In a response to the collapsing Cooperative Medical Scheme, the government issued a policy on subsidy for barefoot doctors in 1981, at the same time allowing barefoot doctors who proved qualified to the level of specialized medical higher middle schools through exams to be certified as village doctors (State Council, 1981b). In 1985, when the Minister of Health Chen Minzhang announced the abolition of the title of barefoot doctors (Ministry of Health, 1988), the 1.2 million barefoot doctors were disaggregated as 643,022 village doctors and 650,072 health personnel (i.e. people working at village clinics without the certificate of a village doctor) (Ministry of Health, 1989, p. 79). The MOH also allowed “multiple forms” of practice, virtually endorsing the privatization of village clinics though still maintaining formal

government regulation of village doctors (State Council, 1985).

For township health centres, there was an immediate human capital crisis due to an exodus of the rural health workforce in the early 1980s, exposing the mobility of human resources at the township level. In Pinggu, Beijing, 166 technical “backbones” (most of whom had been “sent down” from urban hospitals) left the county, almost eradicating its technical capacity (Pinggu Health Bureau, 1986b, p. 20). Similarly, in a township health centre in Heishan County, Liaoning, among the 11 graduates from specialized secondary school assigned there since 1962, only the director of the township health centre was still working there in 1982, while 13 of its 19 staff members were temporary employees (Liu, Xu, & Wang, 1996). Some of the local health administrators said the policies were “eradicating” rural health professionals (Liu et al., 1996). Zhang (1982) summarized the main reasons for the exodus as concerns for children’s educational and employment prospects, relatively low incomes and weak incentives, and lack of prospects for career development. Those who exited tended to be the more qualified professionals.

6.4.2 Evolution of the two-sector structure (1985-2002)

The reform between 1978 and 1985 seemed to have created a self-reinforcing belief in the role of pro-market mechanisms in managing health care. The government further issued autonomization policies, which drove hospitals to depend increasingly on service charges (based on fee-for-services with price distortion that allowed cross-subsidies from expensive technologies priced above cost to physician services priced below costs) rather than on financial subsidies, as demonstrated in a declining share of public finance in total health expenditures (State Council, 1989, 1992, 2000). Revenue-based bonuses were adopted in hospitals to incentivize doctors to generate more revenue for the hospitals. Facilities could retain and redistribute profit without the need to transfer revenue and wait for appropriations. This, along with the subsidy

of medical services from drug mark-ups and technology and profit margins, created perverse incentives for over-prescription (Liu & Mills, 2005). The technology-biased price schedule, drug price mark-up, and increased hospital autonomy in redistributing revenue collectively formed a high-power incentive linking hospital managers and doctors in over-prescribing expensive technological procedures and medicines. They allowed public hospitals to develop with little direct government subsidy (from 12.6% of total hospital revenue in 1990 to 8.7% in 2002, calculated with data from China Health Yearbooks 1991 and 2002). The period also saw a medical equipment race as a result of the revenue that could be gained from high prices for high-technology medicine. While the quota of equipment distribution was one CT scanner for every 0.3 million population according to a national policy, the number of CT scanners in practice was much higher and, for example in three cities (Yanjin, Anshan and Qingzhou), amounted to 3.4 (in 1993), 2.6 (in 1994) and 8 (in 1997) times the national quota (Li, 1997). Pharmaceutical sales accounted for 47% of all revenue of comprehensive hospitals under health departments (Ministry of Health, 2003).

The former barefoot doctors also generally relied on profits from drug sales to earn their living. In 1989, 59.4% of all clinics operated by barefoot doctors were fee-for-service private clinics, including 11.1% as group private practices and 48.3% as solo private practice. The remaining 40.6% were collective clinics, most of which were village clinics. Others went back to farm work (Liu & Cao, 1992). Barefoot doctors fell from the core of a village community-based health system, including being the guardian of the CMS, to the bottom level of auxiliary care extended from and supervised by urban-based formal and privileged health care professionals. Funding became a challenge not just for service, but also for training. The status of barefoot doctors within villages was also eroded as agricultural revenue increased, and some village doctors started selling “illicit drugs” (Huangsongyu Commune Health Centre, 1989).

The poor capacity of township health centres created a vicious cycle of underutilisation and losses. Liu et al. (1996) discussed the poor conditions of township health centres in the 1980s and the 1990s as follows. Some township health centres went bankrupt, some found maintaining buildings and equipment a challenge, and others had to downsize the number of beds, close departments and reduce patient volume, or even rely on barefoot doctors to provide all services. Dangerous working space accounted for 12% of the total built space in township health centres on average, and the percentage could be as high as 31% in poor areas. Many township health centres still relied on the “*lao san yang*” (three old items), i.e. clinical thermometers, stethoscopes and sphygmomanometers, to provide medical diagnosis. Liu concluded that low perception of quality of care affected patient demand, and exacerbated financial losses and brain drain. While hospitals were able to cross-subsidize using the reformed price schedule, the demand for relatively sophisticated technology was low at the level of township health centres. Price increase on technologies which benefited hospitals therefore further contributed to shrinkage of demand for township health centres. In one township health centre, for example, price reform had to be reversed as it had caused drastic decline of patient visits and the consequent loss was greater than 50% of expenditures (Liu et al., 1996). Specialization of township health centres through integration with urban hospitals’ specialist clinical departments did not work out, as economic competition eroded the intended collaborative morale (Liu et al., 1996). 40% of the staff at township health centres did not have professional education in 1983, and by 1988 the percentage was 58% (see Chapter 4). Township health centres generally could not benefit from the new opportunity to develop extra-budgetary revenue. As reflected in Chapter 4, the number of beds stopped growing in township health centres while the number of hospital beds grew steadily between 1978 and 2002.

Weak institutional factors were also reflected in the delay and difficulty in addressing the chronic lack of finance, human resources and capital in primary care

with the collapse of the CMS and the end of state support (assistance to collectives). Minister of Health Cui Yueli, a more pro-market radical reformer according to previous analysis (Huang, 2000), noted in a speech to all directors of provincial health bureaus that 18.2% of township health centres could generate only 30-49% of expenditures from revenues (excluding government grants), while 33.11% could not even earn 30% of their expenditures. While Cui asked for more fiscal contribution from local government, he mentioned in the same speech that reform of their management should remain stable and along the same “autotomizing” principles (Ministry of Health, 1989). In 1990, the Ministry of Health managed to place primary health care on the policy agenda and issued a policy requiring establishment of a network of rural primary care facilities along with cooperative medicine nation-wide by 2000 (Ministry of Health, 1990b). However, the policy was quickly dropped on the ground that it could potentially add to the burden of peasants, which seemed particularly dangerous as tensions between local (i.e. county and township) governments and peasants were already tense due to tax collection and fee charging for various public services, as local government replaced communes (General Office of Central Committee of Communist Party of China & General Office of State Council, 1993). When the Urban Employees Basic Medical Insurance Scheme replaced the Labour Insurance Scheme, the new scheme adopted policies introducing competition among empanelled health facilities, abolishing the previous requirement of referral from facilities at lower levels in order to claim expenditures in hospitals (State Council, 1998). Hence, both frustration of efforts to restore rural health coverage and the development of insurance schemes encouraging patient choice and competition across different levels of care reinforced pre-existing dynamics of hospital bias.

The dynamics of the health system reforms between 1978 and 2002 did manage to ensure the survival of primary care facilities, however, the short-term-oriented reforms (i.e. technology-focused price adjustment, the economic management system,

and the pharmaceuticals and materials sales mark-up policy) collectively biased the health system towards adoption of advanced technology, and discriminated against primary care facilities. Although the post-1978 reform brought new opportunities and challenges to both hospitals and primary care facilities, hospitals benefited relative to primary care facilities. Overall, this period witnessed a continuous downstream evolution of an institutional balance that favoured hospitals, as hospitals tended to adapt better to the changing social economic order.

6.4.3 Turning towards universal health coverage (2002-2009)

The period between 2002 and 2009 marked a U-turn in the government's commitment to health and universal health coverage, and a conjuncture in the coevolution of hospitals and primary care providers. In the World Health Report 2000, China's health system was ranked 188th in terms of fairness of financial contribution among 199 countries and regions (World Health Organization, 2000). Added to the pressure for reform was the outbreak of Severe Acute Respiratory Syndrome (SARS) in 2002, which exposed the weak health care coverage, particularly in rural areas. A policy window was opened with a change of administration, which shifted the development agenda from a single-minded focus on economic development to a balanced approach of socio-economic development. The Chinese government continuously put greater public funding into health care, and actively expanded social insurance schemes. Also, a series of rural health financing experiments provided the empirical evidence needed for designing new insurance schemes (Liu & Rao, 2006). In a rare move, a government research organisation published a report criticizing the previous pro-market health system reform as "unsuccessful" (Ge & Gong, 2007). Following a series of debates and important conferences, the central government eventually produced a milestone policy document on comprehensive health system reform in 2009 (Central Committee of Communist Party of China & State Council, 2009). The remainder of this sub-section deals with the three most relevant aspects of

the reform plans.

In terms of financial coverage, two major schemes that covered population groups left out during the previous reforms were set up and expanded during this period: the New Rural Cooperative Medical Scheme (usually abbreviated as NCMS to distinguish it from the “old” Cooperative Medical Scheme) in 2003 and the Urban Residents Basic Medical Insurance in 2007. The reform brought significant increase in use and coverage of health services at all levels of care (Meng et al., 2012). However, at least in the short term, it seemed that policies were not well coordinated towards improving the balance between hospitals and primary care providers. Health insurance schemes were biased towards reimbursement for hospitalizations, on the grounds that these were thought to be related to illnesses with catastrophic expenditures (Yip & Hsiao, 2009). The new health insurance schemes encouraged competition between providers rather than establishing a referral chain, which also favoured hospitals as they were better-capacitated. Both the emphasis on reimbursement for inpatient services and competition fed growing demand for hospital care. In order to boost enrolment, even the recently established NCMS abolished referral requirements within counties in 2007, so patients could access any facilities within their counties (Ministry of Health, Ministry of Finance, & State Administration of Traditional Chinese Medicine, 2007). Similarly, when the Urban Residents Basic Medical Insurance Scheme was established, it did not require any form of referral from primary care facilities to hospitals, though the document did require reimbursement rates to be slightly higher for primary care facilities.

In terms of strengthening of health service delivery, direct financial support was used to subsidize health facilities, besides increased subsidy from the demand-side via insurance payment. The reform policies since 2002 had sought to strengthen the basic-level health delivery system, including both primary care facilities and county hospitals. However, county hospitals were prioritized as they were considered the

leading facilities within counties, and were supposed to trickle down support given to them (Central Committee of Communist Party of China et al., 2009). Efforts to establish urban community health centres, expand insurance funds and increase government support also improved the function and capacity of urban primary care facilities. Bhattacharyya et al. (2011) reported increasing policy emphasis on and growth of urban community health services (converted from primary and secondary hospitals in urban areas) through their inclusion in the empanelment of social health insurance between 2001 and 2008. The proportion of government subsidies in the revenue of primary health institutions increased between 12.4 and 17.7 percentage points on average to about a third of total expenditures in various types of primary care facilities (excluding village clinics) (Meng et al., 2011). The same study also reported improvements in training, equipment, human resources, preventive and curative functions, etc.

Profit from pharmaceutical sales incentivised over-prescriptions and accounted for an unusually large proportion in revenue of health facilities. Getting rid of drug price mark-up was thus a major potential area for efficiency gain. As early as 2007, the National Congress of the Communist Party of China had declared “separation between pharmaceutical sales profit and prescription” (*yiyaofenkai*) as one of the key principles of health reform (Hu, 2007). Two policies on tightening control in terms of pharmaceuticals, i.e. zero-mark-up policy and delinking practitioners’ income from service revenue, were promulgated nationwide by 2009 as key components of the essential drugs policy (State Council, 2009).

6.4.4 Implementing health system reform (2009-2013)

In the years after 2009, detailed plans for health system reform were formulated and implemented. The sub-section below focuses on how these policies responded to and reshaped the path of coevolution between hospitals and primary care providers in

the post-juncture years.

The government continued to increase financial spending and extended coverage of social health insurance programmes. Medical insurance coverage rates rose continuously, from 29.7% in 2003 to 95.7% in 2011 (Meng et al., 2015, p. 188). Per capita funding of the NCMS rose from 30 *yuan* in 2003 to 370 *yuan* in 2013, which was almost the same as the premium level for the Urban Residents Medical Insurance Scheme (Meng et al., 2015). The increased funding covering both rural and urban areas created an opportunity to shift the balance towards primary care. While the 2009-reform did change some of the financial difficulties of primary care facilities, hospital centrism built up and reinforced over previous decades seemed to have prevailed and was even further accentuated, as the trend analysis in Chapter 4 has shown. The subsection below focuses on development related to pharmaceutical policies and payment policies, since these were most critical as influences on the coevolution of hospitals and primary care providers.

While the government declared a clear goal to strengthen primary care facilities, it had other pressing concerns, particularly the exploding demand due to rapid insurance coverage expansion as well as population ageing. The complexity of hospital reform and patients' pressing need for health care seemed to affect the structural balance, and reform in hospitals became difficult and thus delayed. Separation of revenue from pharmaceutical sales with provider revenue was delayed. An effect of this delay allowed hospitals to become ever more important in pharmaceutical sales. Based on calculation of data from the *Chinese Health Statistical Yearbooks* on the drug revenues in urban community health institutions, township health centres and public hospitals between 2008 and 2012, only 11%-15% of all drugs revenue went to urban community health institutions and township health centres, while 85%-88% went to public hospitals. Another consequence of the hospitals' role in pharmaceutical sales was their intertwining with other interests in the pharmaceutical sector (Ministry of

Health, 2009; National Health and Family Planning Commission, 2013). Scandals about payments from drug manufacturers and distributors to providers in promotion of pharmaceutical sales exposed the lasting problem of reliance on informal payment from the drug distribution system to pay physicians (Huang, 2013; Zheng, 2013). Hospitals generally continued to enjoy drug price mark-ups of 15% and rights to sell non-essential drugs, with an unknown amount of kickback for staff from the distributors and manufacturers. Health bureaus appeared to lack the regulatory power to control hospital charges, except temporarily (He, 2012).

Meanwhile, the previous lack of human capital accumulation in primary care seemed to factor in and led to challenge regarding patient confidence and satisfaction with community services. Even though patients recognized the convenience of community health services, they reported several barriers in increasing their utilisation: lack of capacity of service providers (which was also related to the lack of training in practical skills), low public awareness of the services, lack of evidence-based practice guidelines, lack of insurance coverage in some areas (Wang, Gusmano, & Cao, 2011b). Despite low levels of utilisation and staff qualification, government agencies in charge of public subsidies, pressed by the need to demonstrate quick improvement, invested most subsidy in capital assets rather than human capital (Meng et al., 2011) (also reflected in Chapter 8). In addition, the difficulty in recruiting well trained medical graduates into primary care seemed to persist (Anand et al., 2008; Hou et al., 2014). Data on recruited medical graduates at primary care levels are not available nationally, but a survey of 546 medical graduates from Fudan University showed that none went to primary care facilities between 2007 and 2010 (Fu, 2011b). Furthermore, it appeared that the already biased professional dimension of the coevolution of hospitals and primary care providers was shifting further towards hospitals as an unintended result of the post-2009 reform policies (particularly essential drug policies and emphasis on preventive services). For example, the essential medicines policy and separation of

revenue and expenditures management removed drug price mark-ups at primary care facilities, but also led to reduced income of providers. This gave rise eventually to a brain drain of experienced doctors to county hospitals, taking with them the patients (Zhou et al., 2014). Evidence of issues about imbalance in terms of professional institutions as an interrelated set of policies are further reflected in Chapter 8.

Overall, previous downstream development of hospital advantage seemed to continue playing out. Without addressing the two fundamental challenges in shifting the path of coevolution of hospitals and primary care providers, namely the profit-driven incentive structures of hospitals, and the weak institutionalization of the primary care professions, the government's response to some pressing concerns contributed to a self-reinforcing process of hospital centrism.

6.5 Discussion

6.5.1 Weaknesses and strengths

The analytical approach of the sub-study involved four sets of institutions of two sectors including numerous health facilities over a long and dynamic period of time in a vast and diverse country. The volume of interesting detail was simply too large to handle in a single study. Therefore, the sub-study tried to strike a balance between the necessary level of detail and the necessary breadth, but is unavoidably limited in that details and variations were lost about decision-making processes. Interesting details were sacrificed to make the analysis feasible. However, the breadth of the study in terms of period and coverage of institutional factors seemed instrumental in allowing a systematic examination of the historical roots of the current balance between hospitals and primary care providers.

The sub-study also had limitations regarding the sources it covered and the way these sources were interpreted. The central level archives were not accessible for post-

1949 periods, and for the local archives, documents were not disclosed for years after 2000, when collection took place. The complexity of the issue of coevolution of the hospitals and primary care providers also made it impossible to use all sources across all periods. In order to address the issue of “representativeness”, I systematically covered the journal dataset for the years between 1919 and 1949, national yearbooks and the local archives after 1949 up to 2000, and also tried to ensure no critical secondary references were left out. As each dataset had its own tendency to over-represent certain actors’ perspectives, the other information from important historians’ works and journal publications helped reduce the limitations. Despite efforts to “immerse” myself in history, potential challenges in interpretation may still arise from the author’s upbringing in post-1980 China, where there has been a dominant narrative of a strong state in development. The approach of the study also had its strengths. The path dependence analysis facilitates the study of health system development in its own history and political economy as emphasized by Bloom (2014). The approach also enhanced understanding of complex historical dynamics through using a systems perspective (Peters, 2014). The development and application of the approach marked progress from recent use of path dependence on health systems development (Balabanova et al., 2011).

6.5.2 Coevolution of hospitals and primary care providers

Table 6-2 sums up the dynamics of coevolution of hospitals and primary care providers from 1835 to 2013 (including the precursor period between 1572 and 1835) seen through a framework of path dependence. Each period studied earlier, shown in a single column in the table, seems to constitute a cycle of coevolution of hospitals and primary care providers, with the theoretically informed stages of critical juncture, post-juncture development, conjuncture and post-juncture development. First, a two-track structure evolved with the introduction of modern medicine in China from 1835 to 1949, before which health care had been largely organized privately and focused on

ambulatory care. During the critical juncture between 1835 and 1844, missionaries introduced a hospital-centric model of the medical mission, effectively establishing the initial professional binding between hospitals and professionals trained in Western medicine. The model became a consensus and spreaded slowly. Then the philanthropists brought substantial additional funding and an agenda for medical modernization and reinforced the central role of hospitals in medical professionalism.

During the conjuncture between 1928 and 1934, as the country was unified and with the rise of social medicine, a state medicine programme was launched to shift the focus of development to rural primary care. However, in the post-juncture development between 1935 and 1949, Chen Zhiqian's failed proposal of using medical school graduates as rural doctors to supervise lay health workers marked the ultimate split of hospitals and primary care providers into two tracks. By 1949, the split had turned into two very different roads of medical development: repackaged hospitals based on upgraded mission hospitals focusing primarily on curative care with increasingly professionalized staff, hovering above a primary care network that was poorly resourced.

Then, from 1949 to 1978, the previous two-tracks structure evolved into a two-sector system, embedded respectively in agricultural rural and industrializing urban contexts. The years from 1949 to 1953 appeared to be a critical juncture in shifting government priority in health development towards an exclusively urban system of health welfare, along with an overall turn towards urban industrial development. Efforts to control costs in urban areas had limited success in the late 1950s.

The conjuncture between 1965 and 1968 saw a major shift of focus towards rural and primary care, enabled by a centralization of political structures. The years after 1968 witnessed popularization of a rural system with barefoot doctors and Cooperative Medical Scheme. However, this post-juncture development also demonstrated the

difficulty of moving resources away from urban areas, and saw the relatively weaker institutionalization of rural primary care in the poorly resourced Cooperative Medical Scheme and the absence of strong professional and bureaucratic stake.

Finally, from 1978 to 2013, the dominance of hospitals expanded based on increased funding. At the critical juncture between 1978-1985, the government adopted policies which allowed generation of extra-budgetary funds through distorted incentives. Meanwhile, the already unstable CMS that financed rural primary care suffered from weakness in their institutions especially related to financing and governance and then collapsed. The loosening of financial control and the fast-growing private payment for the critical juncture reinforced institutions that favoured hospital development.

When the next conjuncture happened between 2002 and 2009, policies were made with the intention to expand insurance coverage and strengthen primary care. However, distorted incentive structures at hospitals survived the reform, owing and contributing to the continuous growth of new and complicated interests and a reliance on resource generation at hospitals to cover costs and meet demand. Both within the conjuncture and the post-juncture development, the increase of public money also disproportionately benefited hospitals.

Table 6-2. Historical logic of coevolution of hospitals and primary care providers in China (1835-2013)

Path dependence stages	First cycle (1835-1949)	Second cycle (1949-1978)	Third cycle (1978-2013)
Pre-existing conditions	1572-1835: indigenous medicine mainly based on ambulatory care	--	--
Critical juncture	1835-1844: hospitals were established to provide institutionalized curative care to establish civil engagement, exhibit technical advantage and win over religious converts	1949-1953: urban insurance and urban hospital construction created strong interest coalitions with hospitals, while rural health care focused on preventive care	1978-1985: introduction of strong financial incentives to generate extra-budget revenue favoured hospital development, while rural reform left the collective-based rural primary care facilities in lasting difficulties
Post-juncture development	1844-1901: hospitals became the recognized model for developing missionary medicine 1901-1928: hospitals were reinforced by well-resourced philanthropists as centres of high standard scientific medicine, and became the base for modern medical professionalization.	1953-1965: mediated by government policies, the two tracks started to develop into two sectors of health services embedded in China's rural and urban division; while hospitals had large outpatient departments, township health centres became the rural equivalent of hospitals	1985-2002: further autonomization of policies drove hospitals to develop a model based on heavy use of technologies and expensive medicines, while the primary care sector suffered from continued financial difficulties; rapidly widening gaps between primary care and hospital facilities
Conjuncture	1928-1934: state medicine rose to the national agenda, thanks to establishment of a central administration, the influence of international experiences, local experimentation and a historical turn towards the importance of rural areas. A nation-wide plan was published to extend modern medicine to the level of primary care	1965-1968: the “June 26th Directive” prioritizing rural health care and the centralized leadership structure during the Cultural Revolution led to policy changes that favoured coverage expansion based on cooperative medicine and barefoot doctors	2002-2009: amid the crisis of SARS and change of administration, policy agenda shifted towards universal coverage, in which primary care was emphasized as a key area for strengthening

Post-juncture development	1935-1949: the planned training of primary care professionals failed to materialise, further consolidating the split of hospitals and primary care providers into two trajectories with hospitals focusing on curative care and relying on professionals, and primary care mainly defined as epidemic control and provided by lay health personnel.	1968-1978: policy makers failed to substantially shift professional doctors away from hospitals, and eventually placed a further emphasis on rural self-sufficiency via rhetorical packaging, without solving the underlying weak institutions (financing, bureaucratic and professional stake, etc.)	2009-2013: as reform of hospital incentives was delayed, rapid insurance expansion channeled funding disproportionately towards hospitals and relatively weakened primary care providers
----------------------------------	---	---	--

The analysis suggests three cycles of coevolution between hospitals and primary care providers, and a comparison of the cycles reveals similar patterns. Each cycle started with a critical juncture (1835-1844, 1949-1953, and 1978-1985) that created a split between hospitals and primary care that was not symmetrical but biased towards hospitals. After these critical junctures, the institutional complex underpinning the two sectors saw a process of self-reinforcement that enhanced hospital dominance, via strengthening the professional, bureaucratic/philanthropic or financial stake in hospital dominance and consolidating a hospital-centred model of medical services, alongside a relatively weakly institutionalized primary care sector.

At a later point of conjuncture, major centralization of power in central leadership saw policies introduced to correct previous policy-bias towards hospitals. On the one hand, these change in policies towards greater primary care emphasis seemed to be the result of a political change external to the health delivery (i.e. the establishment of a national administration between 1928 and 1934, the Cultural Revolution (including its eve) between 1965 and 1968, and the shock of SARS and the change of political administration and agenda between 2002 and 2009). On the other hand, the conjunctures marked reactive processes against previous focus on hospitals.

With reinforcing processes preceding, the reactive processes came into the existing institutional environment that exhibited strong resilience that the reactive processes could not overcome and had to adapt. Therefore, the impacts of these reactive policies on the path of coevolution of hospitals and primary care providers were limited and further distorted the development of the two sectors. The result was repeated attempts to develop primary care as an antithesis of hospitals in all three cycles, with attempts to strengthen the four institutional pillars underpinning the coevolution of hospitals and primary care providers (professional, organizational, financial and governance institutions). The weakness of such institutionalization turned into disadvantage for primary care. Overall, the existence of the three

conjunctures suggests that policy windows for change were not rare within the Chinese political system, however, the frustrated reforms after the conjuncture imply that the path dependence effects caused by policy details and feedback effects need to be attended to.

Health system changes must be put in the broader context of each period. The first cycle was perhaps most important for the split between hospitals and primary care providers regarding types of professionals. This period saw the activism of medical reformers, coinciding with the rise of medicine as a professional realm, throughout the period when influence of modern medicine remained small and isolated in a world of traditional medicine.

The second cycle appeared to have played the biggest role in distinguishing hospital and primary care governance structures and their embedding into urban-biased state and rural communes by 1978. This period saw the strong force of political leaders coinciding with the radical extension of state power as carriers for an extensive health delivery system with basic modern technology. The previous division between hospital and primary care in relation to the dimension of traditional medicine vs biomedicine became less important as the two styles of medicine were integrated.

The third cycle appeared to be the most responsible for the large financial gap between hospitals and primary care providers by 2013. This period saw the influence of financial incentives coinciding with the continuous and high-speed economic growth of the Chinese economy, producing massive and modern hospitals which now dominate the health system in China.

To sum up, the institutional coevolution between hospitals and primary care providers in China seemed to be dynamic and cyclical. The reinforcing and reactive processes seemed to have formed the basis of cyclical policy shifts between emphasis on hospitals and emphasis on primary care providers, alongwith changing political,

social and economic landscape. However, as hospital-centric development was introduced earlier, the reactive processes were compromised by reinforcing path dependence which built hospital dominance on the urban-biased socio-political structure. The dominance of hospitals professionally, politically, organisationally, and financially persisted over all the three cycles, and seemed to build on each successive cycle. The reinforcement seemed particularly strong during the three critical junctures, when new powerful professional, political or financial forces were introduced. Together, they repeatedly benefited hospitals disproportionately, resulting in the eventual hospital dominance of 2013.

6.6 Conclusion

Chapters 5 and 6 have studied the institutional coevolution of hospitals and primary care providers between 1835 and 2013 through the lens of path dependence. First introduced by medical missionaries in 1835, the hospital-bias seemed to have experienced three cycles of “(pre-juncture—)critical juncture—post-juncture development—conjuncture—post-juncture development” and eventually resulted in the contemporary division of hospitals and primary care providers skewed towards the former. The development of institutional balance between hospitals and primary care providers suggested a dominance of hospitals in the four institutional pillars underpinning the coevolution of hospitals and primary care providers (professional, financial, governance-related and organizational). The approach of path dependence analysis seemed helpful in making sense of the persistence of hospital-centrism in China between 1835 and 2013 by identifying a long term and cyclical logic grounded in history.

While discussion of the implications of the findings is provided in Chapter 9, here it is worthwhile to note that by tracing the roots of current balance between primary care providers and hospitals over their long co-evolutionary history to 1835, the thesis is not seeking to assign blame to some foreigners that set in motion a perverse health

system structure that plagued China's modern history. Instead, it has provided some rich understanding of the centrality of temporal dynamics to the issue of balance between primary care providers and hospitals.

Indeed, the historical-institutionalist analysis of the coevolution of the hospitals and primary care providers over the history of China suggests that the apparent path dependence of hospital centrism was not rooted in foreclosing policy episodes, as key moments of major change emerged periodically. Instead, one key source of path dependence seemed to be the tendency to respond to short term policy needs at critical junctures, which ended up creating potentially adverse and irreversible consequences. The other key source of path dependence seemed to be the repeated frustration of implementing reform after the conjunctures, suggesting the crux of overcoming path dependence in this context lies in establishing mechanisms to maintain momentum, perhaps by increasing bureaucratic and professional stakes in primary care, and through coordinated and cumulative efforts to harness political, bureaucratic and professional dynamics to secure reform implementation. The thesis will revisit the findings from these historical analyses after examining a contemporary gatekeeping pilot aimed at strengthening primary care.

Part III A contemporary gatekeeping pilot

Chapter 7 Impact evaluation of the gatekeeping pilot

7.1 Introduction

Part II tracked the historical trends of the balance between hospitals and primary care providers in China and traced the historical coevolution of the two sectors. It revealed the challenges facing primary care strengthening. Part III turns the focus towards a contemporary pilot that attempted to strengthen the functioning of primary care via gatekeeping in the setting of two rural townships in China. Although initiated locally, the pilot was a national pioneer and drew considerable public attention. This part consists of two chapters, which together constitute a mixed-methods case study of the pilot. Chapter 7 provides a quantitative impact evaluation of the pilot, and Chapter 8 assesses qualitatively how the pilot functioned.

As identified in Chapter 2, evidence regarding the effectiveness of gatekeeping has been scant and of low quality, and particularly so in low- and middle-income countries. Despite claims that gatekeeping would be an effective solution to the issue of weak primary care in China (Zhou et al., 2014), no empirical evaluation has been reported. In high-income countries, studies on gatekeeping generally focus on shifting care between specialists and generalists (i.e. general practitioners, internal medicine residents, and family medicine doctors), who are both qualified with full medical degrees. The context of China, as shown in previous chapters, involves a greater professional imbalance between hospitals and primary care providers, where practitioners usually do not have full medical training. Therefore, it is of interest to see how effective a Chinese version of gatekeeping could be in achieving a rational

distribution of patient care seeking. An impact evaluation of the effects of the pilot in shifting patient ambulatory visits and spending at primary care facilities is presented in this chapter.

7.2 Methods

7.2.1 The gatekeeping pilot

The gatekeeping pilot took place in the context of the New Rural Cooperative Medical Scheme (NCMS) and in a suburban district of a large metropolitan city in northern China. The NCMS is a nation-wide rural insurance scheme (i.e. for people with rural household registration) with a heavily tax-subsidized premium and high co-payment rates; it was launched in China in 2003 (Meng & Xu, 2014b). As in other areas, the NCMS fund in the pilot district (a suburban district of a large city) was pooled at district (comparable to county) level and managed by an NCMS service centre under the District Health Bureau¹¹. Each county has been in charge of making its own policies, though counties usually adhere to the guidelines made by the central government and their superior provincial governments, and these need approval from their superior municipal government.

The district had a population of 0.42 million, among which 0.2 million were registered as rural residents. Rural per capita income was 16,865 *yuan*¹² in 2012. Over 99% of the eligible population were enrolled in the NCMS. During the three years between 2012 and 2014, each enrollee paid a premium contribution of 100 *yuan* every year and participated through the unit of a household, while government subsidy in the

¹¹ County and district are comparable and are the level of government immediately below municipality. No specific law or regulation was found regarding the differences between county and district. Generally speaking, a county has a higher proportion of rural residents, while a district has a higher proportion of urban residents.

¹² The annual average exchange rate between Chinese *yuan* and US dollars was 631.25 : 100 (National Bureau of Statistics of China, 2015).

premium increased from 540 to 900 *yuan* per year. Local government paid the premium for people entitled to a subsistence allowance or “five guarantees” (including food, clothing, fuel, education and burial expenses) and for other people to whom policies for subsistence allowance applied.

Before the pilot, patients generally had unrestricted access to health care facilities within the municipality, including municipal and district level hospitals and primary care facilities (township health centres and village clinics, most of which had been integrated with township health centres). However, the district’s NCMS reimbursement policy for ambulatory services favoured care-seeking at primary care facilities, as NCMS schemes in other areas frequently did. To encourage utilisation of services at primary care level, favourable deductible and reimbursement rates applied if patients used ambulatory services at township health centres (community health centres) or village clinics (community health stations).

The deductible for ambulatory services at primary care facilities was 100 *yuan*, while the reimbursement rate was 50%. The deductibles for ambulatory services at “secondary hospitals” and “tertiary hospitals” were 500 *yuan*, while the reimbursement rates were 35% and 30% respectively. It is worth noticing that the local policy’s classifications of “secondary hospitals” and “tertiary hospitals” were not strictly following their technical accreditation. According to local NCMS documents, policies for secondary hospitals applied for all district-level hospitals, including the District People’s Hospital, technically a tertiary hospital. Meanwhile policies for tertiary hospitals applied to all hospitals beyond the district. Special ambulatory services, including radio- and chemo-therapy for cancers, renal dialysis for terminal renal diseases, anti-rejection drugs used after kidney transplantation, combined liver and kidney transplantation and heart transplantation, and paediatric aplastic anaemia and haemophilia, had higher reimbursement rates (75% in secondary hospitals and 60% in tertiary hospitals).

The pilot was launched in 2 (among a total of 17) townships of the district in 2013. Some basic information about the pilot and non-pilot areas is shown in Figure 7-1. The 2 pilot townships on average had 4,235 enrolees, and the 15 non-pilot townships on average had 15,345 enrolees. Both groups of townships saw a small decline of enrollment, as average enrolment fell to 14,793 in non-pilot townships and 4,116 in pilot townships by 2014. This was a continuation of a previous decline from 2012 to 2013. According to local directors of township health centres, the decline in enrollment was due to a shrinkage of eligible population (i.e. a net emigration of population). In particular, some young people went to schools and colleges in the district seat or outside the district, graduated and found jobs in urban areas, became urban-based or changed their household registration type from rural to urban. A slightly larger percentage of these emigrants were male.

Average percentages of female enrollees were 51.3% in both pilot and non-pilot townships in 2013, and both slightly increased from 2012 to 2014. Average age was 45.5 years old in non-pilot townships and 44.2 years old in pilot townships, and both also slightly increased from 2012 to 2014. Each township had a township health centre. One of the pilot townships had no village-level community health station, while the other had three. As the non-pilot townships tended to be larger, they had on average 8.6 village-level community health stations.

In July 2013, two rural townships in the study district introduced a gatekeeping pilot programme. The pilot involved a change of incentives for both service users and providers. From the perspective of local beneficiaries of the NCMS, the reform required patients to obtain a referral letter from primary care providers (i.e. township health centres and their subordinate village health stations) for accessing care at ambulatory departments of secondary hospitals. To access ambulatory departments of tertiary hospitals, patients needed a referral letter from the District People's Hospital if they wished to claim reimbursement for their expenditures.

Patients could opt out of the system through paying out-of-pocket and forego their claim for reimbursement, as health facilities accepted privately funded patients. In other words, the pilot differed from a strict gatekeeping policy where patients without referral letters from general practitioners would be denied access to hospital consultants. In case of emergency, patients could go directly to emergency departments in hospitals and would not need referral for reimbursement.

On the supply side, the township health centres (including their subordinate village health stations) in the pilot areas were given an annual global budget for ambulatory services reimbursement and were responsible for all the expenditures covered by the NCMS for ambulatory services at both hospitals and primary care facilities. The annual budget was calculated according to the level of ambulatory reimbursement per enrollee in 2012 (235 *yuan* in one township and 133 *yuan* in the other)--the year prior to the start of the pilot, with a minor 5% increment year-on-year. In case of surplus at the end of the financial year, the remaining funds would be kept by the facilities as a bonus for gatekeeping and so-called family-practitioner-style services (i.e. chronic illness management, health records, etc), though not for salary (see next chapter). In case of deficit, the facilities would bear the cost of the amount needed to reimburse patients. Payment for primary care facilities was not related to number of referrals. By contrast, ambulatory services in district-level hospitals were reimbursed based on fee-for-service.

While the referral requirement incentivized patients to go through service providers level-by-level, i.e. from primary care providers to secondary and then to tertiary if needed, the capitation-based ambulatory reimbursement budget incentivized (at least in theory) providers to retain and treat patients in their facilities. Combining the user and provider incentives, the policy was intended to rationalize patient flows and hence was expected to shift ambulatory patient visits and service expenditures from hospital ambulatory departments to primary care providers. In other words, the

expected outcomes included: ambulatory visits to primary care facilities would go up; ambulatory expenditures in primary care facilities would increase; ambulatory visits to tertiary hospitals would reduce; ambulatory visits to tertiary hospitals would reduce; ambulatory expenditures at tertiary hospitals would reduce in those facilities with reduced total ambulatory expenditures. It was less clear for the secondary hospitals, as they were intended to gain patients from tertiary hospitals, while patients that previously would have visited secondary hospitals were supposed to visit primary care facilities instead.

Organization of the pilot implementation involved the municipal and district health bureaux, township health centres and village doctors in pilot areas. They mainly disseminated the information through brochures and health education lectures. Other reform pilot programmes in the study district in July 2013 included payment reform launched in July 2013 in the entire district, which changed reimbursement for inpatient expenditures from payment based on fee-for-service to payment based on diagnosis-related groups for short-term hospitalizations and payment per inpatient day for long term hospitalisations (longer than 60 days).

Reform policies both within and beyond the gatekeeping pilot townships underwent adjustment in the year 2014. In order to “compensate” the beneficiaries in the pilot townships for the restriction of choice, the reimbursement rate for primary care facilities in the pilot townships was slightly increased, to 52% in 2014. An intended scaling up of the gatekeeping pilot did not take place. However, for ambulatory services in hospitals outside the district, referral policies were implemented in non-pilot townships of the district in 2014. In other words, if patients sought ambulatory services in tertiary hospitals without a referral letter from the District People’s Hospital, only 80% of reimbursement would be paid to patients. It should be noted that patients in pilot areas would not get any reimbursement if they went to a tertiary hospital directly.

7.2.2 Difference-in-differences study design

The study estimated the impacts of the gatekeeping pilot on utilisation and total expenditures at different levels of care. The original intention was to include an analysis of the effects on ambulatory care sensitive hospitalisations. This became problematic first due to data issues, as the claims dataset did not use international standard disease coding (e.g. ICD-10) for diagnosis but rather used string text. In an attempt to translate the diagnosis into ICD-10 coding, most of the recorded diagnoses were unmatched with ICD-10. In addition, there was another reform pilot, on inpatient services, launched in the same district where the gatekeeping pilot townships were located, potentially influencing the outcome. Hence, health outcome was not studied here.

Ideally, for impact evaluation, participation in the gatekeeping intervention would have been assigned randomly. However, in this study, investigators were not involved in assigning the intervention. In other words, the study was observational. Indeed, the choice of two relatively smaller townships as pilots for gatekeeping was due to a concern about the potentially controversial nature of the pilot and intended to minimise risks. As Figure 7-1 summarises, the chapter adopted the study design of a difference-in-differences analysis (introduced in Chapter 2). The enrollees in the two townships under the gatekeeping pilot programme formed the treatment group, and those in the 15 other townships (sub-districts) formed a control group. Using data both before (including 6 waves, with each quarter as a wave) and after (also including 6 waves, with each quarter as a wave) the launch of the gatekeeping pilot, effects on outcome variables were tested through difference-in-differences analysis which compared the changes in outcome variables over time between pilot townships and non-pilot townships.

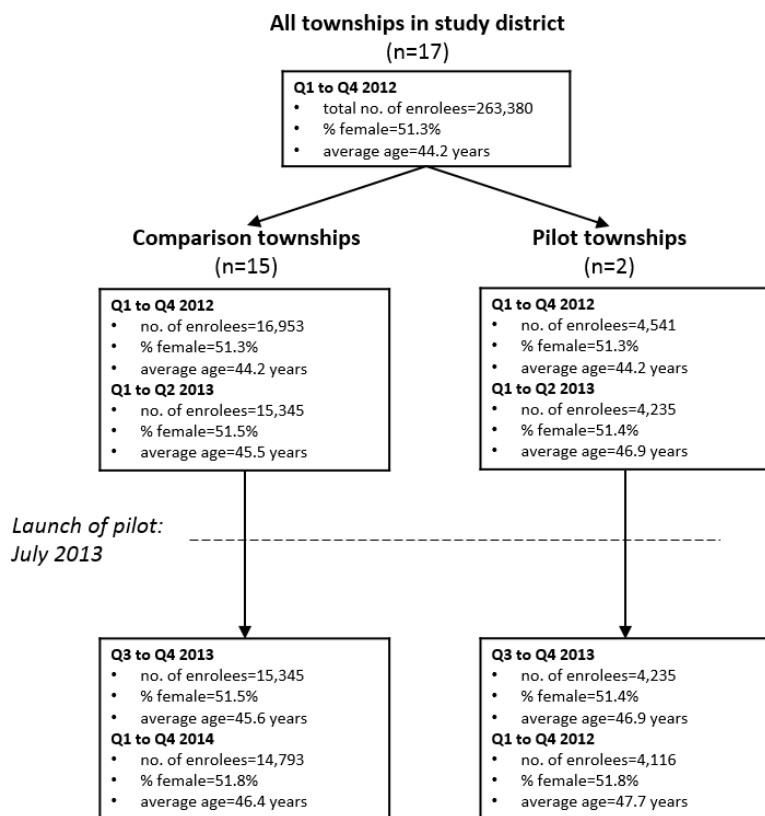


Figure 7-1. Schematic representation of study design

Note: unless specified, all the numbers in the boxes are averaged numbers for all townships in the specific group.

7.2.3 Data

The sub-study covered the period from January 2012 to December 2014 and used claims and enrolment data regarding the enrolees of the district, provided by the Municipal Management Centre of the NCMS, where data were centralized. The claims database recorded information on each claimed episode of patient care, including total amount of expenditure, amount reimbursed, amount of expenditures not covered by the NCMS¹³, name and level of facilities, date of the visit, and personal information on the beneficiaries (age, gender, personal identification number, and village and

¹³ As actual reimbursement rate was an indicator of performance of NCMS, the expenditures not covered were routinely collected.

township of residence).

From January 2012 to December 2014, the reimbursement procedure relied on manual settlement. In other words, patients had to pay the full amount of their ambulatory expenditure when they utilised ambulatory care at health facilities. Township health centres and village cadres then collected patients' receipts and brought them to the township financial department, where eligible expenditures were reimbursed. Staff at the township level then manually entered the claims information into the same claims database following the same structure. This information was centralized at the municipal level and became the claims database used in the study. The municipal NCMS centre also maintained a full list of enrolees for each year, including information on age, gender, personal identification number, and village and township of residence.

The list of enrolees was merged with the list of claims (each claim was taken as one visit) into a full list of enrolees with information regarding their utilisation of services. Data were then aggregated at the level of each quarter of a year so that seasonality caused by the change of claims process was relatively smoothed. The data were also aggregated at the level of townships, as the township health centres were the key unit of experiment, though the level of village was also used to test the robustness of the analysis.

Outcome variables were defined and constructed as follows. Ambulatory visits to each level of (i.e. primary, secondary, and tertiary) facilities per enrolee per quarter, were the total number of patient observations with recorded visits to facilities at the corresponding levels of facilities in each quarter of the year in every township, divided by the number of enrolees in the township during for that year (it was not possible drop out or enrol during the year). Ambulatory expenditures at each level of (i.e. primary, secondary, and tertiary) facilities and at all facilities per enrolee per quarter were

defined as the ambulatory expenditures in the corresponding levels of facilities.

7.2.4 Econometric analysis

The following difference-in-differences model was used for the analysis of the effects of the pilot programme:

$$Y_{it} = \alpha_i + \eta_i + v_t + \delta \times Treatment_i \times After_t + \beta X_{it} + \varepsilon_{it} \quad (1)$$

Function (1) is a model of difference-in-differences analysis involving multiple periods (in this case 12 quarters over four years). α_i is a constant variable, η_i represents the fixed effects of each township, and v_t represents the fixed effects of each time period (t representing the number of quarters since the start of the data period). The existence of v_t also adjusts for seasonal effects. $Treatment_i$ is a dummy variable, which takes the value of 1 if the township (i) was in the treatment group (among the gatekeeping pilot areas), 0 if not. $After_t$ is a dummy variable taking a value of 1 if the time period (t) was after the start of the pilot and 0 if before.

The coefficient of the interaction term $Treatment_i \times After_t$, i.e. δ , captures the effects of the pilot policy. ε_{it} is a random error term with zero mean. The main effects of $Treatment_i$ are omitted due to their perfect collinearity with η_i , while the main effects of $After_t$ are omitted due to their perfect collinearity with v_t . In function (1) the covariates—available time-variant variables of township with possible effects on the outcome variables (X_{it}) are included (i.e. average age, and proportion of female enrolees). In addition, a simple difference-in-differences model (i.e. not including v_t or βX_{it}) and an unadjusted model (i.e. not including βX_{it}) based on function (1) were used to estimate the effects.

Three sensitivity analyses were conducted. The first verified that the results were not sensitive to alternative levels of aggregation by the use of the village level as the unit of aggregation. The second verified that observation weighting and clustering on

standard errors did not affect the results. The third was a pre-trend test to see whether the treatment group and the control group had similar trends before the pilot. For the pre-trend test, the following equation was used with data from before the pilot:

$$Y_{it} = \alpha_4 + \eta_i + \beta \times T + \rho \times Treatment_i \times T + \varepsilon_{it} \quad (2)$$

In function (2), β is the pre-pilot trend of the townships in the control group, while ρ captures the difference in the pre-pilot trend between treatment and control groups. T is a linear time variable taking the value from 1 to 12 for each quarter from the first quarter of 2012 to the last of 2014.

7.3 Results

Table 7-1 presents the average outcomes in the pilot and non-pilot townships before and after the launch of the pilot, as well as simple calculation of the difference in differences. From the descriptive results, the pilot had an obvious impact on increasing visits to primary care facilities, as well as expenditures at this level (although the growth in expenditures was small relative to the growth in visits). Average ambulatory visits per enrollee per quarter increased in primary care facilities in pilot townships from 0.261 to 0.429. 0.112 out of the increase (0.168) was attributable to the pilot programme. Ambulatory expenditures at primary care facilities in pilot townships also increased from 49.613 *yuan* to 57.612 *yuan*. 7.310 *yuan* out of the increase (7.999 *yuan*) was attributable to the pilot programme.

The influence on secondary hospitals was less obvious. The ambulatory visits at secondary hospitals reduced from 0.093 to 0.075; 0.012 out of the decrease (0.018) was attributable to the pilot programme, while the ambulatory expenditures at this level increased slightly, from 67.054 *yuan* to 71.273 *yuan*, with 1.758 *yuan* out of the increase (4.219 *yuan*) attributable to the pilot programme. Average ambulatory visits to tertiary hospitals in pilot townships reduced from 0.011 to 0.007; 0.002 out of the

reduction (0.004) was attributable to the pilot; ambulatory expenditures at tertiary hospitals fell considerably, from 30.861 *yuan* to 14.195 *yuan*; 14.007 *yuan* out of the decrease (16.666 *yuan*) was attributable to the pilot programme.

Overall, the pilot townships saw a reduction of 4.448 *yuan* in average total ambulatory expenditures at all three levels, from 147.528 *yuan* to 143.080 *yuan* (the net effect of the pilot programme would be 4.939 *yuan*, considering the average total ambulatory expenditures rose in non-pilot townships by 0.492 *yuan*).

Table 7-2 reports the results of the statistical analysis for visits using three difference-in-differences models: simple difference-in-differences without covariates, multiple-period difference-in-differences without covariates, and multiple-period difference-in-differences analysis with covariates. The main text here shows only the results from the unweighted models using township as the level of aggregation. The results are similar to the results generated with the weighted model using township as the level of aggregation, as well as results for weighted model using villages as the level of aggregation. The unweighted model using villages as the level of aggregation showed some discrepancies in results compared with the other three models, due to the change in number of enrollees as a result of urbanization. The results from all three alternative models are shown in Appendix Table 3.

Ambulatory visits at primary care facilities per enrollee per quarter saw a significant increase, by 0.101 or 38.7% of 0.261--the average number of visits to primary care facilities per enrollee per quarter in the pilot areas before the pilot ($p=0.013$). By contrast, the number of visits to secondary hospitals per enrollee per quarter reduced by 0.01, or by 10.7% from 0.093--the average number of visits in the pilot areas before the pilot, but the evidence is weak ($p=0.332$); visits to tertiary hospitals reduced by 0.001, or 7.8% of 0.011--the visits to tertiary hospitals before the pilot, but again the evidence is weak ($p=0.428$).

Table 7-3 shows the regression results regarding expenditure. As for visits, the main text here only shows the results from the unweighted model using township as the level of aggregation. Results from the weighted model using township as the level of aggregation, as well as results for weighted and unweighted models using villages as the level of aggregation, are shown in Appendix Table 4. Ambulatory expenditures at primary care facilities per enrollee per quarter increased by 7.445 *yuan*, or 15% of the 49.613 *yuan* in pilot areas before the pilot, though evidence supporting this was not very strong ($p=0.116$). Ambulatory expenditures at secondary hospitals increased by 5.225 *yuan*, or 7.8% of the 67.054 *yuan* in pilot areas before the pilot. Again, the evidence was not very strong ($p=0.146$). Spending at tertiary hospitals reduced significantly, by 13.17 *yuan* ($p=0.020$) per enrollee per quarter, or 42.7% from the 30.861 *yuan* in pilot townships before the pilot. The total ambulatory expenditures were 147.528 *yuan* before the pilot.

Table 7-1. Outcome variables of enrollees in pilot and non-pilot townships before and after the start of the gatekeeping pilot

Characteristics	Pilot townships			Non-pilot townships			Difference in differences
	<i>Before introduction</i>	<i>After introduction</i>	<i>Difference</i>	<i>Before introduction</i>	<i>After introduction</i>	<i>Difference</i>	
<i>(Per enrollee per quarter per township) *</i>							
Ambulatory visits to primary care facilities	0.261	0.429	0.168	0.196	0.252	0.056	0.112
Ambulatory visits to secondary hospitals	0.093	0.075	-0.018	0.095	0.089	-0.006	-0.012
Ambulatory visits to tertiary hospitals	0.011	0.007	-0.004	0.014	0.011	-0.002	-0.002
Proportion of ambulatory visits at primary care out of total ambulatory visits	0.602	0.814	0.212	0.587	0.696	0.109	0.103
Total ambulatory expenditures at primary care (yuan)	49.613	57.612	7.999	62.000	62.689	0.689	7.310
Total ambulatory expenditures at secondary hospitals (yuan)	67.054	71.273	4.219	71.632	74.093	2.461	1.758
Total ambulatory expenditures at tertiary hospitals (yuan)	30.861	14.195	-16.666	21.535	18.876	-2.659	-14.007
Total expenditures on ambulatory care	147.528	143.080	-4.448	155.167	155.659	0.492	-4.939
Proportion of total ambulatory expenditures at primary care facilities out of all ambulatory expenditures	0.355	0.442	0.087	0.363	0.413	0.050	0.037

Notes: * The numbers were calculated at the level of township per enrollee per quarter.

Table 7-2. Regression results on visits

Regression models	Ambulatory visits at primary care facilities per enrollee per quarter	Ambulatory visits at secondary hospitals per enrollee per quarter	Ambulatory visits at tertiary hospitals per enrollee per quarter	Proportion of ambulatory visits at primary care in total ambulatory visits
Simple difference-in-differences (DD)	0.112*** (0.000)	-0.0124 (0.369)	-0.00168 (0.465)	0.103 (0.318)
Multiple-period DD	0.112*** (0.000)	-0.0124 (0.380)	-0.00168 (0.475)	0.103 (0.329)
Multiple-period DD with covariates	0.101** (0.013)	-0.00999 (0.332)	-0.000862 (0.428)	0.0798* (0.094)
Pre-trend test	0.000274 (0.984)	-0.00230 (0.587)	-0.00112** (0.046)	-0.0128 (0.238)
Mean outcome in pilot townships before introduction of pilot programme	0.261	0.093	0.011	0.602

Note:

a) p values are shown in brackets.

b) $*0.05 \leq p < 0.1$; $** 0.01 \leq p < 0.05$, $*** p < 0.01$.

Table 7-3. Regression results on expenditures

Regression models	Total ambulatory expenditures at primary care per enrollee per quarter	Total ambulatory expenditures at secondary hospitals per enrollee per quarter	Total ambulatory expenditures at tertiary hospitals per enrollee per quarter	Total expenditures on ambulatory care	Proportion of total ambulatory expenditures at primary care facilities in all ambulatory expenditures
Simple DD	7.310 (0.112)	1.758 (0.831)	-14.01** (0.026)	-4.939 (0.784)	0.0372 (0.609)
Multiple-period DD	7.310 (0.120)	1.758 (0.835)	-14.01** (0.029)	-4.939 (0.789)	0.0372 (0.617)
Multiple-period DD with covariates	7.445 (0.116)	5.225 (0.146)	-13.17** (0.020)	-0.499 (0.965)	0.0165 (0.529)
Pre-trend test	-3.668 (0.251)	1.372 (0.729)	0.777 (0.856)	-1.519 (0.799)	-0.0123 (0.523)
Mean outcome in pilot townships before introduction of pilot programme	49.613	67.054	30.861	147.528	0.355

Note:

a) p values are shown in brackets.

b) $*0.05 \leq p < 0.1$; $**0.01 \leq p < 0.05$, $*** p < 0.01$.

The pre-trend test attempted to check for potential differences in the pre-pilot trends of the outcome variables between pilot and non-pilot areas. Regression results of the test are shown in Table 7-2 and Table 7-3. There were no significant differences between pilot and non-pilot areas in terms of ambulatory visits at primary facilities and secondary facilities, proportion of ambulatory visits at primary care out of total ambulatory care, and ambulatory expenditures at each and all levels of facilities per enrollee, and the proportion of expenditures at primary care facilities. For ambulatory visits at tertiary hospitals per enrollee, the coefficient representing the pre-trend was -0.001 ($p=0.046$, with average number of visits at 0.011 per enrollee per quarter before the pilot), suggesting that visits to tertiary hospitals had been on a moderately declining trend compared to non-pilot areas. This appeared to further weakened the support for a declining trend of patient visits to tertiary hospitals. Overall, the results in general strengthened the confidence in results of the difference-in-differences analysis.

7.4 Discussion

7.4.1 Strengths and limitations

The sub-study used the difference-in-differences method to evaluate the impact of the gatekeeping pilot. The approach has both strengths and limitations. The difference-in-differences analysis established and used counterfactuals to calculate the net effect of the pilot intervention, and was therefore superior to the only previous study of the effect of gatekeeping from LMICs which used a before-and-after study. The study further controlled for covariates, while the previous study used aggregate level data.

A potential limitation of the difference-in-differences approach is that it attributes any difference in trends between the pilot and non-pilot townships to the pilot programme. As in previous studies about gatekeeping, the gatekeeping pilot was among a package of payment reforms, as well as a range of other on-going policy changes. The sub-study analysed the key policy documents that might have influenced the pilot and information about the facilities, and checked their implementation across

pilot and non-pilot townships. No particular difference in policies and implementation seemed likely to have considerably affected the outcome.

The sub-study used claims data, like several previous studies on gatekeeping in high-income countries (Garrido et al., 2011). A key problem with claims data in general, and in this sub-study, was related to the fact that they were collected for reimbursement rather than to answer the specific research questions (Ferver, Burton, & Jesilow, 2009). There were some key pros and cons regarding the use of claims data in this sub-study.

First, if patients did not expect to be reimbursed for the services and medicines received at facilities, they might not bother to claim for reimbursement (particularly when the amount was small). Therefore, there might have been issues of underreporting of visits with small amounts of expenditure. Nevertheless, the likelihood of such “under-claiming” seemed mitigated by the network of reimbursement management team reached all villages.

Second, another type of underreporting that potentially weakened the study was bundling. This was the situation where persons who processed the claims combined receipts together for more than one episode of ambulatory care visits. As the local NCMS managers admitted, bundling was a common practice particularly because claims were manually entered into the database during the study period. This might result in under-reporting of the number of visits.

Data originally acquired covered the five years from 2010 to 2014. Based on claims data from the two years of 2010 and 2011, when an electronic system was used to automatically generate records of care seeking, the average number of ambulatory visits was 6.35 per year. The number decreased to 1.68 on average for 2012, 2013 and 2014. Due to this radical change, the first two years were excluded from the analysis. The implication is that the actual change in number of visits would have been

substantially larger, though the proportion seemed likely to remain the same.

The changed reimbursement procedure also gave rise to artificial seasonality, and the number of observations fluctuated during the study period. Another issue created by bundling was that the time of visits was affected. This issue of bundling affected both pilot and control areas, and hence it was assumed that the dataset would still be useful for the analysis. The bundling issue is unlikely to influence the information on ambulatory expenditures across different levels of care, as claims processing staff would need to calculate the amount of reimbursement according to the particular level of care.

Third, the information regarding diagnosis and socio-economics of enrolees was problematic. As mentioned earlier, the diagnosis information was not standardized according to the international disease coding system, which prevented further analysis regarding the effects of the pilot programme on hospitalisations that would be sensitive to changes in primary care performance. Due also to the practice of bundling, the diagnosis information for ambulatory visits was confused, which also prevented potentially meaningful investigation. Socioeconomic information on claimants in pilot and control areas was not available, so their potential effects on the outcome variables could not be controlled for.

Despite these drawbacks, there was also important advantage in using claims data here. A major advantage was generic to secondary data analysis, in that claims data were available electronically, and using claims data was economical (Ferver et al., 2009). Few resources were needed in collection and entry of the data. Secondly, the use of claims data here had the advantage of automatic capturing of information about the “population”. Using claims data avoided the potential pitfalls of sampling, low response to surveys, loss of follow-up, etc.

Thirdly, the claims data used here captured all recorded use of services (though

primarily for the purpose of reimbursement) of patients. Therefore, recall bias--a problem in surveys--was avoided in the sub-study. Fourth, the use of claims data also avoided the problem of relying on administrative data, which were prone to manipulation because of performance monitoring of public servants. Finally, the accuracy of the data (again from the perspective of reimbursement) was checked by a system of monitoring; as reimbursement was involved, both patients and the staff entering and managing the data were concerned to ensure data accuracy.

7.4.2 Findings

The findings of this first impact evaluation of gatekeeping in China add some new knowledge to the literature on gatekeeping. First, there is strong evidence ($p=0.013$) that more patients (an average increase of 38.7%) visited primary care facilities because of the gatekeeping pilot programme. Expenditures at primary care facilities also increased, by 15%, though the evidence is not very strong ($p=0.116$). The increased visits to primary care facilities (0.101 per enrollee per quarter) far outweighed the decreased visits to secondary hospitals (-0.001) and tertiary hospitals (-0.001).

Furthermore, the treatment intensity in the additional primary care visits seemed to be low. Had gatekeeping been effective in shifting patient services to primary care facilities from secondary hospitals, then service expenditures at primary care facilities would be expected to increase more than the number of visits, as secondary hospitals are expected to treat sicker and more resource-demanding patients. The increased primary care visits exceeded the increase in primary care expenditures, suggesting that most patients shifting from secondary hospitals to primary care facilities were not receiving much treatment in primary care facilities. This could have been the case if the visit was mainly to get a referral letter to higher level of care. As the reimbursement rate was only marginally changed, it was unlikely that the increased visits were caused by increased demand due to lower cost.

The changes in secondary hospitals (district-level hospitals) were not significant in both patient visits and expenditures, suggesting that secondary hospitals could have been shifting some patients to primary care facilities while gaining from tertiary hospitals. There was no evidence of a decreasing number of patient visits to tertiary hospitals. However, the expenditures in tertiary hospitals (outside the district) reduced significantly, by 42.7% ($p=0.020$). The opposite effects on visits and expenditures at tertiary care facilities could have been due to the influence of bundling (mainly affecting information regarding visits), which could have become particularly problematic with the smaller number of visits to this level of care as compared to other levels. The effect of gatekeeping on total expenditures on ambulatory care was not obvious.

The gatekeeping pilot programme therefore seemed to be effective in shifting expenditures from extra-district hospitals to facilities within the district. The significant decrease of treatment expenditures happened in extra-district tertiary hospitals, suggesting that the District People's Hospital with tertiary facilities may be able to play a gatekeeping role in preventing care seeking at higher level. Since the reimbursement rate was higher for care at lower levels, it is likely that patients benefited from lower out-of-pocket expenditures. The decrease of expenditures at tertiary hospitals in this pilot was not a big issue for hospital income, as these patients were not their main customers. However, scaling up gatekeeping could potentially mean more challenge from hospital related interests.

7.5 Conclusion

Gatekeeping has been under-studied, particularly in low- and middle-income countries (LMICs). This first impact evaluation of gatekeeping in a LMIC used a difference-in-differences analysis to evaluate a gatekeeping pilot programme in rural China. The analysis suggests that gatekeeping between primary care and hospitals increased patient visits to primary care providers relative to hospitals; however, the

largely unchanged level of expenditures suggests that most patients may have been visiting for a referral letter. While a small proportion of patients did seem to move to primary care for treatment, the evidence was not strong. The drastic decrease of patient expenditures in extra-district tertiary hospitals implies that the District People's Hospital may have had a strong role preventing care seeking at a higher level. This seems to suggest that strong service capacity is key for effective gatekeeping.

Despite the enormous and growing amount of routinely collected and digitized claims data, health systems studies in China have rarely used this form of data. Methodologically, the study has made a contribution by demonstrating the possibility of doing policy-relevant research using such real world data.

Chapter 8 A qualitative systems analysis of the gatekeeping pilot

8.1 Introduction

Chapter 7 provided an impact evaluation of a gatekeeping pilot programme in two rural townships in a large municipality in northern China. The primary concern of Chapter 8 is to understand qualitatively how the gatekeeping pilot functioned, more specifically: the intended and the actual functioning of the gatekeeping pilot; factors that likely explained the gap between the intended and actual functioning; and consequences of the gatekeeping pilot. Drawing on recent developments in the field of health systems and policy research (Adam, 2014; Alliance for Health Policy and Systems Research, 2009), and expecting complexity of gatekeeping in rural China, the sub-study combined causal loop analysis with a health system building blocks categorisation and a qualitative policy analysis method to study the functioning of the gatekeeping pilot.

Indeed, previous studies from China have suggested that a large proportion of patients treated in hospitals could be managed more cost-effectively at lower levels of care (Lei et al., 1996; Wang, Yin, & Song, 2011; Xu, Wang, et al., 2007), implying huge potential for gatekeeping. However, a literature review exposed a paucity of research articles about reform pilots in China involving gatekeeping (Li & Li, 2008). Furthermore, health system changes in recent decades that affected both the primary care sector and hospitals profoundly (Bloom & Wolcott, 2013; Sun, Wang, & Barnes, 2015; Yip & Hsiao, 2015; Zhou et al., 2014) are likely to influence the functioning of gatekeeping.

The various arrangements, context-specificity, multiple and interrelated impacts, as well as the nature of the gatekeeping pilot suggested the need for an approach that

allows sufficient sensitivity to and synthesis of the multiple factors in the interrelating dynamics. Systems thinking has been described as a mind-set which sees systems and sub-components of systems as interrelated to one another, and interprets the interrelationships as the key to knowledge about how things function (Adam, 2014). Advocated as useful for health systems and policy research, systems thinking has proved valuable in revealing key elements of success and failure in implementing complex interventions, including the role and importance of relationships, actors in health systems, environmental factors, anticipating potential unintended consequences, and systematically evaluating the implementation process and reactions to feedbacks within the systems (Adam et al., 2012; Adam, 2014; de Savigny & Adam, 2009; Gilson, 2012). Combining qualitative methods with systems thinking can add depth to analysis of health systems issues, and adding visualization can help convey complex interpretations and findings (Adam, 2014).

8.2 Methods

Causal loop analysis is a method among the tools of applied systems thinking. It maps out and qualitatively models the dynamics amongst a number of interconnected factors using causal loop diagrams (CLDs). Recent application of CLDs was seen in the field of health policy and systems research (Rwashana & Williams, 2007), neonatal mortality (Rwashana et al., 2014), medical dual practice (Paina et al., 2014), and integrated Community Case Management (iCCM) of malaria, pneumonia and diarrhoea (Sarriot et al., 2015). In these studies, CLDs made explicit cause-and-effect relationships and facilitated understanding and interpretation of interacting factors and feedback loops that contribute to important policy issues. Causal loop analysis has not been used to study gatekeeping.

As a tool, CLDs do not automatically generate the information needed for constructing them. Sterman suggested that data collection and analysis should be based on qualitative methods (Sterman, 2000), however, there has been little guidance on

how to rigorously generate CLDs from qualitative interviews. It has also been unclear how to cover the range of health systems issues that could influence the functioning of complex interventions like gatekeeping. Therefore, the sub-study linked causal loop analysis with the WHO classification of health system building blocks and the “Framework” approach for qualitative data analysis. The overall process of the study consisted of five stages as shown in Table 8-1.

Table 8-1. Overall process of the qualitative systems analysis

Stage	Content
1	Developing preliminary thematic framework and research tools
2	Fieldwork and interviews
3	Initial analysis of interview transcripts
4	Interpretation of data and tabulation
5	Construction of causal loop diagrams and analysis

The first stage followed the “Framework” approach and developed a preliminary thematic framework and research tools, with the aid of the WHO categorisation of health system building blocks. National and local policy documents were collected from the municipal and district health bureaux and central government and analysed. Then a preliminary thematic framework was developed, leading to question guides that included questions about implementation and intended mechanisms of the gatekeeping pilot programme, as well as systems level factors that potentially influenced the gatekeeping programme. For systems level factors, the framework and question guide were constructed based on the WHO categorisation of health system building blocks (World Health Organization, 2007), with questions focusing on the interactions between building blocks (i.e. service delivery, health workforce, health information, medical technologies, health financing, and leadership and governance).

For the second stage, the fieldwork was carried out in two phases (November 2014 and July 2015) during the pilot gatekeeping programme. For this qualitative study, semi-structured interviews with key stakeholders were conducted to identify the

effects and mechanisms of gatekeeping and its constraints. Interviewees were selected based on the intention to cover the perspectives of all key groups of stakeholders, and ensure sufficient number of interviewees within each group to represent variations of views and information. The following categories of stakeholder were interviewed: ambulatory patients with experiences of the gatekeeping policies, physicians and managerial staff from a district hospital and three township health centres (two pilot township health centres and a non-pilot typical township health centre), and administrators of the municipal and the district NCMS agencies and the district health bureau.

Table 8-2 presents the list of all interviewees and their main characteristics. Six patients in the pilot townships with experiences related to referral were selected among patients visiting primary care facilities. Ambulatory patients were asked randomly who had experience related to gatekeeping, i.e. patients who had requested to visit higher levels of care and were either referred or rejected, or patients who did not request but were referred by primary care practitioners on the practitioners' initiative. As nobody was rejected, those who had been referred were recruited. Unfortunately, there was no way to identify patients referred from the two townships at the district hospital, due to both the small numbers and the fact that the referral letter was not presented by patients to hospital staff—it was used only for claiming reimbursement back in their townships.

Eight doctors with experience of dealing with patient referral in the pilot district were interviewed. They included two from the district hospital who had handled the referral process, recruited through consultation with staff at the hospital's department of medical affairs. The other six were from primary care facilities. As there were barely any independent village doctors in the district, only one village doctor employed by one of the pilot township health centres was available for interview. Five facility managers were interviewed, including two from the two pilot township health centres and two from a comparison township health centre in the same district—the

comparison township health centre was considered typical township health centre in the city. Similar responses regarding issues not directly involving gatekeeping enhanced my confidence about the relevance of findings from the pilot-related facilities to other areas. Three health administrators, from the district health bureau, the district NCMS agency, and the municipal NCMS agency (which initiated the pilot programme), were also interviewed. Within each group (except for patients), one interviewee was interviewed in both 2014 and 2015 to check for policy and implementation consistency over time. In total, 20 people were interviewed (see Table 8-2). The interviews were recorded, then transcribed by a professional company, and checked by myself.

Table 8-2. List of interviewees and their characteristics

Interviewees	Serial numbers	Type of facilities /institutions	Gender	Age group	Year of interview
Patients	P01	Primary care	Male	51-60	2014
	P02	Primary care	Female	51-60	2014
	P03	Primary care	Male	61-70	2014
	P04	Primary care	Male	51-60	2015
	P05	Primary care	Female	51-60	2015
	P06	Primary care	Female	61-70	2015
Doctors	D01	Primary care	Female	41-50	2014
	D02	Primary care	Female	21-30	2014, 2015
	D03	Primary care	Female	41-50	2014
	D04	Primary care	Female	41-50	2014
	D05	Primary care	Female	41-50	2015
	D06	Primary care	Female	41-50	2015
	D07	Hospital	Male	41-50	2015
	D08	Hospital	Female	41-50	2015
Facility managers	M01	Primary care	Male	41-50	2014, 2015
	M02	Primary care	Male	51-60	2015
	M03	Primary care	Male	51-60	2015
	M04	Primary care	Male	51-60	2015
	M05	Hospital	Female	41-50	2015
Health administrators	A01	District health bureau	Male	51-60	2014
	A02	District NCMS agency	Male	41-50	2014, 2015
	A03	Municipal NCMS agency	Male	41-50	2015

Note: NCMS stands for New Rural Cooperative Medical Schemes

The third stage involved initial analysis of interview transcripts. This was done drawing from the “Framework” approach for data analysis developed by Ritchie and Spencer (2002), and using the qualitative data analysis software platform of NVivo 11 for Windows Version 11.0.0.317 (QSR International, 2013). The analysis was carried out in 3 steps. First, I familiarized myself with the range and diversity of the data by listening to all the recordings and transcriptions, and making notes. Second, a thematic framework was developed based on both the preliminary framework and field notes, and transferred into a structure of nodes in NVivo. Third, the transcriptions were coded according to these nodes, and the nodes were further developed and refined in repeated rounds.

The fourth stage involved interpretation of data and tabulation. I interpreted the coded data and identified factors related to the analysis, and associations between them. These factors were classified into causes and effects (either direct or indirect). For each set of causes and effects, key system variables were extracted and causal links were constructed correspondingly. A set of causal links included an upstream variable (cause) and at least one downstream variable (direct and indirect effects), as well as arrows denoting the direction and polarity of each individual causal link between the variables. The causes, effects, and the sets of causal links were then tabulated, along with the sources--the serial number of interviewees that provided the supporting evidence.

For the fifth and final stage, a causal loop diagram was constructed and used for analysis. The constructed links were transferred into a draft of a causal loop diagram. First, the causal links were connected with each other via identical variables to get feedback loops. Second, delay marks were added for links that were associated with delays. Third, each feedback loop was given a specific name reflecting the general theme that it described and added the signs representing the nature (reinforcing or balancing) of the feedback loops. Fourth, the serial numbers of causal loops that each link fitted into was added to the table created in Stage 4, so that the process was easily

traceable. The finalized form of the intermediate results is presented in Appendix Table 5.

The workflow of these steps was iterative. The interviewees responded to questions constructed from the categorisation of the health system building blocks. In other words, there was no prior set of system variables to code their responses. Neither was the table created in Stage 4 constrained by a refined structure. The resultant draft diagram had several places where the effects did not link back to a cause directly or indirectly or where the effects and causes were at different levels of detail. The constructed links were re-evaluated, and the causes and effects in these links were refined when necessary and reasonable, by revisiting some of the source coding. Neglected logical stages were added if needed, based on logical reasoning.

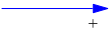
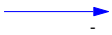





The diagram was drawn with Venism® PLE for Windows Version 6.4a (x32) (Ventana Systems Inc., 2016). Symbolic representation was used in accordance with Sterman (2000, pp. 138-139): an arrow denotes that there was a directional causal link between the two variables; positive and negative signs express the polarity of the causal links; “B” within a circular arrow represents that the loops were balancing, while “R” within a circular arrow represents that the loops were reinforcing. The only exception was that the colour green was used for links introduced by the gatekeeping pilot (the colour blue is used for other links), and dashed arrows for links that were intended by policy makers to happen, but which did not materialise in reality (see Table 8-3). The diagram was used to analyse potential lessons for policy making from the study.

8.3 Results

This section presents the CLD (including the whole diagram and each feedback loop) and explains the functioning of the gatekeeping pilot using the CLD (see Figure 8-1). It starts from the intended policy effects of the gatekeeping pilot programme (R1’

and B1'). Then the section analyses related factors and the feedback loops they formed, which challenged the feasibility of the two intended feedback loops. This leads to the three feedback loops (R1a, R1b and R1c) related to primary care facilities and three feedback loops that involve hospitals (B2, R2, and R3). Finally, the section explains the policy resistance facing gatekeeping itself (B1).

Table 8-3. Symbolic representation in the causal loop diagram

Symbol	Meaning
X  Y	a positive causal link (all else being equal, a change in the variable X leads to a change in the same direction in variable Y, compared to when X is held constant)
X  Y	a negative causal link (all else being equal, a change in the variable X leads to a change in the opposite direction in variable Y, compared to when X is held constant)
X  Y	a causal link with delay (change in Y take place after change in X after a delay)
X  Y	a causal link introduced with the gatekeeping pilot programme
X  Y	a causal link that was intended by policy makers but not achieved
	a balancing loop (sum of total numbers of negative causal links is odd)
	a reinforcing loop (sum of total numbers of negative causal links is even)

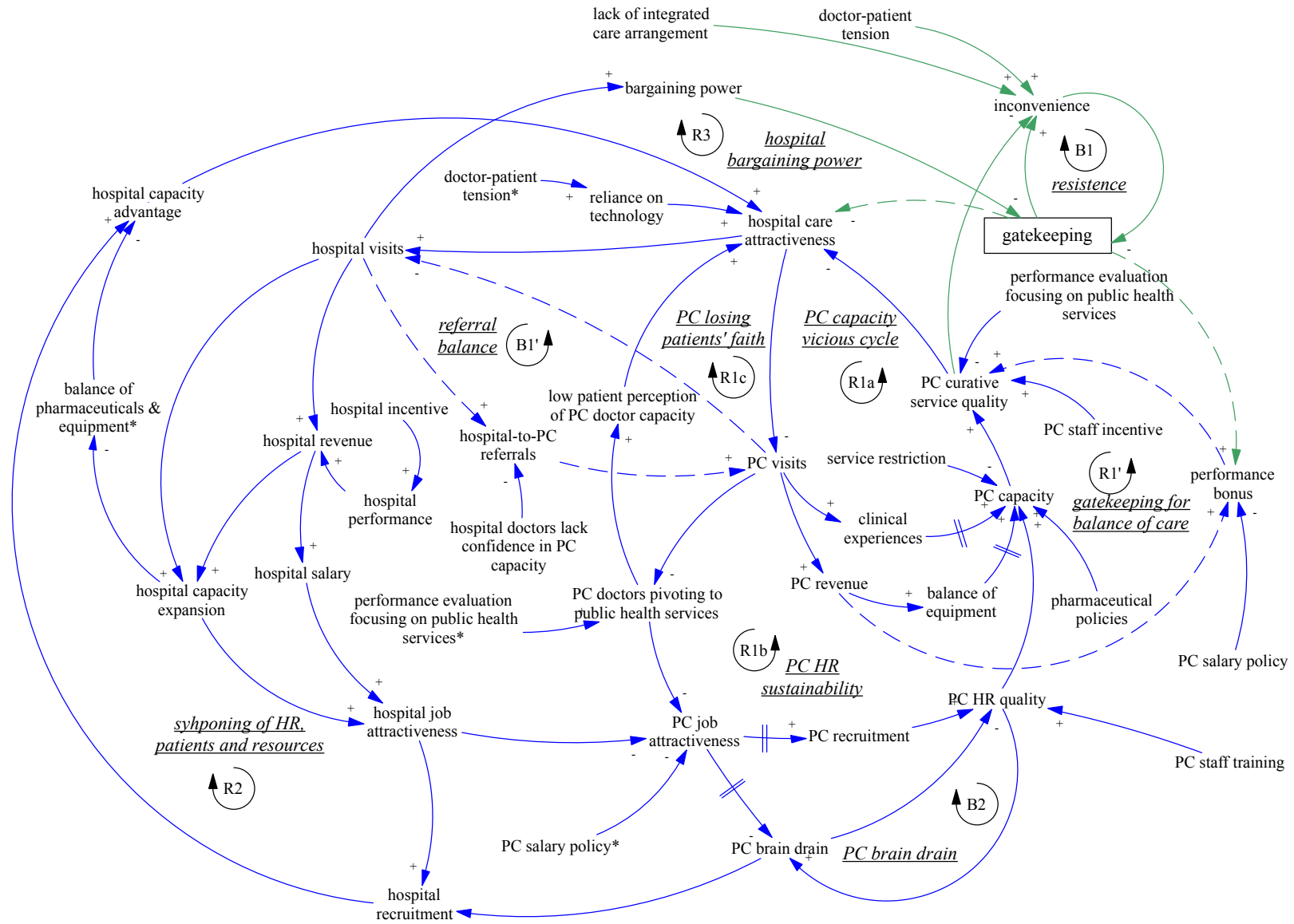


Figure 8-1. Causal loop diagram for gatekeeping pilot

8.3.1 Intended dynamics

The gatekeeping pilot programme was intended to influence the incentives of both providers and patients. From the demand side, the gatekeeping policy was expected to curb direct care seeking at hospitals without referral from lower levels of care, by preventing patients from claiming for their expenditures on such visits. Ideally, this would lead to reduced patient visits to hospitals, more patient visits to and more revenue of primary care facilities. From the supply side, the gatekeeping pilot programme installed a fundholding role in the township health centres. Potential savings and expenditures were supposed to be used as performance bonus for the facility and by facility managers to improve services. This was expected to reduce patients' need to go to hospitals (simplified as "hospital care attractiveness" in the diagram). In short, the gatekeeping policy was intended to establish a reinforcing feedback loop (R1', "gatekeeping for balance of care", made of blue arrows in Figure 8-2), and to a virtuous cycle towards improved functioning of primary care by reducing hospital care attractiveness and enhancing bonuses for primary care providers.

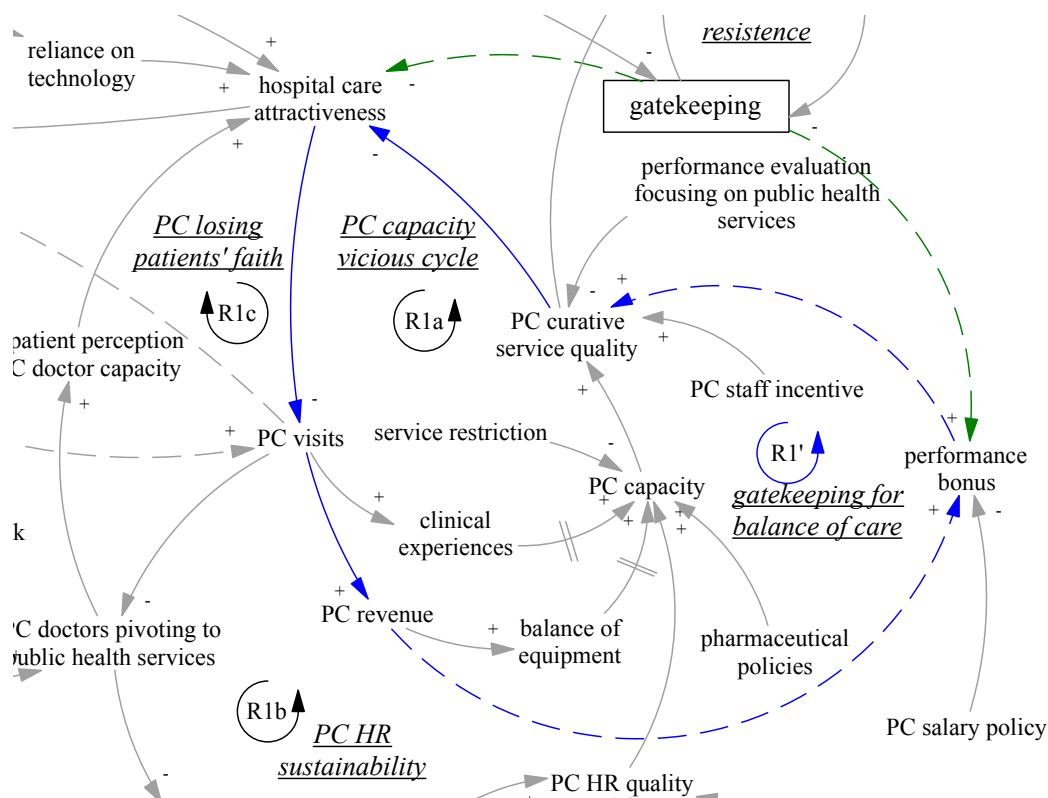


Figure 8-2. R1', an intended reinforcing feedback loop regarding the effects of gatekeeping for balance of care

Another intended feedback loop was on the referral interactions between hospital visits and primary care visits. While primary care visits were supposed to replace a significant but unknown proportion of hospital ambulatory visits, a portion of hospital patients were supposed to be referred back to primary care facilities particularly for follow up care. Therefore, there was an intended balancing feedback loop (B1', "referral balance") between the visits to hospitals and referrals back to the primary care facilities. If the feedback loop functioned as intended, it would contribute to the balance of patient visits between primary care facilities and hospitals.

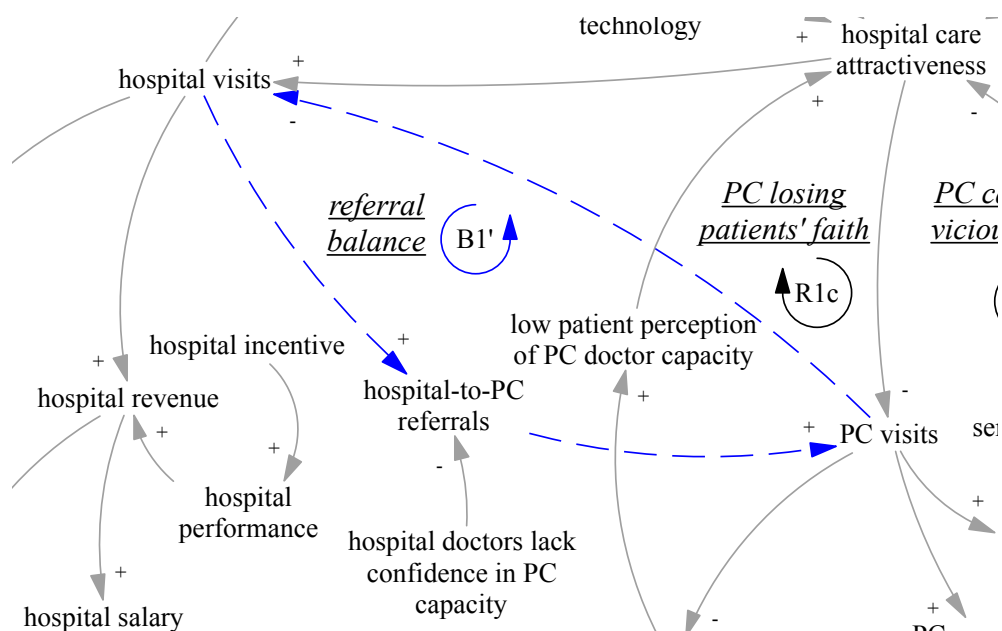


Figure 8-3. B1', an intended balancing feedback loop concerning referral balance

8.3.2 Low incentive of performance-based salary

The R1' loop was obstructed due to several factors mentioned by the interviewees. First, within the loop, the facility managers had limited ability to use the performance bonus to incentivize service improvement. This was mainly due to a nation-wide change in the salary policy (Ministry of Human Resources and Social Security, 2009; State Council, 2009). In order to maintain financial sustainability of primary care facilities and curb the profit-oriented over-prescription of drugs and services, a previous revenue-based salary system based on user fees had been replaced by a performance-based salary system based on a generally fixed total budget for salaries. While the policy change enhanced the financial sustainability of primary care facilities, it also minimized micro-economic incentives within the facility. The words of a director of a township health centre illustrated the ineffectiveness of using a performance bonus to stimulate doctors:

“The performance bonus comes from a [fixed] sum of salary... [the staff]

believe the performance bonus is a portion of earning that one deserves... He [or she] actually consider the money as part of his regular entitlement. If you deduct his bonus, he [or she] would be very unhappy... If one gets more, others will get less... For those who receives extra bonus, the amount cannot be too large. As a result, the stimulus is small... The intended aim of the performance bonus could barely be achieved.” (M02, interviewed in 2015)

The manager then recollected that before the reform in 2009, when the bonus for staff had been tied to their contribution to service revenue, the staff had been much more motivated. After the abolition of revenue-based bonus, he found it difficult to motivate his staff and had been relying on ineffective persuasion. The director of the other pilot township health centres used an increasingly objective performance index system to quantify the services and justify his decisions on salary distribution. The performance-related payment was kept lower than 300-400 *yuan* per month, which had a small effect on staff motivation. He purposely kept the bonus variations small as higher variations would create within-group tension and undermine the much needed teamwork spirit (M01, interviewed in 2014). In other words, the leveraging power of the performance bonus appeared minimal. The expected incentive of fundholding was therefore obstructed and less effective than desired.

8.3.3 Vicious cycle of primary care

The R1' loop was further weakened by a deterioration of primary care capacity, which appeared to be in a vicious cycle. A doctor at a pilot township health centre who previously had worked at a retail pharmacy selling drugs said:

“I feel there is little difference with a pharmacy. In a pharmacy, one also asks patients about their conditions and then dispenses medicines. Here things are basically the same... For some patients I would suggest blood test, [but we cannot provide that]... I can only give them some drugs that fit with the symptoms. We have all the examination devices but nobody to operate them.”

(D04, interviewed in 2015).

The township health centres were in the process of conversion to community health centres, therefore eliminating some main functions related to the mini-hospitals that they used to be. Inpatient care had been virtually eliminated, as had surgical operations. Reduced clinical experience and dysfunctional equipment of primary care facilities contributed to the decline of primary care capacity. The two township health centre managers complained that:

“The skills of all our doctors have deserted.” (M03, interviewed in 2015)

“We are not doing any [surgery] ... Our director Liu used to do all kinds of surgeries from head to foot.” (M04, interviewed in 2015)

The regulations unwittingly reinforced a process of breaking down the capacity and identity of what doctors meant at primary care level, and constructed a vicious cycle of primary care capacity. As primary care doctors were seen by both themselves and patients as losing their key competence, the patients who sensed serious illnesses just turned to hospitals without visiting township health centres. The vicious cycle was also reinforced by the unintended consequences of other policies that did not attend to the complexity involved in reforming primary care.

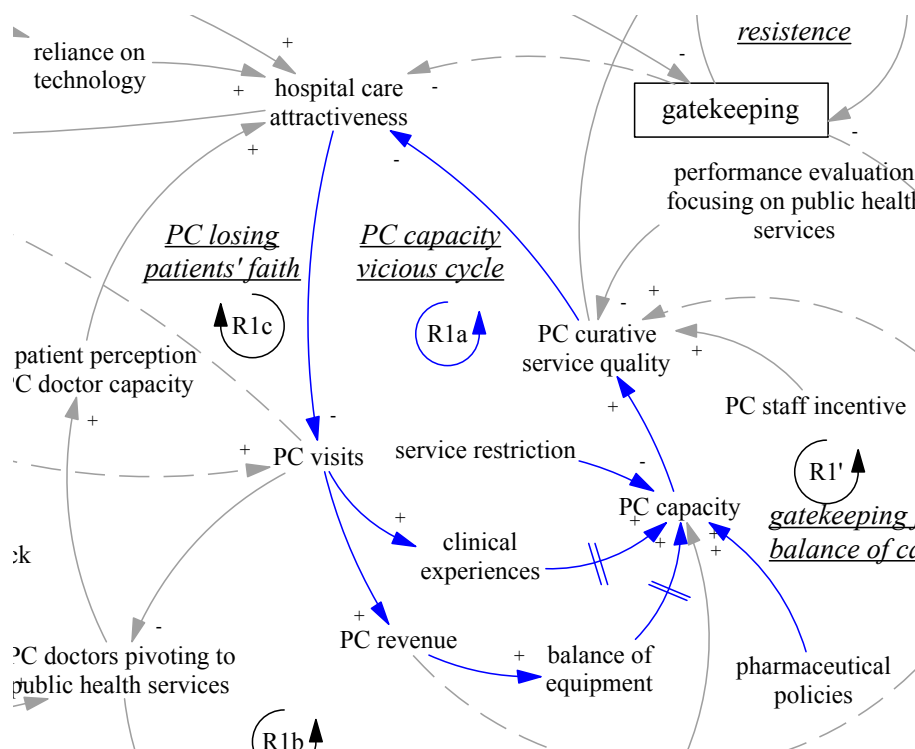


Figure 8-4. R1a, a reinforcing feedback loop concerning a vicious cycle of primary care capacity

The imbalance between hospitals and primary care facilities, in terms of use of medical technologies and pharmaceuticals, further limited the ability of primary care practitioners (particularly in comparison with hospital doctors). This fed back towards patients’ preference for hospital care. Particularly, the nation-wide essential pharmaceutical policy (Ministry of Health et al., 2009), while reducing purchasing prices for patients, restricted the pharmaceuticals that primary care doctors could prescribe. The restriction was made worse by the difficulty of transport in the countryside. A patient complained passionately about the restriction of access to pharmaceuticals:

“Now you have to go to large hospitals for serious illness. We also need to talk about the problem with pharmaceuticals... Even if [the primary care doctors] have the competence [of diagnosis and prescription], [they] cannot prescribe certain drugs... They are restricted by the level of facilities ... As the saying goes,

a capable housewife cannot cook a meal without rice [note: a Chinese saying meaning nobody can make something out of nothing].” (P03, interviewed in 2014)

The performance evaluation system was also driving the primary care facilities further away from providing curative care. In both the offices of the chief directors of the pilot township health centres, there hung a huge board of performance indicators where curative care took only about 1/5 of the space, with the rest devoted to performance management for the other subjects—mainly essential public health services. The pressure was intensified by the need to report to two agencies—a community health management centre (an arms-length agency under the district bureau of health) and a disease control department under the bureau of health, who had overlapping supervisory roles on the performance of primary care facilities on so-called “public health services” including managing health records, follow up of patients, etc. (M03 and M04, interviewed in 2015). Based on the interviews, there seemed to be variations on how much was achieved in reality, but the shifting focus of work from clinical curative care to disease management was clear. A young doctor in a pilot township health centre, who had trained in clinical medicine, spent most of her time in a department of community health, and was doing little clinical work because of the intense pressure of performance evaluation, and her youth (and hence lack of trust by patients, and low patient volume). She said:

“I certainly regretted, because I studied (clinical) medicine and was prepared for that. But since it was the requirement of work, I had no choice.” (D02, interviewed in 2014)

This pivot was also reinforced by the reduced patient visits and contributed to a downstream reduction of the attractiveness of jobs as a primary care doctor. None of the interviewed doctors who were spending their time in disease management were happy about the situation. This was also an issue of sustainability of human resources

at primary care facilities (R1b, “PC HR sustainability”). Recruitment in the facilities mainly targeted medical graduates with a four-year associate degree (a full medical degree would require minimally 5 years of training). Even such recruitment was difficult according to a district health administrator (M03, interviewed in 2015).

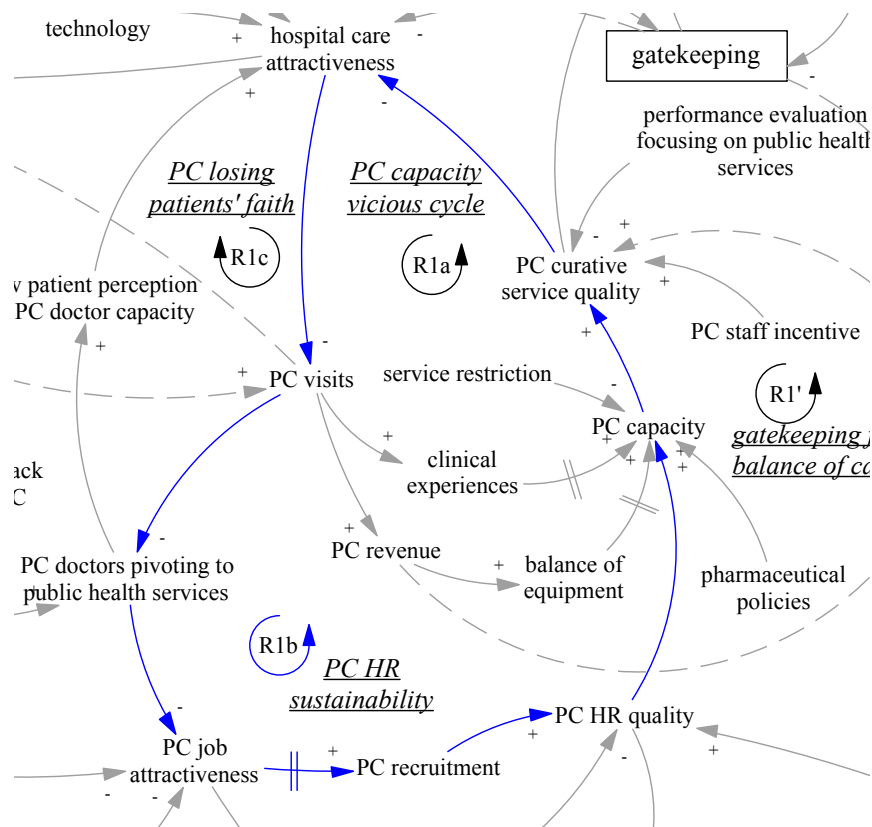


Figure 8-5. R1b, a reinforcing feedback loop concerning the sustainability of primary care human resources

Related to this was another reinforcing loop that further challenged the intention to move away from the situation of patients bypassing primary care facilities. As the focus of primary care doctors shifted towards public health services, the patients noticed that their service function was reduced. Low patient perception of the capacity of primary care doctors fed back towards hospital care attractiveness (R1c, “PC losing patients’ faith”). There appeared to be a breakdown of doctors’ professional status in the township health centres, not only from the perspective of the disillusioned doctors, but also of the nostalgic patients who said that in the past the township health centres

could deal with all kinds of cases including some major surgery:

“they could cut out [diseased] lobes of lung, sections of breast”. (P03 interviewed in 2014)

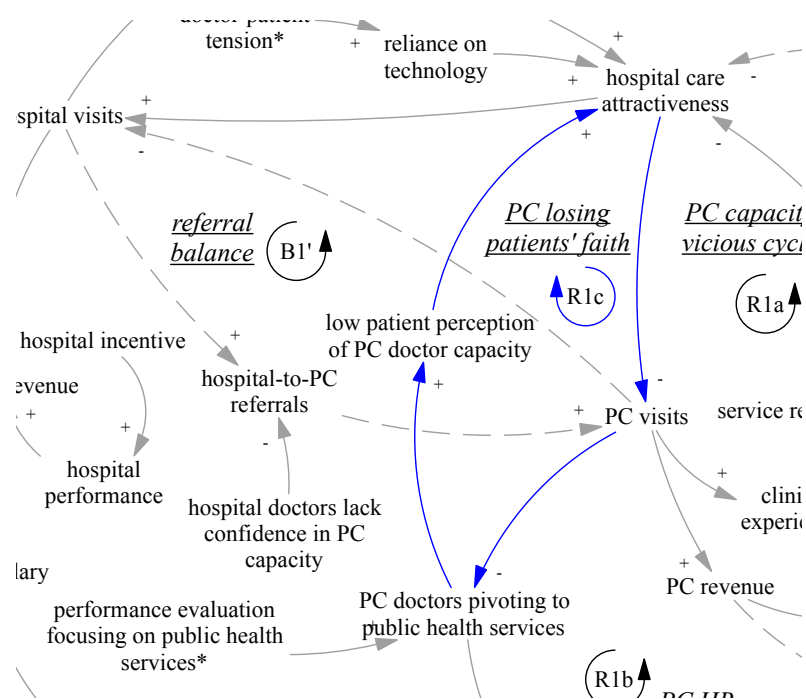


Figure 8-6. R1c, a reinforcing feedback loop concerning primary care losing patients' faith

When considered together, R1a, R1b and R1c formed a very strong tendency towards further declining functions and capacity in primary care facilities, and erosion of the professional status of doctors. The feedback loop B1' regarding post-hospitalization referral to primary care was also not achieved, as the hospital doctors lacked confidence in the capacity of primary care facilities, rendering the feedback loop illusory. The hospital manager interviewed said:

“now almost every young person goes to college and gets a full degree, how could those who could not get enrolled in a full degree university programme be trusted to treat people's illness?” (M05, interviewed in 2015)

A doctor from the district hospital expressed a similar doubt:

“I don’t think they can solve any real problem. Those with real problems would be referred to tertiary hospitals... Regarding back referral [from hospitals to township health centres], to be frank, we operate according to the demand of patients... If the patients believe it is inappropriate, we have to give up... There are very few back referrals [in practice].” (D07, interviewed in 2015)

8.3.4 Challenges from hospitals

Besides the challenges within the primary care sector, difficulties for gatekeeping also came from the interface with hospitals. In order to improve the level of skills, there was a training system in which newly recruited medical graduates at primary care level underwent further training at hospitals. According to a hospital doctor and an officer of the district health bureau, many seemed ready to give up their position in primary care if they were offered a position in the hospitals. Therefore, there was a balancing loop of human resources at primary care level (B2, “PC brain drain”), in which those with better training and career prospects would leave primary care for hospitals. The lack of enforcement of a university medical education programme to train rural students as doctors targeted for specific rural primary care facilities (so-called “order-based training programme”, or medical graduates admitted on such special programmes were required to work in primary care facilities in rural areas) was common, which facilitated the brain drain (A02, interviewed in 2015).

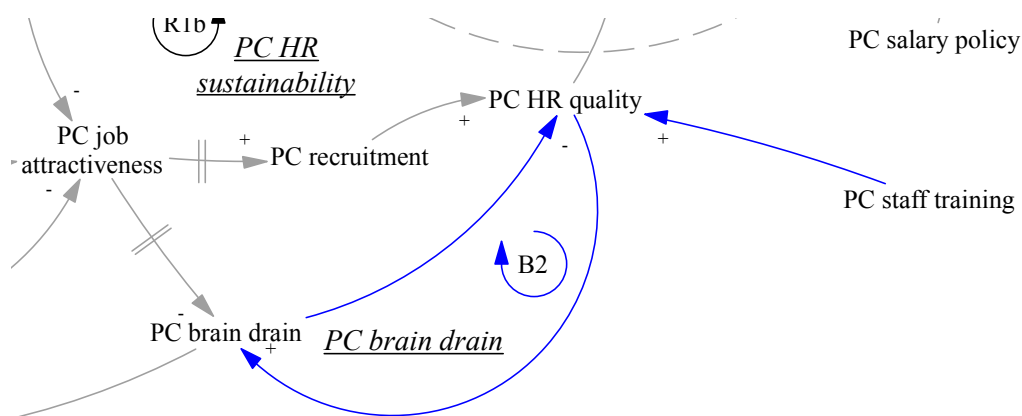


Figure 8-7. B2, a balancing loop concerning brain drain from primary care

The self-reinforcing nature of the balance between primary care facilities and hospitals becomes particularly clear, when the feedback loop R2 (“syphoning of HR, patients and resources”) is examined. Discussion with the district hospital manager (M05, interviewed in 2015) revealed that the brain drain of primary care was mainly limited by the already depleted reservoir of capable primary care doctors. In fact, the hospitals were actively recruiting graduates with not only a university medical degree but also masters graduates (three extra years of medical training). The result perhaps was not just reduction in recruitment at primary care level but also a deterioration of quality and a further divergence of professional status and aspiration. As the same manager in the district hospital argued, some who preferred to stay at primary care facilities tended to have low career aspirations. They were happy working at primary care facilities, because there was lighter workload and steadier income compared to working in hospitals. Related to this was the increasing hospital visits associated with the hospital incentive structure linked with revenue generation. The hospitals (mainly the district hospital) attracted a lion’s share of revenue (A02, interviewed in 2015). The hospital used it to build up its advantage in equipment and infrastructure. In short, the comprehensive structural advantage of the hospital fed back to its functional advantage in that it attracted ever more patients.

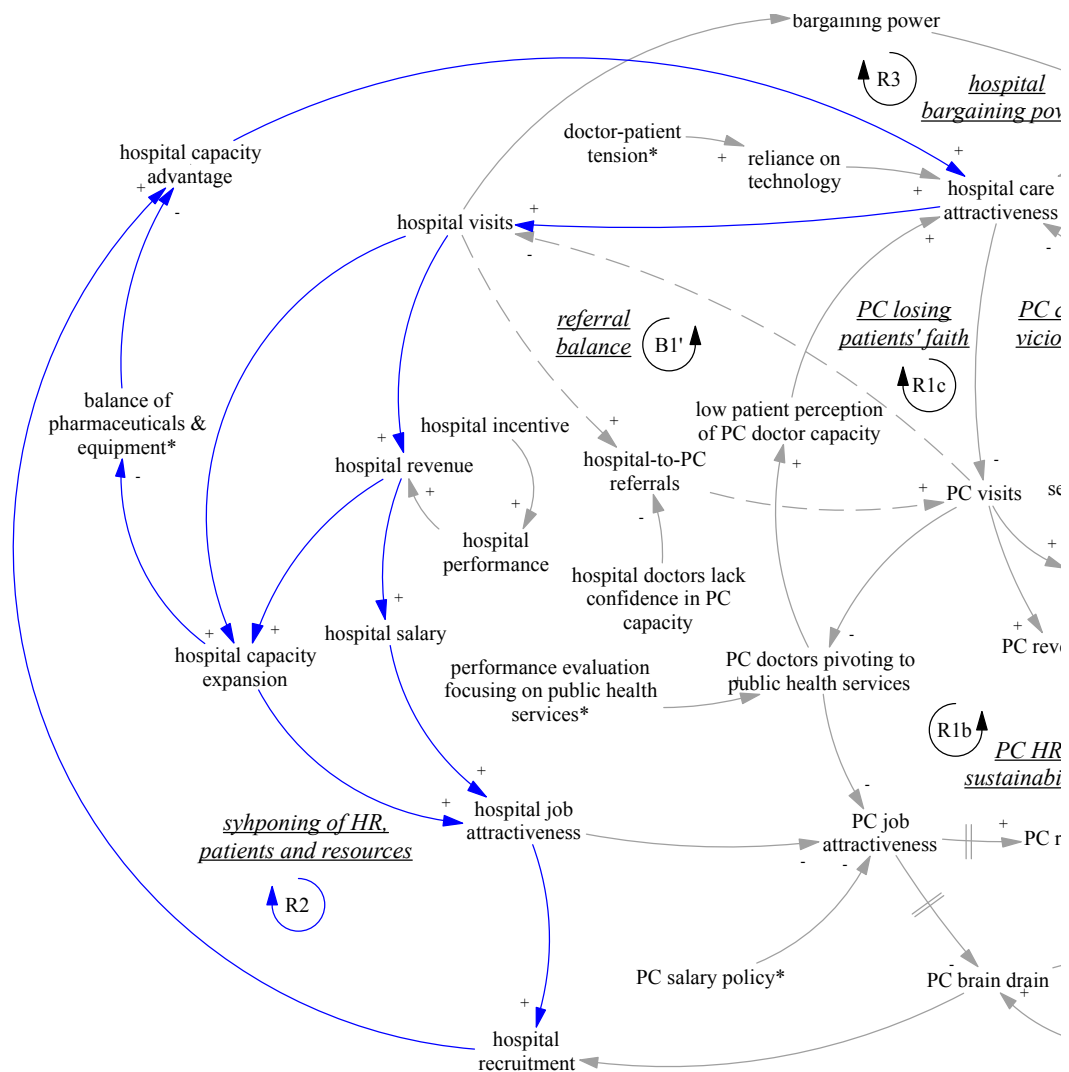


Figure 8-8. R2, a reinforcing feedback loop, concerning syhponing of HR, patients and resources

The advantage of hospitals also fed back to the policy making process. The large patient volumes in hospitals provided them with strong bargaining power and reduced the prospects of strict gatekeeping policies (R3, “hospital bargaining power”), particularly as local government was required to provide care for most patients within the geographical range of the district. In other words, the opposition from the interests related to hospitals were challenging the sustainability of the gatekeeping pilot in its current design. Indeed, the municipal NCMS administrator was considering replacing the pilot programme by moving the capitation-based ambulatory care budget to the district hospital, as it was believed to be more capable of functioning as gatekeepers

(A03, interviewed in 2015).

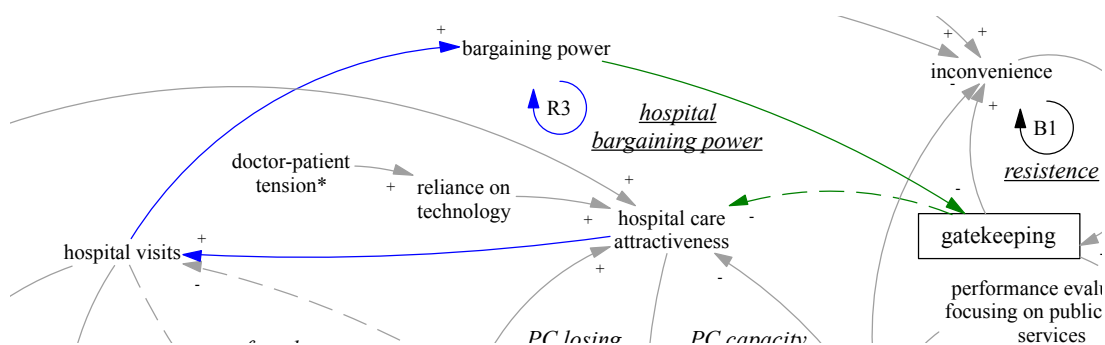


Figure 8-9. R3, a reinforcing feedback loop concerning hospital bargaining power

8.3.5 Gatekeeping backfires

Finally, the gatekeeping policy backfired due to weak primary care capacity (B1, “resistance”). Patients found primary care facilities to be very restrictive in services, technologies, and pharmaceuticals, and felt they received little extra benefit when they came to visit primary care facilities for referral. The extra visits became a burden to primary care facilities, too.

“The patients came to us to be referred to the district hospital, [and then] to the district hospital to be referred to municipal hospitals. Will you say that is not troublesome for them? It is understandable that patient complained... They are not willing to come here to get referral. [They would say] I see doctors elsewhere but [why do] I need you to give me a certificate.” (D06, interviewed 2015)

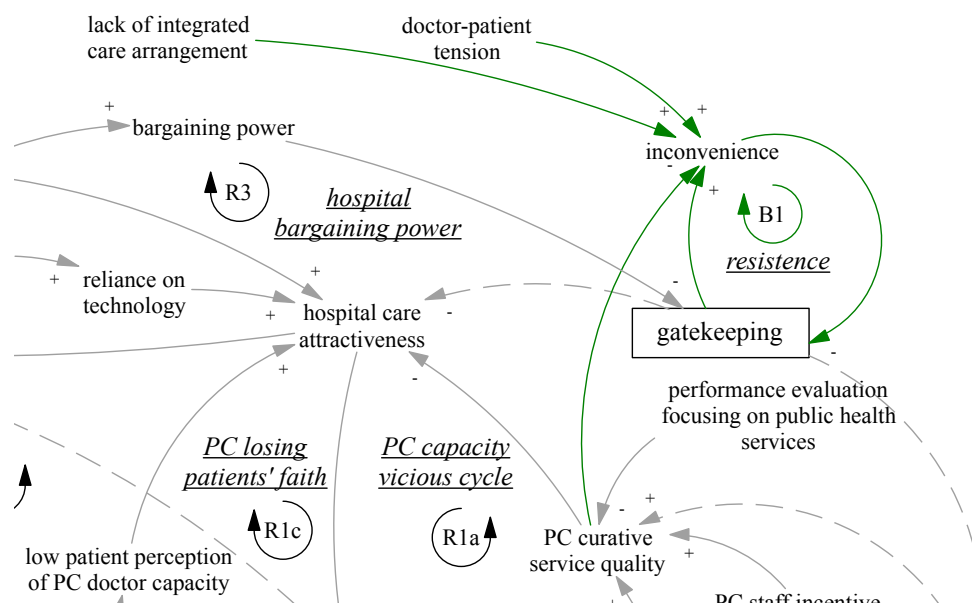


Figure 8-10. B1, a balancing feedback loop concerning resistance to the gatekeeping policy

Most doctors and patients considered the policy an inconvenience, though some also acknowledged that the policy brought additional opportunity to make contact with patients. The tension was also increased by the lack of patient awareness, despite government's effort to publicize the policy change. In several cases, patients went to hospitals first and when they wanted reimbursement they found that they had to get a referral from primary care doctors. Pressured by patients (with whom doctors had a potentially tense relationship) and limited by the capacity to provide clinical services that could replace patients' care seeking at hospitals, primary care practitioners usually just wrote referral letters for patients (P01, interviewed in 2014). Furthermore, there was little arrangement (e.g. priority access compared to self-referred patients) to make it easier for patients to see doctors in hospitals. Even if they got a referral from primary care facilities, patients still went through the same process and joined the same queue as other patients (D02, interviewed in 2015). The referral requirement thus became largely ritualistic and indeed an extra trouble for both doctors and patients, which added to their resentment of the policy. In particular, gatekeeping hurt local elites who had more say in the political process (e.g. people's representatives), and these people

pressured local leaders to abolish strict gatekeeping policies (A02, interviewed in 2015). With the weak capacity of primary care and other issues of the policy's implementation, the value of gatekeeping became mostly an inconvenience that turned into resistance against the pilot policy.

8.4 Discussion

8.4.1 Limitations and value of the approach

One limitation is that the sub-study did not let interviewees or independent experts validate the model, which has been recommended (Carey et al., 2015). After a failed attempt to explain an earlier draft of the CLD to some municipal policy makers, it was considered difficult to use the CLD as a communication tool to policy makers who had little prior training, and to explore this further was beyond the capacity of the study. The findings should therefore be seen as the understanding of the researcher, generated through a rigorous process.

The approach used in this sub-study seems to have advantages in understanding the complexity involved in shifting balance of care through interventions like gatekeeping. The use of the WHO categorisation of health systems building blocks facilitated a systematic mapping of factors related to gatekeeping. In the sub-study, applying the categorisation facilitated the identification of issues directly related to the mechanisms of gatekeeping such as financing (e.g. the ineffective performance-based bonus), but also less directly related to gatekeeping such as pharmaceutical policies and technologies (e.g. restriction of access to medicine).

Using the CLD has allowed the sub-study to bring together the separate analyses to understand the interrelationships between different factors within and across categories of building blocks. One particular advantage is related to dealing with unintended consequences of policies indirectly related to gatekeeping (e.g. the restriction and change regarding the service functions of primary care practitioners

contributed to a deterioration of service capacity of primary care facilities). The CLD also allowed the sub-study to identify both local patterns of feedback loops and how these feedback loops formed a holistic picture of all the key factors related to gatekeeping.

The approach bridged analysis of the gatekeeping pilot with analysis of the system within which the gatekeeping pilot was embedded. The approach thus brought into the qualitative evaluation of gatekeeping the three dimensions of interrelationships, perspectives and boundaries, highlighted in the systems literature (Williams, 2015). It revealed the richness of interrelationships among different factors within the health system that were directly or indirectly related to gatekeeping functioning, reflected the multiple perspectives of different groups of stakeholders, and encouraged a deeper understanding of the boundaries by highlighting the linkages between the intervention and the system, as well as by examining unintended consequences of the gatekeeping pilot.

Furthermore, the approach of qualitative systems analysis developed in this sub-study was explicit and transparent. A systematic review of the recent use of system science and systems thinking for public health suggested that studies using systems modelling methods should make the formulation of models (in this case a CLD) explicit enough for readers to judge the rigour of the studies or to repeat the process (Carey et al., 2015). The complicated process and lack of transparency of interim stages made causal loop analysis prone to issues regarding accountability. The danger of misunderstanding the system based on a model with suboptimal rigour is amplified by the assumed interconnectedness of the factors. However, there has been little guidance on how to develop CLD based on rigorous qualitative methods and data. This sub-study has established an example of a transparent and rigorous approach to qualitative systems analysis of a complex health systems intervention.

8.4.2 Findings regarding gatekeeping and implications beyond

The chapter has presented the first evidence on the intended and actual functioning of gatekeeping in a pilot in rural China. The sub-section here discusses the findings regarding both the intended and the actual functioning of the gatekeeping pilot; factors that likely explained the gap between the intended and actual functioning, both due to historical legacies and the consequences of recent other policies.

Within the study context, the intended mechanisms of gatekeeping in changing patients' utilisation pattern of care were only partially realised. The intended supply-side incentive on treating a greater number of patients at local facilities did not seem have functioned as expected, as the salary policy was too rigid with a level of pay too low to either attract or incentivise gatekeeping-related clinical work. On the demand side, a large number of patients appeared to be going through primary care reluctantly to get referral in a generally ritualistic process. This helps explain the findings from the previous chapter, which found that patient visits to primary care facilities increased by a large proportion but spending increased much less. The implementation of the approach of gatekeeping in the studied pilot led to dissatisfaction of both doctors and patients. This contradicts a patient survey done in Shenzhen (Gan et al., 2016) that showed willingness of local residents to accept community health centres as gatekeepers.

Besides public resentment, potential adverse effect is delay of diagnosis or misdiagnosis (Crawford, 2014). The sub-study did not investigate this issue directly, but the weak primary care capacity suggests that this would be hard to avoid, if a significant number of patients rely on the current primary care providers. Furthermore, given the different capacity of primary care facilities and hospitals, implementing gatekeeping only for the NCMS could potentially exacerbate inequity by restricting their access to facilities of lower service quality.

The sub-study identified three aspects that led to the sub-optimal functioning of gatekeeping. First, the weak conditions of primary care, particularly regarding the skills of primary care doctors in comparison with those in hospitals, seemed to be a fundamental barrier facing the reform. The nation-wide gaps in qualifications and income between primary care doctors and hospital doctors were sustained over the recent decade when social health insurance coverage was extended to the whole population (see Chapter 4). Therefore, it was understandable that patients in the pilot townships were not satisfied when their eligibility for direct access to ambulatory services in hospitals were taken from them.

Second, the study has further revealed reinforcing feedbacks that turned into a series of vicious cycles for primary care development, in terms of the weakened service capacity of primary care, the decreasing patients' trust of primary care and questionable sustainability of human resources for primary care. The study has shown the danger of neglecting the professional aspiration of primary care practitioners and patients' appreciation of their competence, which seems still to hinge on the ability of primary care practitioners to provide curative care.

The lack of progress in reforming hospitals exacerbated the imbalance between the two sectors. Despite reform in primary care, the inflationary incentive structure in hospital care remained unchanged. Hospitals were systematically absorbing human resources, patients, and other resources, contributing to greater imbalance in the system. Hospitals (particularly the district hospital in the pilot area) have become increasingly the main provider of curative care and received most of the total medical expenditures. This is corroborated by findings of Chapter 4. The self-reinforcing nature of the imbalance between hospitals and primary care facilities could mean increasing difficulty in future reforms.

Third, the effectiveness of gatekeeping was hampered by the unintended

consequences related to conflicts among different priorities required of primary care development. Primary care facilities have been loaded with much aspiration for the ultimate goal of universal health coverage in low- and middle-income countries. There coexisted multiple policy initiatives in the pilot as well as China-wide: strengthening the function of primary care facilities in curative primary care, strengthening the function of primary care facilities in preventive primary care for the increasingly prevalent noncommunicable diseases, curbing over-prescription related to the previous incentive structure, and reducing pharmaceutical prices. These intersecting reforms provided plenty of scope for clashes and inconsistencies. The findings suggested challenges in changing the functions of primary care facilities, as primary care facilities have relied for years on mechanisms similar to those in the hospital sector (revenue generation, recognition of professional status focused on treating diseases, etc.).

Technological regulations, some of which aimed at standardizing primary care facilities and improving the alignment of their service with a primary care orientation, appeared to undermine the basis of trust in primary care providers' technical capacity. The effort to strengthen chronic disease prevention (e.g. focus on performance indicators of "public health services" including follow up care of chronic patients) was important as a corrective action to the previous focus on curative care. However, it might undermine efforts to provide more and better curative care at primary care facilities, and even break down the appreciation of professional status and competence of primary care practitioners by both patients and colleagues.

In relation to this, the performance-based salary policy reform and a virtually fixed budget payment system, by eliminating the previous incentive to over-prescribe, seemed to have affected the facility manager's entrepreneurship and ability to motivate staff. The essential drug policies, which seemed to have unintendedly led to limited access to pharmaceuticals at primary care facilities, also restricted the range of services

available at this level. Previous studies have suggested these were common challenges facing primary care facilities in China (Zhou et al., 2014), though this sub-study further elucidated the underlying dynamics.

8.5 Conclusion

This sub-study has used a qualitative systems analysis to understand how gatekeeping functioned under constraints in a pilot in rural China, and the factors that affected its functioning. The approach combined the WHO categorisation of health system building blocks, a causal loop analysis, and the “Framework” approach for qualitative data analysis. Using this approach, the sub-study developed a CLD based on information collected from semi-structured interviews and used the CLD to facilitate further analysis of the functioning of the gatekeeping pilot in the context of its system. The qualitative systems analysis found that the pilot policy led to a substantial proportion of patients visiting primary care simply to get a referral, which resulted in dissatisfaction of both doctors and patients.

The study has suggested that the gatekeeping pilot failed to alter the dynamics involved in an increasingly imbalanced local health system. The current salary policy was too rigid with a level of pay too low to either attract or incentivize gatekeeping related clinical work. If scaled up and strictly adopted in settings with weak primary care, gatekeeping of the kind implemented in the pilot could lead to other undesirable outcomes. These might include public resentment and other unintended consequences in equity and quality of care (e.g. delayed diagnosis), which could undermine the momentum for shifting the balance from hospitals to primary care providers.

The study has suggested a number of underlying systems factors that restricted the functioning of gatekeeping in the pilot area. The weakness of primary care capacity (particularly in terms of human resources) lay at the heart of ineffective gatekeeping. Primary care facilities were also trapped in vicious cycles. Particularly dangerous was

the phenomenon that the primary care doctors were losing patient trust and professional aspirations. Unintended consequences of a number of concurrent policies also impeded strengthening of primary care functioning. Strict regulation on pharmaceuticals and the technological imbalance between primary care and hospitals limited the medicines and technologies available to primary care facilities. The delayed reform of perverse hospital incentives also contributed to the barriers to successful functioning of gatekeeping.

Part IV Discussion and conclusions

Chapter 9 Discussion

This chapter addresses the following topics: the methodological issues regarding synthesizing findings from all four sub-studies; the overall findings when all sub-studies are considered together; relevance and generalizability of findings; and methodological reflections.

9.1 Methodological issues regarding synthesizing findings from sub-studies

The journey of the thesis started from treating gatekeeping as an intervention to achieve stronger functioning of primary care in China and expanded to a complex health system issue—the balance between hospitals and primary care providers. Eventually, the thesis became a multi-methods mix of sub-studies. Overall, the research design encompassed two historical sub-studies: a descriptive study to assess the structural and functional balance between hospitals and primary care providers (between 1949 and 2013) and a historical-institutionalist analysis of the coevolution of hospitals and primary care providers (between 1835 and 2013), and two contemporary sub-studies: an impact evaluation of a gatekeeping pilot since 2013, as well as a qualitative study of functioning in configuration with other elements in a dynamic system.

The task of this section is to discuss some key methodological issues about synthesizing the findings from the four sub-studies, covering the following points: First, the section discusses the overall conceptualization of balance between hospitals and primary care providers, which was the conceptual basis for all four sub-studies. Second, the section discusses the synthesis of the historical sub-studies. Third, the section discusses the synthesis of the contemporary sub-studies. Fourth, the section discusses the thesis's combined use of historical analyses with contemporary

evaluation.

9.1.1 Conceptualization of balance between hospitals and primary care providers

This framework contains hospitals and primary care providers as two sectors within a single dualist health system structure and provides the conceptual basis for all four sub-studies and their synthesis; how the framework was used in the thesis therefore deserves attention.

A central cross-cutting issue was the meaning of and distinction between primary care providers and hospitals. As reviewed in Chapter 2, there are multiple dimensions to define primary care (and thereby primary care providers): based on a range of clinical professions or disciplines, a type of setting or a system of service delivery, a set of services provided by primary care providers (e.g. general practices), or services with a set of normative attributes at micro or macro level. It has also been clarified in Chapter 4 that this thesis used the government policies' operational definition of "*jiceng weisheng jigou*" (literally basic level health facilities), which included township health centres and village clinics in rural areas, and community health centres and stations in urban areas. This indeed has been the general approach commonly used in studies about China, though not always. For example, in 2015 alone more than one thousand Chinese research papers were published on "*jiceng weisheng jigou*". On the other hand, in trying the Chinese version of the Primary Care Assessment Tool, Shi included as primary care providers doctors at tertiary hospitals who were used by patients as regular first point of contact (Yang, Shi, et al., 2013).

The underlying rationale behind previous mainstream studies on primary care providers in China appeared to be that the perceived weakness of primary care providers constituted a serious policy issue. This has been conceptualized in Chapter 4 as structural and functional balance between primary care providers and hospitals. In Chapters 5 and 6, the institutional dimension was further included in the analysis.

The analyses therefore linked description of the situation with both major policy issues and the potential policy targets at the operational level. Had the thesis focused on “primary care provided by hospitals” then increasing the primary care attributes of hospital services would likely lead to further weakening of the functioning of primary care providers.

The thesis also looked at the combination of the two sectors within a single dualist structure. As mentioned earlier, this combination brought analysis of hospitals and primary care providers under a single framework. It facilitates the discussion of the interrelationship between the systems factors that relate to both sectors. The framework highlighted the fact that the two sectors share related sources of inputs (human, physical and political resources) and serve related functions. Their division, professionally, politically, and organizationally and financially, mattered greatly in affecting the structural and functional balance between them. Such issues would be difficult to capture in analysis only on hospitals or primary care providers.

To sum up, the conceptualization of hospitals and primary care providers as two sectors within a single structure, and their operationalization with a consideration of contemporary impactful policy targets, constituted the common conceptual ground for all four sub-studies and the synthesis of their findings.

9.1.2 Synthesis within the historical part

The historical trend analysis described the trend of structural and functional balance between hospitals and primary care providers. The second historical sub-study analysed the coevolution of hospitals and primary care providers, focusing on the four sets of institutions (professional, organizational, financial, and governance). The historical-institutional analysis provided an explanation of the findings of the quantitative historical trend analysis of structural and functional balance between hospitals and primary providers. Synthesizing findings from the historical trend

analysis and historical-institutionalist analysis therefore focused on generating a historical narrative, which then acted as a basis for generating policy implications.

9.1.3 Synthesis within the contemporary part

The qualitative systems analysis provided an approach to see gatekeeping functioning under constraints of the pilot. Meanwhile, it also provided a “whole-system” view of the factors that constituted the constraints. By linking the evaluation of the pilot with evaluation of the system containing the intervention of the pilot, the analytical approach recognized that the system dynamics were working with and without the intervention. The quantitative impact evaluation could be seen as an outcome evaluation. The contemporary qualitative systems analysis provided not only an explanation for the results of the quantitative impact evaluation, but also a deeper understanding of the functioning of the contemporary system relatively unaffected by the gatekeeping intervention.

9.1.4 Synthesis across the two parts

1) Rationale of synthesis

Purposes of historical studies and contemporary evaluation are usually divergent. While historical studies rarely tries to prescribe solutions, programme evaluation focuses on relatively specific factors that can be altered to generate specific impacts. The discussion considers a way to synthesize and the potential benefits of combining the two parts in a single framework.

Occasionally, studies have suggested potential areas of convergence. Evaluation has generally come to recognize the existence of complexity, one of which is the temporal dimension. Policy analysts have also been drawing results from evaluation to build arguments related to path dependence. “Every programme is a complex system inserted into a complex system” (Pawson, 2013, p. 82), where “[c]omplexity comes

because about because inputs, stakeholders, institutional features and outputs are all in a state of constant change, with or without the help of the programme” (Pawson, 2013, p. 49). Following this logic, framing the issue of balance between hospitals and primary care providers in a health system as a historical system provided a basis for synthesizing these two strands of investigation. The contemporary systems analysis provided a cross-sectional snapshot of the systems dynamics where the pilot unfolded. The synthesis focuses on the benefit of historical analyses.

2) Added value of history

- Function 1: Broadening the scope of investigation

History offers insights that transcend the temporal barrier of the contemporary world “historical amnesia”. It provides a temporally rich understanding of the “causes of causes” and particularly traces the contemporary phenomena to historical events/processes and sequences. Rather than seeking answers from contemporary associations between potential policy targets and outcomes, a historical study looks for the consequences (or policy feedback in historical-institutional terms) of past events and their sequences. As Berridge (2011) suggested, a key function of history in public health is to make people rethink.

While they highlight historical factors and sequences not malleable in the current period, historical studies also highlight the contemporary dynamics set in motion historically. Some such dynamics may replicate themselves, while others may face impediments that lead to breakpoints. From the perspective of “a pure historical study”, it is appropriate to just trace backwards towards identification of historical factors and dynamics that cannot be influenced, identifying the historical roots/causes of contemporary policy issues. However, from the perspective of contemporary policy interests, this is not adequate. A historical academic effort, by facilitating reflection and demonstrating the plasticity of some current factors, should open up new

opportunities or new insights that will adjust our understanding by removing preconceptions. From a policy perspective, the interpretation of history should ultimately be geared towards helping contemporary actors develop better interventions to either change the course of history or to ride the tide of history.

Re-identifying causes prompts reconsideration of solutions. Historical sub-studies therefore help propose new theories on the key drivers of change, and these new theories may lead to new interventions that can potentially be tested in evaluation. By developing theories that explain the coevolution of hospitals and primary care providers, the historical sub-studies (particularly the historical-institutionalist analysis) could highlight the mechanisms that institutions become path-dependent (coordination difficulties, asset specificity, reinforcing and reactive feedback). Evaluation within conventional health service research has been criticized as being generally technocratic with little consideration of power (Mills, 2012). Developments in programme evaluation have also been urged by system thinking experts to go outside the boundary of contemporary intervention and particularly to engage consideration of power in developing evaluation projects (Williams, 2015). The historical-institutionalist analysis in particular highlights the critical role of power and politics.

From the perspective of history, evaluation can only capture piecemeal and limited processes. Major changes are more likely processes of social transition that may well be beyond the conceptual framework of a project of evaluation. What comes out of the historical analyses might not fit in the framework of a programme intervention. This is discussed in Section 9.3. Here it suffices to say that what historical research can add is to bring in a critical perspective on the role of evaluation and indeed policy interventions in terms of impact.

- **Function 2: Building a systematized narrative**

Historical research can “systematize” the various aspects of interest into a single

historical narrative. This is different from the whole-system approach that claims to address all relevant factors within a system and the interrelationship among these factors. Systems analysis has been criticized for sacrificing depth for breadth. The contemporary evaluation has identified a range of systems level issues that went beyond the mechanism pathway of gatekeeping. They also seem to reflect numerous challenges that are difficult to handle. The findings of historical analyses could be used to reinterpret findings from contemporary evaluation in order to suggest how to introduce systematic change. They also provided a temporal frame in developing reform strategies. In a sense, this combination can better address the inter-linkages of a complex reform that would be difficult to capture in separate programme intervention evaluations. The historical studies provide a historical narrative that helps bind contemporary issues together as the joint effects of institutional evolution.

The historical-institutionalist analysis has also captured the interplay between politics and systems/policy, by focusing on feedback effects where policy changed the rules of politics by preconditioning the preferences of actors and the coalitions they formed. The historical-institutionalist analysis, by focusing on the feedback effects of policies made in earlier periods, also highlighted the interrelationship between policies and politics. In addition, the historical-institutionalist analysis (reinforced by the quantitative trend analysis) distinguished the long-term “stable” period and short-term “dynamic” period. Findings from historical analyses facilitated a judgement of the temporal context of policies and therefore helped optimize policy suggestions. By bringing the numerous contemporary factors together, bringing the policies and politics together, and bringing evaluation of a policy implemented at a particular time point into a continuous historical process, the thesis allowed an integrated and yet focused analysis of key policy issues, while capitalizing on the comparative advantages of multiple research approaches.

3) Barrier to synthesising

A key barrier to synthesising findings from the historical and contemporary parts lay in the different levels of data representativeness. The historical and contemporary studies dealt with different geographical ranges. The historical study utilized data about the entire country, while the study of the pilot was restricted to the local district. As mentioned before, the pilot was chosen as a site for this case study due to its pioneer nature. Despite the obvious limited scope of the contemporary pilot, none of the key contemporary factors in the system seemed to be confined to the local pilot. A certain degree of theoretical generalizability can be drawn for similar settings in China, though there is much room for strengthening with further evidence. Findings from the historical studies act as a triangulation and generally support the contemporary evaluation, as will be clear in Section 9.2.

9.1.5 Summary

The thesis constituted a multi-level mixed-methods study. The contemporary qualitative systems analysis provided not only an explanation for the results of the quantitative impact evaluation, but also a deeper understanding of the functioning of the contemporary system relatively unaffected by the gatekeeping intervention. The historical-institutional analysis provided a theoretical narrative behind the quantitative historical trend analysis of structural and functional balance between hospitals and primary providers. The two historical sub-studies also constituted a historical explanation and deepened understanding of the contemporary system related to the balance between the two sectors, and in combination with the contemporary evaluation generated a historical narrative relevant to contemporary policies.

9.2 Discussion on results

This section discusses the results of the thesis in three sub-sections. The first describes the situation and evolution of the balance between hospitals and primary care

providers; the second evaluates the effectiveness of and challenges facing the gatekeeping pilot in shifting balance towards primary care providers; the third interprets the overall results and discusses the diagnosis, prognosis and solutions of the balance between hospitals and primary care providers.

9.2.1 Situation and evolution of the balance between hospitals and primary care providers

This sub-section draws mainly from the results of the historical chapters and the qualitative systems analysis.

1) Situation

In Chapter 4, a system of metrics was used to measure the structural and functional balance between hospitals and primary care providers in China between 1949 and 2013. As compared with previous studies (for example, Starfield (1994) and Mills (1990)), the sub-study expanded horizontally to cover a systematic range of indicators closely related to the balance and also longitudinally to a longer period (64 years). There seemed to be a diverging trend of the development of hospitals and primary care providers in terms of structural balance by the late 1950s. Hospitals acquired a much higher quantity and quality of practitioners, as well as number of beds, than primary care facilities. From the late 1950s, primary care providers (solely represented by township health centres) started to be equipped with a fast growing number of beds, which was approaching. After the late 1960s, the quantity of primary care practitioners started to increase dramatically, particularly with the huge number of de-professionalized barefoot doctors, and surpassed to that of hospitals around 1978. Due to the absence of data regarding service utilisation, there is a gap of information regarding functional balance of care before 1978.

The years between roughly 1980 and 2002 seemed to witness a stagnation of

development of structural capacity in primary care facilities and deterioration of their functional role, while hospitals maintained stable development in general. Significant gulfs grew between hospitals and primary care providers in terms of human resources, beds and other assets. For example, the number of hospital staff increased by 50%, while the number of primary care staff increased by only 16%. The years after 2002 saw a major boost of both sectors; however, additional resources seemed to have been absorbed primarily by hospitals, and hospitals played an increasingly larger functional role. Hospitals accounted for about 80% of revenue and 88% of value of assets between 2002 and 2013. Hospital admissions more than tripled between 2002 and 2013, while admissions at primary care facilities doubled. Hospital visits increased more than 100%, while primary care visits grew by about 80%. The stratification of quality of doctors and other aspects of technical capacity between primary care faculties and hospitals also worsened in this period.

The sub-study found a similar situation regarding focus of resources in hospitals to that found by Mills (1990) and World Health Organization (2014, p. 6). The thesis also exposed the overlaps of function and resources between hospitals and primary care providers in the context of China, reflected in the huge proportion of hospitalisations provided at township health centres in particular, as well as the large proportion of ambulatory care provided at hospitals. Quantitatively, earlier development of hospitals, was followed by development of the primary care sector with much fewer physical resources but similar number of doctors, providing by the 1980s a similar amount of care. Hospitals became increasingly dominant in the recent three decades. Qualitatively, the primary care sector has a huge number of rural/barefoot doctors with significantly lower qualification and lower pay. Primary care providers were more like a separate sector of basic medical services.

The qualitative systems analysis of the gatekeeping pilot demonstrated weak conditions of primary care in the pilot, particularly regarding the clinical skills of

primary care doctors in comparison with those in hospitals. The sub-study further revealed reinforcing feedbacks that turned into a series of vicious cycles for primary care development, in terms of weakened service capacity of primary care, decreasing patients' trust in primary care and questionable sustainability of human resources for primary care. The lack of progress in reforming hospitals also exacerbated the imbalance between the two sectors as far as the local pilot was concerned.

2) Evolution

The balance between hospitals and primary care providers was not the effect of a clearly discernable policy. On the one hand, the balance was the accumulated effect of a number of institutions and policies; on the other hand, the balance fed back to the policy-making arena via pre-allocated resources and pre-shaped preferences and interests. The study structured the investigation of these institutions and policies to a pair of four sets of key institutions. In light of these theoretical tools, the balance between hospitals and primary care providers resembled a social structure that had stable characteristics and law-like regulating effects on a range of human individual and organisational behaviours. This point was particularly clear in the study of the evolution of this structure.

The thesis was not interested in judging whether the development of a de-professionalized primary care sector was the right thing to do. It was mainly concerned with the contemporary legacies. The historical-institutionalist analysis suggested the existence of three dynamic cycles (1835-1949, 1949-1978, and 1978-2013) of a path dependent process of institutional balance of care.

In the first cycle started in 1835, hospitals were selected as the ideal organizational structure of Western medicine institutions to both showcase technological advantage and support the preaching mission. Overtime, there developed a binding of professional doctors with hospitals, which were also in addition relatively decently

resourced and governed. Between 1928 and 1934, a national health service agenda emerged in a conjuncture, with policy commitment to public provision of modern medicine to the entire population, particularly those in rural areas. The original policy blueprint developed during this period included a professional primary care sector doctor as a supervisor of community health personnel. However, such a design was dropped in the implementation after 1935, mainly due to lack of human resources and resource mobilisation to fund professional primary care, but also because of the lack of cooperation from medical schools. Hospitals and primary care split into two tracks of medical modernization. The curative aspect of primary care remained largely traditional.

The cycle then repeated twice, first between 1949 and 1978, then between 1978 and 2013. Each period saw the initiation of a medical development programme that favoured hospitals, which further strengthened hospitals and then was challenged by a major policy turn towards an agenda of improving balance of care through attempting to strengthen primary care. The second and third cycles witnessed hospitals and primary care facilities growing into a system of two sectors—professionalized urban insurance-based medicine and de-professionalized cooperative medicine by 1978, and hospitals growing into the dominant providers of health services by 2013.

The history suggests the working of two main feedback mechanisms: reinforcing and reactive. The initial development of hospital-centred Westernization of medicine established a privileged status of hospitals that was repeatedly strengthened. Hospitals saw consolidation of professional status, increasing links with patients with higher entitlement or affordability, strong political interests and organisational consolidation in housing large ambulatory service departments. However, the hospital-centric model was not suitable for the whole of China as it was primarily an urban model, had too few positions of doctors, and served a very small proportion of the total population.

In each cycle, the small reach and high cost of the hospital-centric model laid grounds for a reactive process where the value of primary care was emphasized (during the three historical conjunctures identified). The schism between primary care providers and hospitals became the clarion call for revisionary efforts in favour of primary care. However, such efforts were confronted by issues related to coordination difficulties (particularly facing strong network effects in managing multiple interrelated aspects of reform), as well as asset specificities (particularly in terms of human resources for primary care), and self-reinforcing processes, as the demand at hospitals (serving urban privileged population) was realized first, and they tended to resist changes. The two-way movement eventually saw the strengthening of primary care blurred with political rhetoric. The reactive process eventually witnessed the distortion of primary care to an antithesis of hospitals-- a de-professionalized path of development that also mimicked the hospitals' practice of comprehensive ambulatory and inpatient care.

From a historical perspective, the findings from the descriptive history and the analytical history (for the last two cycles, i.e. 1949-1978 and 1978-2013) seemed to converge and corroborate each other. The development path of the balance between hospitals and primary care providers since 1949 was deeply rooted in the changes between 1835 and 1949, which as just mentioned both formed a hospital-centric initial introduction of modern medicine and initiated a track of primary care development where curative care was side-lined by epidemic control. The years between 1949 and 1978 saw the development of two relatively comparable sectors in terms of quantity of services, human resources and beds, despite huge gulfs in institutional underpinning and quality. The post-1978 period saw the centralisation of care and resources in hospitals.

The cost of Western medicine and the coalition between hospitals and more powerful and usually urban interests meant that low-cost primary care was the only

feasible historical option for rural coverage. The decision set in motion a development path with poorly resourced primary care, and this hospital dominance seemed to be very difficult to shift.

As the qualitative systems analysis revealed, efforts to strengthen human resources of primary care providers faced the challenge of syphoning towards hospitals of patient demand (and the financial accompanying revenue), human resources, etc. The causal loop analysis identified the interrelated feedback loops that kept benefiting hospital development disproportionately. Combining findings from the qualitative systems analysis and historical-institutionalist analysis, the historically created development path of hospital centrism remained in place in the pilot which also led towards a weaker role of primary care providers in curative care.

9.2.2 Can the gatekeeping pilot shift the balance towards primary care

Despite previous claims about the potential of gatekeeping in deflecting patient flows towards primary care providers, evidence about effectiveness of gatekeeping in low- and middle-income countries has been rather thin. In this thesis, the impact evaluation of a gatekeeping pilot found a significant increase of ambulatory visits to primary care facilities (an average increase of 38.7%, $p=0.013$), while increase of spending at that level was not significant. The changes in secondary care (district-level hospitals) were not significant in both patient visits and expenditures, suggesting that secondary facilities could have been shifting some patients to primary care facilities while gaining from tertiary care facilities. There was no strong evidence of a decreasing number of patient visits to tertiary care facilities. The expenditures in tertiary care facilities (outside the district) reduced significantly, by 42.7% ($p=0.020$), though the district hospital was a tertiary hospital.

The qualitative systems analysis identified two aspects that led to the sub-optimal functioning of gatekeeping. First was the pre-existing weak capacity of primary care

service providers and the dynamics related to an increasingly hospital-centric balance of care, including the inflationary incentives in hospital care. Second, the effectiveness of gatekeeping was hampered by the unintended consequences of other related policies, particularly the reform of the primary care sector around 2009, which put in place a management arrangement that basically weakened the incentive for provision of primary medical care, and served to reorient primary care doctors to public health work without securing their professional status. The sub-study showed the danger of neglecting the professional aspirations of primary care practitioners and patients' appreciation of their competence, which seemed still to hinge on the ability of primary care practitioners to treat patient illnesses. Findings also shed light on some unintended consequences that are likely to come with gatekeeping. For example, the implementation of gatekeeping in the studied pilot led to resentment by both doctors and patients. This finding contradicts the optimistic results of a patient survey done in Shenzhen (Gan et al., 2016), which showed willingness of local residents to accept primary care facilities as gatekeepers.

The results of the two contemporary sub-studies complemented each other. While the difference-in-differences analysis isolated the effects of the gatekeeping pilot on functioning of primary care, it had some self-contradictory findings, particularly the increase of visits to primary care facilities without significant extra spending. The difference-in-differences analysis could only suggest that the resource utilisation of such visits seemed to be lower. I speculated that it was because large numbers of patients were visiting for referral. This was supported by interviews for the qualitative systems analysis. Furthermore, the qualitative systems analysis provided a rich representation of the causal relationships that restricted the effectiveness of gatekeeping. The causal loop analysis presented the interrelated feedback loops that kept benefiting hospital development disproportionately, while primary care development was stuck in vicious cycles. The feedback loops seemed to be driving a destruction of the professional career of primary care practitioners. Combined, the two

contemporary sub-studies provided not only empirical evidence regarding the limited effectiveness of gatekeeping in shifting balance of care but also the factors that limited effectiveness. From a systems learning perspective, the contemporary sub-studies used the gatekeeping pilot as a case and gained important knowledge regarding challenges facing shifting the balance towards primary care in the current health system.

9.2.3 Diagnosis, prognosis and solutions

The discussion here follows the logic of earlier methodological discussion in Section 9.1 on how the historical analyses add to the contemporary evaluation.

1) Diagnosis

- Unclear role of primary care providers

A challenging issue for sustainable development of primary care is the lack of a clearly defined role of primary care providers. The historical and contemporary analyses converged on the confusion related to the overlapping resources and functions between primary care providers and hospitals. The thesis has shown in the metrics of balance between hospitals and primary care providers that both sectors provide a large volume of ambulatory care and hospitalisations, and primary care facilities had a significant number of beds. Other elements of the confusion relate to the fact that primary care was also a carrier of public health work. The qualitative systems analysis suggests that the drastic pivoting to public health services is potentially both an adaptation to the weak functioning of primary care in serving patients and a result of unclear definition of the role of primary care doctors.

The continuing overlaps on functions and resources between the two sectors, and lasting weakness of primary care providers seemed to fit the criticism of World Health Organization (2008, p. 11) in regard to over-simplification of primary care in resource-constrained settings that primary care has been “reduced to a stand-alone health post

or isolated community-health worker” and “synonymous with low-tech, non-professional care for the rural poor who cannot afford any better”. The thesis has shown that the situation was formed and reinforced historically. From the introduction of Western medicine, primary care has found it increasingly difficult to find a proper place in modern medicine. Primary care has been designed as a low-cost substitution of professional medicine defined by hospitals. The study also reflected the gap between the rhetorical importance of primary care providers to cover a range of functions at low cost and the expectations of good doctors in the mind of patients and providers. Both historically and contemporarily, there seems to have been a tendency that other priorities crippled the development of a primary care profession.

- Repeatedly distorted momentum of primary care strengthening

The role of the sequence of historical events seems central in the analysis, reflected particularly in its effects on the post-juncture periods. This finding is consistent with the work of Hacker (1998). Historical situations played a significant role in predisposing the dominance of hospitals. The first critical juncture was defined as much by the missionaries’ agenda to set up a place of preaching as by the pre-existing traditional medicine world that had left a niche for Western medicine hospitals. The latter two critical junctures reflected the relatively weak institutionalization of the primary care sector as compared to hospitals. The fact that a hospital-centric model of increasingly professionalized medicine was set up at the critical junctures with the absence of attention to the primary care level before the conjunctures of attempts to shift balance towards primary care seemed central to the continuing development of a hospital-centric modern health system. The three critical junctures of the establishment of mission hospitals between 1835 and 1844, the formation of the urban-only entitlement between 1949 and 1953, and the collapse of the cooperative medical scheme between 1978 and 1985, marked significant changes in the agenda of health system development. The three critical junctures all led to subsequent periods of

increasing returns when hospital-related institutions were rapidly strengthened, particularly in the four aspects of professional, financial, organizational and governance institutions. Most conspicuous of the four seemed to be the professional institutions, which were rooted in the first critical juncture initiated by medical missionaries.

Path dependent development seemed unable to foreclose opportunities to change policies, suggested by the repeated conjunctures that aimed at shifting the balance towards primary care, seen in the state medicine programme between 1928 and 1934, the cooperative medicine scheme between 1965 and 1968, and the health system reform between 2002 and 2009. The repeated constraints seen on the agenda to shift the balance of care in favour of primary care that came out of the conjunctures during the post-juncture periods also suggested the importance of the logic within the system that was continued. By the end of the conjunctures, hospital-related institutions had grown stronger while primary care-related institutions remained weak. The post-juncture periods also saw processes of normalization of such primary care institutions, first ceding priority to epidemic control in post-1935, then to rhetorical reconstruction in post-1968, and finally to public health functions in post-2009. On this note, the focus of Hacker (1998) on how history affects the moments of key decision making seems inadequate for complex issues like balance of care, while the attention to systems logic proposed by Tuohy (1999) seems an important complement.

The historical analysis indicated room for optimism about the availability of opportunities for major policy changes as well as scepticism that a single bold movement could simultaneously address all the issues. While the repeated emergence of conjunctures to change policy suggested the biggest problem is not the lack of policy opportunity at key moments of change, the repeated frustration of efforts to shift balance towards primary care providers after historical conjunctures suggested the critical importance of making suitable policies, which required good empirical

evidence regarding both the effectiveness of specific interventions and in-depth understanding about the policy process needed to bring changes to the system.

The distorted reform initiatives ended up reinforcing the previous contrast between hospitals and primary care providers. The post-1935 implementation of the state medicine programme contributed to the diverged track of primary care from comprehensive care involving professional doctors towards primarily focusing on epidemic control. The post-1968 implementation of the cooperative medical scheme also led to the normalization of weakly financed and de-professionalized primary care under a revolutionary rhetoric. The two processes were historical precedents of the post-2009 process of de-professionalizing primary care practitioners, and the resistance to change in hospitals. The contemporary challenges all seemed attributable to the historically strong institutionalization of the hospital sector versus the weak institutionalization of the primary care sector. Particularly, the strengths of hospitals were embedded in a wider social context (e.g. urban-rural gap). Through the lens of path-dependence analysis, these contemporary challenges were not independent and should not be considered separately.

The problems are unlikely to be solvable with one wave of reform. As previous experience and the theoretical analysis of path dependence suggest, the strong hospital dominance is likely to establish network effects that will challenge reform and obstruct policies developed in isolation to address fragmented issues. Problems with professionalization of primary care is likely to face particular challenges due to the asset-specific characteristics of both professional skills and patients' trust. With the challenges in terms of self-reinforcing effects getting ever stronger, the absence of a strong primary care coalition (discussed in the following paragraphs) is likely to make it difficult to address asset specificity. Hence, the primary-care-oriented reformers need to take a long-term perspective with respect to policy making and implementation.

Consistent with the hypothesis of state-centric theorists (Hall et al., 1996; Skocpol, 1985; Skocpol et al., 1986), the state was a key agent that, through both action and inaction, encouraged a hospital-centric development path of balance of care. The complexity of the modern health system seemed to generate a convergence towards greater government interference with third party payers. The complexity of balance of care, which encompasses professional, financial, organisational and governance institutions, was also intertwined with the growing and changing role of the state. The conservatism and inaction of the Qing court were apparently a key factor that enabled missionaries' preference for hospitals to become the mainstream model of modern medicine. Since the establishment of the central government in Nanjing in 1928, the centralized state played an increasingly bigger role in deciding policies related to balance of care. Since then, the state's periodic shift towards attention to a greater role for primary care suggested its ability to stand rather independent from the interference of vested interests that predominantly benefited from a bigger role of hospitals. However, the state's capacity in carrying through changes seemed significantly compromised.

The state's changing preferences seemed to have created interests and power entities that fragmented the decision making arena and made coordination difficult. The analysis has shown that the lack of a strong bureaucratic stake in primary care seems to be a major impediment to shifting the balance towards primary care. This is seen in the post-1928 conflicts between the aspiration of medical colleges for high standards and the requirement of low cost human resources for health, in the post-1968 conflicts between urban demand for hospital care and rural requirements for doctors, in the post-2009 exposure of difficulty in reforming extra-budgetary (pharmaceutical-related) payment for hospitals and hospital doctors. As changing balance of care involves daily decision making of actors with direct influence over both sectors, the state's role extended from that of a negative one involving blocking the influence of vested interests to a positive one involving navigating changes among the various

vested interests, adapting to the power structure regarding daily decision making and adopting policies to fit the situation.

The structural factors were important as backgrounds, and their influences were shown as the balance of care favouring hospitals became embedded in more deeply rooted societal structures. The critical juncture between 1835 and 1844 linked hospitals with the semi-colonial/imperialist footholds of Westerners, while primary care was rooted in rural agriculture. The embedding of a particular balance of care in the rural-urban division was reinforced between 1949 and 1953 with the development of urban-only insurance schemes. The embedding of a hospital-biased balance of care in pro-market reforms with distorted incentives was also key for the period after 1978. The findings of the historical-institutionalist analysis suggested that the changing context was not directly defining the balance of care. The fact that these structural factors did play an important role was because the critical junctures provided the platform for extra-health-system forces to be introduced through the active or passive role of the state in the health system. The predisposition of the balance of care inherited from history shaped the effects of such external structural factors, and was also shaped by such forces.

- Absence of a strong primary care coalition

Related to the unclear role of primary care providers and repeated failure of strengthen professional primary care, there was an absence of a strong primary care coalition. Policy makers needed to know in particular how ready the system was for the intervention to be effective, and how to drive institutional changes. As the post-juncture development involved a series of dynamics among multiple factors including complex historical legacies, building the basis for sustained efforts seemed to be important in navigating through changes of key institutional pillars related to primary care strengthening. This point is reflected in the self-reinforcing development of mission hospitals and the collapse of the cooperative medical scheme. What seemed

to be important is to have strong and sustainable stakeholders to maintain the momentum of primary care development.

The current trajectory of balance between hospitals and primary care providers is a result of the earlier development of hospitals, which triggered three cycles of path-dependent development biased towards hospitals. The history generated a strong bonding between highly qualified human resources and hospitals. The relatively privileged patients that utilised services of hospitals and pharmaceutical industries that depended for their sales on hospitals further reinforced this professional bias. The contemporary unintended consequences of problematic reform of primary care and normalization of weak curative primary care both seemed to be repetition of historical processes. Without a strong primary care coalition, the problems of hospital-centrism will unlikely be solvable.

- Summary

The above three points boil down to the historically created lack of a strong primary care coalition. The lack of a strong professional stake in primary care, the lack of a continuous state bureaucratic stake, the service user/funder relationship, and the organisational boundary between hospitals and primary care providers, fed back to each other. They also confined historical choices and reformers became demoralized (in post-1935), politically radical (in post-1968), or disoriented (in post-2009). As discussed in Chapter 6, this process was also embedded in the unequal social structure of the past, external to the dualist structure.

2) Prognosis

This section discusses the interpretation of results to attempt to extrapolate what is likely to happen. This analysis draws in the expected trajectory based on the diagnosis as well as the changing condition. This process is highly speculative and

contingent on the prediction of factors within and beyond the dualist health system structure.

- Endogenous trajectory

As discussed above, the lack of clear definition of primary care providers and the reinforcing and reactive process set in motion a cyclic process. According to the prediction of this internal logic, the current hospital dominance will likely trigger further action to both contain hospitals' growth and strengthen primary care development. Whether the conjuncture will actually shift the balance of the health system towards primary care depends largely on how the transition during the conjuncture happen.

The historical lack of a strong primary care coalition has been a critical problem facing successful transition towards a primary-care-oriented health system. This contributes to the lack of clearly defined role for primary care, which in turn makes it difficult for primary care coalition to form on a strong basis. Forces behind primary care are going to have very different interests and preferences, easily pulling primary care towards different directions, represented in the history as de-professionalization of primary care, elimination of university-based higher educational track for primary care, and the repeated confusion with public health services. The historically inherited lack of a strong primary care coalition is likely to have consequences in the coming period and be a significant obstruction for its strengthening.

Despite the grim outlook, the coalition for primary care across stakeholders is also likely to receive some strong new allies. There has been an increasing awareness of the importance of high quality general practitioners. Indeed, a number of high-level universities are expanding university degree programmes on general practice, some of whose graduates are being employed at community health centres by some urban government willing to pay generous salaries. This new cohort of professional primary

care doctors may eventually form the basis of primary care professionalism. The government is also trying to make it easier for doctors to open private clinics, to encourage a great number of specialists to undertake services with greater primary care attributes.

- Exogenous factors

For the first time, the state is directly organizing three social medical insurance schemes that essentially covers the whole population. This has implications in terms that financially and governance-wise, the state will have a bigger stake in strengthening primary care. Ongoing reform of public hospitals is finally detaching pharmaceutical sales profit from facility service revenue. While study is needed to evaluate the development, one can speculate that this will lead to organizational containment of hospitals. Overall, the analyses suggest that the next critical juncture will likely work within the constraints and opportunities of government's responsibility. Chinese populations are ageing with a large number of patients suffering from multiple chronic diseases that can be better managed at primary care facilities, and urban need for good primary care has been increasing.

New technology may be disruptive and may blur the organisational boundary between the two sectors. The increasing accessibility provided by information and communication technology may have similar effects to that of the expansion of insurance coverage increasing access to hospitals. Meanwhile, it may also enhance the ability of primary care providers to deliver on “person-centeredness, comprehensiveness and integration, and with a regular point of entry into the health system” (World Health Organization, 2008, p. 42). However, if primary care continues to be weakened without a strong professional basis, technological development might very likely strengthen hospital-centric development yet again.

Urban-rural gap is a lasting characteristic of Chinese society. Previously the issue

of balance of care largely resemble a sub-issue of gap between the urban privileged and the rural underprivileged. In the future, professional primary care is likely to emerge first in urban areas, where the patients' demand can be generated and trust is likely to be created with the rise of the primary care profession. However, rural areas are likely to suffer continuously from lack of strong primary care. Indeed, if the pilot reflects a wide-spread issue, the primary care providers in rural areas are suffering from a collective career crisis. The urban-rural gap would likely be sustained if no special rearrangement is made for rural primary care.

- The fourth critical juncture?

If the above analysis holds, we are likely to be seeing the fourth critical juncture soon. One can speculate that there will be greater divergence between rural and urban areas in development of the balance between hospitals and primary care providers. Primary care is likely to gain a stronger professional foothold and greater stake financially, organisationally and governance-wise in urban areas. However, such changes are unlikely in the rural areas. Therefore, we are going to see a trichotomy of health service delivery: an urban-based hospital sector serving both urban and rural populations, a professionally stronger primary care sector in urban areas and a still de-professionalized primary care sector in rural areas.

3) Solutions

Here the discussion about solutions draw from previous analysis regarding the diagnosis of problems as well as the prognosis of further development of the balance between hospitals and primary care providers.

- **Facilitating the development of primary care profession**

While there are positive signs that the primary care profession may have some welcome beginning rooted in the medical university, one cannot be sure when this

fundamental shift will take place. The conflicts among multiple reform initiatives suggest grave danger of further weakening the basis of primary care particularly in rural areas. The destruction of the current primary care profession will need to be put to an end, via reform of salary policies and improved access to technologies and medicines.

Further challenge is likely to lie in the problem of asset specificity. A primary care profession is a highly specific asset, as are most medical specialties. Developing a strong cohort of primary care professionals requires more than political diktat but also the willingness of doctors and would-be-doctors to take primary care as their specialty. Experience from the UK suggests the importance of securing the professional independence of primary care (as general practice) as a speciality-like discipline, empanelment of patients etc. as well as the construction of a strong professional body (the Royal College of General Practitioners) to maintain high standards of training and practice (Kmietowicz, 2006).

Here I draw from the five elements of professionalism identified by Freidson (2001) to suggest a number of policy options:

First, general practice needs to be defined clearly as a specialized clinical discipline. Further recognition of the status of general practice (including its own professional and academic body) on an equal footing to other specialities is needed.

Second, contract and salary systems for primary care doctors need to improve their pay to attract well-qualified professional primary care doctors. Given the diverse situation in China and the meagre number of well-trained general practitioners, a feasible solution maybe to have some higher level general practitioners to supervise a large number of relatively less well trained ones, particularly in the rural areas.

Thirdly, a “sheltered” position for primary care providers providing security

through a contract with patients based on the number of enrollees seems suitable. Patients referred by primary care doctors should be given rapid access in seeing hospital specialists.

Fourth, a formal schooling and continuous education programme based on universities that produces qualifying credentials is needed. Such credentials should carry equal weight as credentials for training in other medical specialties. Again, a tiered system may be necessary in adapting to the variations of conditions within China.

Fifth, a primary care ideology needs to be fostered that asserts greater value and commitment to the core primary care attributes, i.e. “first contact, accessibility, longitudinality, and comprehensiveness” (Starfield, 1992). In other words, the value of maintaining community health rather than providing advanced specialized services needs to be recognized.

As the study has shown, the issue of redefining the role of primary care requires a systems change, including a recognition of a gradual transition of primary curative care, a process of trust building, etc. To address such coordination difficulties, the governance for primary care will likely need to adopt two changes. On the one hand, a more streamlined and coherent administrative structure to manage these multiple reform initiatives. On the other hand, the primary care development will also need to utilise to a greater extent professional autonomy that can respond to patients’ need and professional requirements more adequately.

- **Facilitating transition towards primary care strengthening**

There are likely to be forces holding back the primary care strengthening process during the coming critical juncture. History provides a critical perspective on the role played by contemporary intervention and evaluation. As the historical study suggests, primary-care-oriented reformers need to harness the power of reinforcing and reactive

feedbacks and put the temporal dimension into the design and implementation of intervention.

Internally, strong dominance of hospitals will trigger a reactive backlash as happened during previous conjunctures. Some external factors beyond the dualist structural also provides additional opportunities: changing demand, slowdown of economic growth and rapid ageing might tilt the balance towards greater role of primary care providers. An emerging middle class might also buy into the services provided by a new breed of primary care providers. A hospital-centric model to take care of the medical need of an ageing society is simply unlikely to be financially feasible for China.

In the past, however, efforts to strengthen primary care have been shaped/inhibited by the reinforcing feedback towards greater hospital dominance from earlier period. Frustration of efforts to shift resources to primary care has led to development of primary care on a de-professionalized path. The privatization of primary care by increasing the role of private clinics might lead to a breaking down of solidarity and eventually give rise to two-tier primary care even within urban areas. While rich people are likely to enjoy good quality primary care, the less well-off are likely to be confined to greater reliance on hospital care.

In the past, policy shifts towards primary care have been seen to be closely linked with the agenda to provide universal health coverage. What has been missing is the necessary change in financial allocation to overcome the challenge placed by structural inequality (particularly in terms of the rural-urban divide). Development of professions require the support of government (Freidson, 2001). Primary-care-oriented reformers need to form a strong coalition around the primary care professionals and use the currently strong government commitment to universal health coverage to usher in a strong social movement towards strong primary care. This coalition needs to be

politically strong but not over-politicized such that reform is replaced by rhetoric. The role of evaluation in facilitating historical changes needs to be reconsidered. Tasks for evaluation is further discussed in Chapter 10.

The historical-institutionalist analysis shed light on the importance of the role of the state, particularly, in responding to innovative technology (progress in bio-medicine) and in influencing health system arrangements (progress in expansion of coverage). The pre-1928 absence of the government's role reinforced the development of hospital-centric modern medicine in China. The post-1949 development of urban-only insurance schemes tied the bureaucratic interests with hospitals, contributing to impediments to reform after 1965. A mistake of policy makers during the cooperative medical scheme period was to focus on propagandist packaging of primary care as the antithesis of professional medicine and bureaucratic control, without addressing the difficulty regarding cooperative medicine (e.g. unstable and inadequate financing), which foregrounded the post-1978 process of rapid hospital-centric development. The post-1978 reinforcement of incentives for revenue generation with the government's distorted intervention on pricing favouring technology and reimbursement based on pharmaceutical mark-up also favouring hospitals proved difficult to roll back. The gatekeeping sub-studies also provided a case showing that well intended government policies may end up in suboptimal outcomes or unintended consequences, and that policy makers need to be informed by empirical evidence.

9.3 Relevance and generalizability of findings of sub-studies

This section discusses the relevance of the study's findings for China as a whole, as well as whether aspects can be generalised to other low- and middle-income countries (LMICs).

9.3.1 Trend analysis of structural and functional balance of care

The descriptive history of structural and balance between hospitals and primary

care providers used data from national statistics and represented the situation of the entire country, though data on some indicators were not available during the first few decades after 1949. Some of the indicators were more relevant to the history of China than other countries. In particular, primary care providers in China had a large number of inpatient beds (the number of beds in primary care facilities was almost comparable to that of hospitals around 1980). China also had a complicated system of medical qualification, and many doctors did not have a full medical degree. However, the overall structure of metrics and the process of generation of the structure may still have value for other LMICs, as this seems to be the first attempt to systematically develop metrics of the structural and functional balance of between hospitals and primary care providers outside the high-income world.

9.3.2 Historical-institutionalist analysis of coevolution of hospitals and primary care providers

The use of a path dependence framework in the historical analysis may limit the transferability of findings to other countries. The historical logic of coevolution of hospitals and primary care providers in China was a combination of unique sequences of historical events, unique state actors and unique structural factors. As the historical-institutionalist analysis suggested, the sequences of key historical events, in particular the earlier institutionalisation of hospitals, were among the central explanatory factor for understanding changes in balance of care over time. The state actors were agents of change, who mediated the sequence of events. The societal factors (such as the urban-rural divide) were the canvas of the historical process but did not determine the shape of institutions and the structural and functional balance between the two sectors.

Despite the unique socio-political structure of China, modernized hospitals and de-professionalized primary care providers have been common historical legacies of health systems in low- and middle-income countries (Abel-Smith, 1994; Bloom, 1985). A large number of developing countries have gone through similar historical stages.

Western missionaries were the main force behind building medical facilities, particularly hospitals, while the colonial state mainly focused on combating epidemics and generally did not spend heavily on medical facilities (Hardiman, 2006, p. 5); health coverage was usually provided initially only for privileged population groups and usually in urban areas; and there have been major pro-market reforms in recent decades. All these processes were conducive to the development of path dependence of a hospital-centric model. This hospital-centric model, as reflected in the Chinese experience, was difficult to change and poses challenge in the movement towards universal health coverage, and resources are likely to be disproportionately funnelled to a privileged hospital sector. In short, while the dynamic path dependence analysis implies caution in generalization of findings across countries, it also suggests that the existence of some common features of LMICs makes the analysis relevant to them.

The analysis of conjunctures in China approximately coincided with the three waves of the international primary health care movement (the 1930s, the period around 1978, since the 2000s). All three waves of movement rejected a Western and hospital focused model of health care for developing countries. The historical-institutionalist analysis of China may be relevant to the body of literature regarding the primary health care movement internationally.

9.3.3 Qualitative systems analysis and impact evaluation of the gatekeeping pilot

The case study of the contemporary gatekeeping pilot used information from the rural population of a district in a metropolitan city in North China. Therefore, generalizability of the study's findings from the case data to larger geographical units cannot be achieved through statistical inference. However, most of the policies involved (with the exception of gatekeeping) were made nationally and implemented nation-wide. The issue of structural and functional imbalance between hospitals and primary care facilities has been a nation-wide phenomenon as reflected in the references cited above from nation-wide studies. On the basis of what Yin (2009)

defined as analytical generalization, which builds generalization upon theoretical comparability [58], this first qualitative evaluation about a pioneer gatekeeping pilot is relevant to comparable settings in rural China, which faces essentially similar challenges.

9.4 Methodological reflection: a cross-disciplinary mixed-methods study approach in health policy and systems research

The section discusses the innovativeness, strengths and weaknesses of the methods used in this thesis. The discussion starts with the overall study design and then covers each of the sub-studies.

9.4.1 Overall study design

1) Innovation

The thesis used an innovative cross-disciplinary research framework; combined the analyses of historical processes and a contemporary pilot; and studied the development of hospitals and primary care providers in a unified framework.

2) Strengths

The research framework reflects the holistic, complex and multi-dimensional nature of health systems that differs from general health technology and program intervention.

As Adam et al. (2012) have argued, systems thinking has proved valuable in revealing key elements of success and failure in implementing complex interventions, the role of relationships, the role of actors in health systems, the importance of environmental factors, and the importance of anticipating potential unintended consequences, and systematically evaluating the implementation process and reacting to feedbacks within the systems. As Adam (2014) has further discussed, combining

qualitative methods with systems thinking can add depth to analysis of health systems issues, and adding visualization can help convey complex interpretations and findings.

By dividing the contemporary part into an impact evaluation and a qualitative systems analysis, the study allowed both components to use advanced research methods. While the quantitative component revealed the effect size, the qualitative component revealed the dynamics of the mechanisms and the contexts from which the separation of intervention mechanisms was possible only theoretically and with problematic simplification. Through the combined use of both approaches, the thesis addressed the several issues of complexity mentioned by the Medical Research Council (2008) guideline on evaluating complex interventions. The qualitative systems analysis facilitated good understanding of the many details of the causal chain, implementation problems and unintended consequences, suggested in the MRC guideline. It further noticed indirect factors that are easily neglected but potentially very important, and placed these factors in a dynamic system.

The study design had the benefits of taking advantage of state-of-the-art research approaches from different disciplinary background. Bloom (2014) argued for the importance of situating health systems development in the context of history and political economy. By separating the historical part from the contemporary part, the study allowed a sufficiently sophisticated development of a theoretical framework based on historical institutionalism, which was used to study historical changes rigorously. On the other hand, the separation also allowed a relative simplification of political, social and economic issues when it came to the qualitative systems analysis, in which attention was devoted to the complexity within the immediate contemporary health systems building blocks. By separating the trend analysis and historical-institutionalist analysis, the study allowed a systematically constructed and quantified measurement of the development of primary care facilities and hospitals from 1949 to 2013, but also provided the factual basis for in-depth historical analysis of the

institutional balance of care.

The study minimised the potentially complex issue of merging data from multiple types of data sources, by keeping separate sets of data for individual sub-studies and conducting sub-studies separately. This overall design therefore allowed each method to play out to its full potential without complicating the process of data management and analysis. In aspects where the methods used intersected with one another, e.g., both qualitative and quantitative components addressed the outcome of the gatekeeping pilot, they allowed the thesis to construct explanatory understanding of the social world of patients and practitioners, to better understand why and how some stakeholders did not appreciate the gatekeeping policy.

The method represents an impact-oriented research approach as policy implications feature critically in the conceptualization of the balance between hospital and primary care providers.

3) Weaknesses

The broad conceptualization of balance between hospitals and primary care providers sacrificed the internal diversity of health delivery. The different research approaches were not directly coherent and required a complex mix of skills. Variation in levels of geographical representatives of sub-studies made synthesis of findings difficult.

9.4.2 Sub-study 1: historical trend analysis of the balance between hospitals and primary care providers

1) Innovation

The study developed a system of metrics of the structural and functional balance between hospitals and primary care providers; and used the system of metrics to

quantitatively describe the development of the hospital-primary dualist structure of the health system from 1949 to 2013.

2) Strengths

The sub-study quantitatively reflected the changing trends of structural and functional balance between the two sectors.

3) Weaknesses

Not all data were available, particularly for the earlier periods. Indicators that reflect the value and quality of care were lacking.

9.4.3 Sub-study 2: Historical-institutionalist analysis of the coevolution of hospitals and primary care providers

1) Innovation

The sub-study innovatively used historical-institutionalist analysis to study the history of health systems; identified three cycles of dynamic path dependence, and the historical-institutional challenges and opportunities facing primary care strengthening in China; provided some groundwork for further study on developing theories of health system development based on China's experience.

2) Strengths

The sub-study analysed historical changes using a rigorous in-depth historical analytical framework and protocol.

3) Weaknesses

Details were sacrificed to some extent for breadth.

9.4.4 Sub-study 3: Impact evaluation

1) Innovation

Innovative use of real world data (medical insurance claims data) to evaluate the effectiveness of the pilot, contributed to filling the gap of lack of impact evaluation of gatekeeping in low- and middle-income countries

2) Strengths

Used counterfactuals to calculate the net effect of the pilot intervention and controlled for covariates.

3) Weaknesses

Any difference in trends was attributed to the pilot programme. The scale of the gatekeeping pilot was limited. The quality of claims data was affected by changed reimbursement procedure.

9.4.5 Sub-study 4: Qualitative systems analysis of gatekeeping pilot functioning

1) Innovation

Innovative use of systems dynamics in a qualitative health systems analysis, provided an example of a standardized protocol of qualitative systems analysis in health systems research

2) Strengths

Revealed the complexity involved in shifting balance of care, the dynamics of the mechanisms and the contexts. The causal loop diagram was explicit and transparent.

3) Weaknesses

The system dynamic model was not validated with the interviewees or independent experts. The scale of the gatekeeping pilot was limited.

Chapter 10 Conclusions and recommendations

10.1 Critical reflection of overall methods

The primary aim of this thesis has been to understand how to strengthen primary care in China, looking at the effectiveness and functioning of a gatekeeping pilot, and more broadly at the history of the balance between hospitals and primary care providers, in order to contribute to policy development in China as well as other similar settings. Through applying a mix of health system research approaches (trend analysis, historical-institutionalist analysis, qualitative systems analysis and impact evaluation), a secondary aim of the thesis was to explore how mixed methods can be used to study a complex system-related issue. To accomplish these aims, the thesis has employed four sub-studies: sub-study 1 described the historical trend of balance between hospitals and primary care providers in China; sub-study 2 tried to understand the historical logic of the coevolution of hospitals and primary care providers in China using a dynamic path dependence analytical framework; sub-study 3 evaluated the effectiveness of a gatekeeping pilot in China; sub-study 4 sought to understand how gatekeeping functions in the pilot in the context of China's recent reforms using a qualitative systems analysis.

The thesis used an innovative cross-disciplinary research framework; combined the analyses of historical processes and the evaluation of a contemporary pilot; and studied the development of hospitals and primary care providers in a unified framework. The research framework reflected the holistic, complex and multi-dimensional nature of health systems that differs from general health technology and program intervention. The method represents an impact-oriented study, as policy implications feature critically in the conceptualization of the balance between hospital and primary care providers. The study design had the benefits of the comparative

advantage of state-of-the-art research approaches from different disciplinary backgrounds. However, the broad conceptualization of balance between hospitals and primary care providers sacrificed the internal diversity of health delivery. The different research approaches were not directly coherent and required a complex mix of skills. Variation in levels of geographical representativeness of sub-studies made synthesis of findings difficult.

Nevertheless, the use of multiple methods for multiple levels of complexity allowed the thesis to tap the comparative advantage of rigorous and innovative research approaches and methods that have been shown to be useful for understanding health policy and systems issues. Multiple methods allowed the study much deeper understanding—an advantage of mixed-methods research (Bazeley, 2004) and a concrete appreciation of the gatekeeping pilot as an intervention among multiple social processes that affected balance of care.

10.2 Summary of findings

10.2.1 Findings of the historical trend analysis

The number of staff at hospitals overtook that of primary care since the 1980s, while the quality of doctors at hospitals was consistently much higher than the usually under-professionalized primary care providers. Hospitals also had stronger physical assets. Functionally, the share of hospitals in both ambulatory and inpatient care was substantial, along with a growing share of total health expenditures in hospitals. The increase in hospital dominance was obvious after 1978, associated with two decades of deteriorating primary care (particularly township health centres) in ambulatory care provision and a recent hospital boom in both inpatient and ambulatory care.

Hospitals and primary care providers were not easily distinguishable. Hospitals had been a key provider of ambulatory care in the last three decades, while primary care providers as a whole provided a substantial amount of inpatient care in township

health centres. However, primary care also developed a large sector of de-professionalized staff at the village level. What was formed seemed a two-tier delivery system with a similarly hybrid structure of inpatient and ambulatory care but a stratified quality of practitioners. At the end of the period studied, hospitals appeared to be enjoying an environment, including premium staff quality and pay, superior assets, increasing share of ambulatory care, and orientation towards inpatient care, that fundamentally gave them an edge over primary care facilities.

10.2.2 Findings of the historical-institutionalist analysis

The coevolution between hospitals and primary care providers seemed to have experienced three path-dependent cycles (1835-1949, 1949-1978 and 1978-2013). Each cycle started with a critical juncture leaving or exacerbating a split between the two sectors biased towards hospitals. Afterwards, the underpinning institutional complex went through self-reinforcement processes strengthening the professional, bureaucratic/philanthropic or financial institutions in hospital dominance. At a later point of conjuncture, major centralization of power in central leadership (i.e. the establishment of a national administration between 1928 and 1934, the Cultural Revolution (including its eve) between 1965 and 1968, and the shock of SARS and the change of political administration and agenda between 2002 and 2009) saw reactive policies introduced to correct previous policy bias towards hospitals.

With reinforcing processes proceeding, the reactive processes came into an existing institutional environment that exhibited strong resilience that the reactive processes could not overcome and had to adapt to. Therefore, the impacts of these reactive policies on the path of coevolution of hospitals and primary care providers were limited and further distorted the development of the two sectors. The result was repeated attempts to develop primary care as an antithesis of hospitals in all three cycles, with attempts to strengthen the four institutional pillars underpinning the coevolution of hospitals and primary care providers (professional, organizational,

financial and governance institutions). The weakness of such institutionalization turned into disadvantage for primary care.

The existence of the three conjunctures suggests that policy windows for change were not rare within the Chinese political system, however, the frustrated reforms after the conjuncture imply that the path dependence effects caused by policy details and feedback effects need to be attended to. Historical sequence played a central role in the evolution of the balance between the two sectors within and across the three cycles.

10.2.3 Findings of the impact evaluation of the contemporary gatekeeping pilot

The impact evaluation suggested that the gatekeeping pilot had a modest impact on shifting patient and revenue from hospitals to primary care facilities in the pilot areas, while the most significant shift of treatment revenue took place within tertiary hospitals rather than between hospitals and primary care facilities. The impact evaluation demonstrated significant results in increased ambulatory visits to primary care facilities and reduction of expenditures at extra-district tertiary hospitals. More patients (an average increase of 38.7%) visited primary care facilities because of the gatekeeping pilot programme, while their spending was less obvious. The changes in secondary care (district-level hospitals) were less clear in both patient visits and expenditures. There was no strong evidence about a decline of patient visits to tertiary care facilities, probably due to data issues. The expenditures at the tertiary level declined significantly, suggesting stronger ability of doctors to enforce referral mechanism at the district hospital with tertiary care facilities.

10.2.4 Findings of the qualitative systems analysis of the gatekeeping pilot

The qualitative systems analysis found that a large proportion of patients visited primary care simply to get a referral, leading to dissatisfaction of both doctors and patients, after the implementation of the pilot policy. The sub-study identified three issues behind the sub-optimal functioning of gatekeeping. First was the pre-existing

weak conditions of primary care service providers and the dynamics related to an increasingly hospital-centric balance of care. Second, there were a series of vicious cycles for primary care development, in terms of weakening service capacity of primary care, decreasing patient trust in primary care and questionable sustainability of human resources for primary care. Lack of progress in hospital reform exacerbated the imbalance between the two sectors. Third, the effectiveness of gatekeeping was hampered by the unintended consequences related to conflicts among different priorities required of primary care development.

The sub-study showed the danger of neglecting the professional aspirations of primary care practitioners and patients' appreciation of their competence, which seemed still to hinge on the ability of primary care practitioners to treat illnesses. Findings also shed light on some unintended consequences that are likely to come with gatekeeping, including resentment by both doctors and patients.

10.2.5 Findings regarding combined research approaches

The thesis, with its combination of four sub-studies, has systematically addressed different levels of complexity: from studying a single intervention relatively out of context, to a comprehensive analysis of a contemporary system in which the intervention is embedded, and a historical system spanning across a prolonged period of time. The use of multiple approaches has allowed the study to capitalize on the comparative advantages of different methods. The thesis has demonstrated the feasibility and value of combining programme evaluation and historical analysis to address a complex health system issue.

10.2.6 Overall interpretation of results

The gatekeeping pilot led to the increase of visits to primary care facilities without significant extra spending. This seemed to be the result of patients visiting primary care practitioners only for referral to hospitals. Besides existing weak conditions of

primary care, the factors underlying the ineffectiveness of gatekeeping included feedback loops that kept benefiting hospital development disproportionately and that were driving a destruction of the professional career of primary care workers, as well as unintended consequences caused by ill-coordinated policies. The dominance of hospitals within health delivery has been a long-term issue since 1949, with institutional roots dating back to 1835. The current dynamics between hospitals and primary care providers were the result of the unclear role of primary care providers, repeatedly distorted momentum of primary care strengthening, and absence of a strong primary care coalition.

According to the prediction of this internal logic, the current hospital dominance will likely trigger further action to both contain hospitals' growth and strengthen primary care development. While urban areas are likely to see development of primary care strengthening based on professionalization, rural areas are likely to get stuck in the currently de-professionalized situation. Therefore, there may be a trichotomy of health service delivery: an urban-based hospital sector serving both urban and rural populations, a professionally stronger primary care sector in urban areas and a still de-professionalized primary care sector in rural areas.

10.3 Policy implications

10.3.1 Policy implications for China

This sub-section deals with policy implications for China in six aspects: the implication regarding the situation and trend of balance between hospitals and primary care providers in China; gatekeeping; reform of the primary care sector; reform of the hospital sector; sequence and coordination of reform policies; and supporting institutions.

1) Situation of balance between hospitals and primary care providers

The thesis has implications regarding the nature of the situation of balance between hospitals and primary care providers in China. The trend analysis suggests a worsening balance between hospitals and primary care providers in recent decades, which strongly supports the ongoing policy agenda on shifting the balance towards primary care. The thesis also suggests that the issue of hospital centrism developed over a prolonged period of time and involved complicated institutional factors. Therefore, policies aiming at short-term solutions without addressing underlying dynamics may have limited effectiveness. The thesis also suggests that the development of primary care providers and the development of hospitals were not independent processes, but instead closely interrelated. Policies need to be devised with both sectors considered as within a single system.

2) Gatekeeping

Under the current situation, systemic gatekeeping is not likely to be as effective in shifting balance between hospitals and primary care providers as previously claimed and would probably incur public resentment. The main barrier to successful gatekeeping is the lack of service capacity of and patient trust for primary care providers. Given that the situation of weak primary care is not limited to the local pilots, gatekeeping across the board seems undesirable. Further piloting of gatekeeping should be considered in combination with more specific arrangements adjusted to local situation. For example, gatekeeping for particular diseases can be piloted with support from high level hospitals in terms of lending experts in the specialty and establishing referral partnerships, and more targeted capacity building of primary care providers.

3) Primary care sector

The thesis findings suggest the need for serious policy attention to the potentially

worsening trend in the primary care sector's function in providing curative care. As strong primary care is critical to deal cost-effectively with the growing burden of non-communicable diseases and population ageing, the weakness of primary care as a medical profession will mean lower performance of health systems. Two kinds of logic are needed to improve the underlying conditions of gatekeeping. On the one hand, the vicious cycle that primary care facilities are facing requires bold and timely measures or experiments. The recent restriction on revenue distribution may need to be loosened to allow primary care facility managers a greater degree of freedom to motivate staff to provide curative care. In particular, it seems necessary and urgent to elevate the competence of primary care doctors, who should also be provided with career prospects.

On the other hand, the findings suggest attention to some fundamental aspects of reforms regarding primary care. There still lacks a clear vision of a professional career for primary care practitioners, which contributed to policy incoherence and even undermined the sustainable development of primary care as a medical profession. Drawing from British experience, formation of a strong independent professional body representing general practice seems a helpful step forward. Rather than shuffling of functions, the policy makers should design reform in which primary care doctors can consolidate their professional basis and the trust of patients and colleagues. Raising salaries and enhancing professional career prospects in general practice at primary care level thus is the priority. In the short run, in order to encourage hospital doctors to move to primary care facilities, pilot programmes may be needed to pay primary care doctors at a higher level to overcome the initial uncertainty. It is likely that urban and rural primary care will diverge as professionalization seemed more likely to happen in urban areas, while rural primary care facilities is likely to be continuously stuck in a de-professionalized situation.

Overall, primary care reform needs a more radical experiment with the aim at

building up professional status in terms of career prospects, the level of training, income, and trust by patients and other health professionals. Lessons can be drawn from the efforts and failure of the 1930s and 1970s, when policies or pilots were made to place highly trained medical professionals in rural primary care settings. A team-based approach using a mix of variously trained staff seems necessary in satisfying both the need to establish a primary care profession and the increasing requirement for multiple functions at primary care level, particularly in rural areas.

4) Hospitals

The thesis suggests the need to push for further hospital sector reform, particularly addressing the payment scheme oriented towards technology (related to devices and pharmaceuticals). As resources have accumulated disproportionately in hospitals, addressing the problem of balance between hospitals and primary care providers should tackle the reform of hospitals and allow greater financial and human resources to flow to primary care facilities. Without addressing rapidly increasing demand on hospital services, it would also be unlikely to be possible to rely on hospital doctors to help strengthen primary care capacity. Financially, the reimbursement scheme related to technology needs to be reformed, perhaps through both adjustment of the price schedule to reduce the reliance of physician income on technology and pharmaceuticals, and greater regulation of hospital technology based on regional planning and health technology assessment. Human resources are a fundamental factor in causing the resource accumulation in hospitals. The bonding between hospitals and professional doctors needs to be loosened to allow a greater number of doctors to flow to primary care.

5) Sequence and coordination of reform

The thesis suggests the need for rethinking of reforms regarding the temporal dimension. Reform of hospitals and the primary care sector should be taken in proper

sequence. Improving primary care professional status and revenue should take place while controlling the unnecessary use of technology in hospitals. Otherwise, hospitals would monopolise the supply of good doctors. Sequence of reform measures needs to take into consideration both the historical legacy and the dynamics in the contemporary system. Part of the implementation failure of the gatekeeping pilot was due to the lack of provider incentives. The providers seemed to have been motivated by a fee-for-service scheme. Thus the reform on payment incentives weakened the motivation to provide services, and undermined entrepreneurship in providing more curative care. Thus, policy makers need to design a process of transition, and not forgo financial incentives directly. The reform of hospital payment mechanisms should also take place before primary care reform to inhibit the worsening brain drain caused by the reform sequence favouring hospitals.

In relation to sequencing reform measures, coordination of reform policies is critical. Addressing the interlocking problems should require attention towards the long term goal and a series of interrelated changes, rather than focusing on the immediate and fragmented issues. Measures should be taken to address the loss of trust of primary care doctors in relation to the change in their service functions. Long-term construction of trust will likely depend on development of primary care as a medical specialty, with equal professional status. Urgent actions will also be needed to restore and strengthen some key curative functions. Services like surgery, appear to be key in determining professional status at township health centres for a long period, and thus need to be carefully handled during a transitional period. Coordination of reform will also require better understanding of the perspectives of patients and doctors, e.g. in relation to the trust in primary care doctors' professional competence.

6) Supporting institutions

The thesis suggests that a supportive institutional environment is crucial. History

suggests that major critical junctures or conjunctures for policy changes were not rare in China. The impacts of previous conjunctures were compromised by the weak institutional legacies of primary care. Reform should therefore focus on enhancing the bureaucratic and professional stake in primary care, along with institutionalizing sustainable financial and organizational changes mentioned earlier.

Regarding creating professional institutions conducive to improved balance between hospitals and primary care providers, history offers important lessons: During the 1930s, it was professional medical elites who embodied the agenda of social medicine, which combined social reform with medicine, professional medicine with public health, curative medicine with preventive medicine, hospitals with primary care providers. The post-1949 reduction in the role of these professional medical elites seemed a loss to the long term momentum of professional primary care. During the 1960s and 1970s, the establishment of the vast Cooperative Medical Scheme and Barefoot Doctors programme did not lead to consolidation of institutions needed to build a sustainable professional or bureaucratic stake in the primary care sector. Instead, as ideological zealotry replaced pragmatic and accountable reform, the rhetorical reconstruction of weakly institutionalized rural primary care paved the way for the post-1978 rapid increase of hospital centrism. Based on the findings from the local pilot, there is risk that this lesson regarding problems related with ideological packaging was not fully learned. The role of the state needs to be developed in partnership with professional and societal interests, with the aim of strengthening the bureaucratic stake in primary care sector.

10.3.2 Policy implications for other low- and middle-income countries

The sub-section deals with policy implications for other low- and middle-income countries, in the following aspects.

1) Situation of balance between hospitals and primary care providers

The thesis quantified the issue of hospital centrism in the case of China, providing evidence to the international agenda on shifting balance from hospitals towards primary care providers. It has suggested that there is value in monitoring and evaluating the situation and trend of balance between hospitals and primary care providers using a system of metrics. Furthermore, the issue of balance between hospitals and primary care providers is likely to involve complicated institutional factors and deserves a long-term perspective in designing policy interventions. Balance between hospitals and primary care providers is an issue across the primary care and hospital sectors, and policies designed solely for one of the two sectors would likely be limited in effectiveness.

2) Gatekeeping

The thesis suggests that gatekeeping is unlikely to be effective in shifting balance between hospitals and primary care providers in settings similar to rural China with weak primary care capacity, and would likely invoke public resentment. Policy pilots need to be attempted in areas with better conditions, and combined with supporting policies, including collaboration with hospitals, perhaps selectively for specific health problems.

3) Primary care providers and hospitals

The thesis has highlighted the need to be clear on the role and content of primary care. Policy makers should focus on the central issue of human resources for primary care, and patient trust of primary care staff. A team-based approach for primary care involving both professional primary care doctors and other grassroots staff of different training will likely be conducive to institutionalizing primary care profession and addressing the multiple expectations for primary care providers. Development of

primary care sector requires the containment of hospital expansion.

4) Sequence and coordination of reform

Given the strong role of historical predisposition of balance between hospitals and primary care providers demonstrated in the thesis, policy makers in other low- and middle-income countries should take existing institutional legacies and entrenched interests into consideration in designing policies to shift the balance between hospitals and primary care providers. At critical junctures of change, policy makers probably need to anticipate and avoid losing opportunities. Policy makers should not only attend to the big moments of radical change, but also address the daily decision making in order to institutionalize policies that benefit the primary care sector.

Lessons should be learnt from China's flaws in reforming its health system, e.g. its extension of insurance coverage before addressing distorted incentives in hospitals, and the combination of delay of hospital reform with rapid primary care reform. Extension of insurance coverage should be informed by consideration of its consequences for balance between hospitals and primary care providers, and should take place alongside strong primary care improvement efforts, to avoid reinforcing an existing hospital-centric model of health service delivery.

10.4 Areas for further research

10.4.1 Metrics and trend analysis of balance between hospitals and primary care providers

Future studies should continue to develop a system of metrics to measure and compare the balance between hospitals and primary care providers in countries. Given the increasing attention to the issue of hospital centrism in low- and middle-income countries, international comparison may be a useful next step for research. Monitoring balance between hospitals and primary care providers at different levels of temporal

and spatial units (from historical to contemporary, and from global, to national and local) may offer valuable information and push for further development of knowledge in this area, e.g. forming and sharing knowledge regarding how to shift balance between hospitals and primary care providers.

10.4.2 History

The historical study about the balance between hospitals and primary care providers in China will be enriched by more local archival work and oral history. This would allow researchers to develop one or more case studies to examine further whether the cyclical model of the development path of balance between hospitals and primary care providers in China proposed in the thesis successfully explains the development of local or regional hospital systems.

Historical institutionalism offers an important lens to develop studies that can help understand contemporary health system issues, and the functioning of interventions. There is the possibility to go deeper with regard to the role of interests (the medical profession, the insurance funds, the teaching hospitals, the different factions in the bureaucracy, etc.).

Third, comparative analysis could be of value both within China at a subnational level and between China and other countries with similar contexts.

10.4.3 Strategies to shift balance towards primary care

More empirical studies are needed about other methods, besides gatekeeping, to shift balance towards primary care. Particularly, studies need to include specific outcomes related to the balance between hospitals and primary care providers, and policy processes related to efforts to strengthen the bureaucratic and professional stakes in primary care and balance between hospitals and primary care providers.

10.4.4 Qualitative systems analysis

The qualitative systems analysis approach shows value in building rigorous study of interrelated issues to understand the functioning of interventions in the system in which interventions are embedded. From this perspective, the qualitative systems analysis approach developed in the thesis seems a valuable addition to impact evaluation. From a systems learning perspective, systems analysis about local cases might offer important opportunities to gain additional learning about the systems the interventions were intended to improve. Hence, qualitative systems analysis could be applied further in combination with intervention studies.

10.4.5 Impact evaluation

The thesis indicated prospects regarding the use of routine data in health systems research. Given rapid increase and digitization of claims and other routine data, this area is likely to see radical growth in the coming years. Researchers should be prepared to make full use of this new opportunity, given the importance of the issue of referral and the ineffectiveness of gatekeeping without strong primary care. There are likely to be more opportunities for future studies in other areas and on a larger scale. Future impact evaluation in relation to balance between hospitals and primary care providers should seek to test the effectiveness of combining referral/gatekeeping with interventions regarding strengthening primary care service capacity.

10.4.6 Using history and systems analysis for health policy and systems research

The thesis exemplifies a possible approach of using a multi-methods research design, involving theories and methods from multiple disciplines, to address a specific issue. Further research might build on this approach and combine different methods. In particular, future studies should further explore the value of combining historical institutionalism and systems analysis as useful tools to increase understanding of complex health system issues.

10.4.7 Research agenda for health policy and systems research

Health policy and systems research is a nascent field, drawing on different approaches rooted in different disciplines and fields. The thesis has sought to acquire deep knowledge by employing different approaches to study a specific topic. Such problem-based mini-fields are likely to be ideal grounds for different approaches rooted in different disciplines to converge while maintaining rigour. In particular, the thesis demonstrates the feasibility and value of combining a historical study with a contemporary evaluation. This is in line with a more sophisticated understanding of health systems as a continuously changing complex system with both internal and external influencing factors.

10.5 Conclusions

This thesis has attempted to evaluate a gatekeeping pilot aimed at strengthening primary care functioning, and studied the history of the balance between primary care providers and hospitals in China, in order to contribute to policy development in China as well as other similar settings. The thesis has developed and used a comprehensive system of metrics to analyse the trend of balance between hospitals and primary care providers in China from 1949 to 2013; developed and used a historical-institutionalist (path-dependence) analytical framework to investigate the coevolution of hospitals and primary care providers from 1835 to 2013; conducted a difference-in-differences analysis of the effectiveness of a nationally pioneering gatekeeping pilot in shifting care from hospitals to primary care settings; and developed and used a qualitative systems analysis approach to study the factors behind the functioning of primary care gatekeeping. Through using these methods, a secondary aim of the thesis has been to explore how mixed methods can be used to study a complex system-related issue.

Ambulatory visits to primary care facilities increased substantially (an average increase of 38.7%), because of the gatekeeping pilot programme, while change in their spending pattern was less obvious. A large proportion of patients visited primary care

simply to get a referral, after the implementation of the pilot policy, leading to dissatisfaction of both doctors and patients. Expenditures at extra-district tertiary hospitals decreased substantially, apparently ceding patient expenditures to a local tertiary hospital within the study district.

There were three issues behind the sub-optimal functioning of gatekeeping. First, the conditions of primary care providers were very weak and situated in increasingly hospital-centric dynamics. Second, a series of vicious cycles continued to trap primary care development, in terms of weakening service capacity of primary care, decreasing patient trust in primary care and questionable sustainability of human resources for primary care. Third, conflicts among different priorities required of primary care development gave rise to unintended consequences, which further hampered the effectiveness of gatekeeping.

The dynamics between hospitals and primary care providers in the pilot were the contemporary manifestation of a long-term hospital-centric structure, which was path-dependent. The bias of balance of health delivery towards hospitals was continuous during the period from 1949, though the rapidly increasing dominance of hospitals was more recent. The historical coevolution of hospitals and primary care providers witnessed three cycles, each of which started with an initial preference for hospitals that triggered both reinforcing and reactive processes, eventually leading to a dominance of hospitals in the four institutional pillars underpinning the dualist structure (professional, financial, governance-related and organizational). Contemporary interventions are unlikely to have substantial effects, unless these historical dynamics are addressed and newly emerging contemporary opportunities seized upon.

The study adds weight to the arguments for methodological innovation in health policy and systems research. The thesis has demonstrated three interrelated dimensions

of methodological innovation to address the complexity involved in a health policy and systems issue by adding systems thinking into evaluation. First, evaluation can be strengthened “horizontally” along the line of the increasingly seen use of systems theory. Second, evaluation can be expanded “vertically” along the less trodden path of historical studies. Third, evaluation can be strengthened “intersectionally” via combining both historical and contemporary components in studies. The study suggests that knowledge about how contemporary challenges in health systems are historically grounded enhances researchers’ ability to critically assess contemporary issues and identify policy implications in a constantly changing world.

The thesis has made a number of policy suggestions. It reaffirms the agenda of attending to the current balance between hospitals and primary care providers in China while highlighting the need to focus on long-term underlying dynamics and institutions. The findings of the study suggest the importance of building a strong primary care profession and the value of a political coalition for primary care strengthening. Contemporary interventions, such as gatekeeping, are unlikely to have substantial effects, unless these historical dynamics are addressed and newly emerging contemporary opportunities seized upon.

The thesis has suggested some areas for further research, including more detailed analysis of the history related to balance between hospitals and primary care providers; more empirical study of other ways of shifting balance between hospitals and primary care providers; and continued methodological innovation including the use of historical approaches in contemporary evaluation of health system interventions or changes.

References

- Abel-Smith, Brian. (1994). *An Introduction to Health: Policy, Planning, and Financing*: Longman.
- Adam, T., & de Savigny, D. (2012). Systems thinking for strengthening health systems in LMICs: need for a paradigm shift. *Health Policy Plan*, 27.
- Adam, Taghreed. (2014). Advancing the application of systems thinking in health. *Health Research Policy and Systems*, 12(1), 1-5. doi: 10.1186/1478-4505-12-50
- Alliance for Health Policy and Systems Research. (2009). *Systems thinking for health systems strengthening* (D. d. Savigny; & T. Adam Eds.): World Health Organization.
- Anand, Sudhir, Fan, Victoria Y, Zhang, Junhua, Zhang, Lingling, Ke, Yang, Dong, Zhe, & Chen, Lincoln C. (2008). China's human resources for health: quantity, quality, and distribution. *The Lancet*, 372(9651), 1774-1781.
- Andrews, Bridie. (2014). *The Making of Modern Chinese Medicine, 1850-1960*: UBC Press.
- Anonymous. (1928). The Social Status of Hospitals and Their Relationships with Practicing Doctors. *Weekly Magazine of Medicine (Yixue Zhoukan)*, 4(4).
- Anonymous. (1977, 21 October 1977). Gaohao daxue zhaosheng shi quanguo renmin de xiwang (搞好大学招生是全国人民的希望) [Carrying out well university admission is the hope of the whole nation], *Renmin Ribao [People's Daily]*.
- Anonymous. (circa 200 B.C.). The Conveyance of Rites (*Liyun*) *The Book of Rites (Liji)*.
- Arthur, W Brian. (1989). Competing technologies, increasing returns, and lock-in by historical events. *The economic journal*, 99(394), 116-131.
- Arthur, W Brian. (1994). *Increasing returns and path dependence in the economy*: University of Michigan Press.
- Atkinson, Paul, & Coffey, Amanda. (2004). Analysing documentary realities. *Qualitative research*, 56-75.
- Atu, Rifat. (2004). *What Are the Advantages and Disadvantages of Restructuring a Health Care System to Be More Focused on Primary Care Services?*
- Austin, A. (1986). *Saving China: Canadian Missionaries in the Middle Kingdom*: University of Toronto Press.

- Balabanova, Dina, McKee, Martin, & Mills, Anne. (2011). *'Good Health at Low Cost' 25 Years on: What Makes a Successful Health System?* : London School of Hygiene & Tropical Medicine.
- Balazs, Etienne, Wright, Arthur F, & Wright, Hamilton Mercer. (1966). *Chinese civilization and bureaucracy: Variations on a theme* (Vol. 200): Yale University Press.
- Balme, Harold. (1921). *China and Modern Medicine*. London.
- Battistella, R. M., & Chester, T. E. (1973). *Reorganization of the National Health Service: background and issues in England's quest for a comprehensive integrated planning and delivery system*: MILBANK MEM.FD.QUART. 51 (4) (pp 489-530), 1973. Date of Publication: 1973.
- Bazeley, Pat. (2004). Issues in mixing qualitative and quantitative approaches to research. *Applying qualitative methods to marketing management research*, 141-156.
- Beijing Municipal Health Bureau. (1955). *Summative Report of Health Work 1955*.
- Beijing Municipal Health Bureau. (1973a). *Beijingshi 1972 nian weisheng gongzuo fazhan qingkuang he 1973 nian gongzuo yaodian* (北京市 1972 年卫生工作发展情况和 1973 年工作要点) [*Development of Health Work in Beijing in 1972 and Key Points for Work in 1973*].
- Beijing Municipal Health Bureau. (1973b). *Benju guanyu jiaqiang nongcun weisheng gongzuo, nongcun yiliaodui jiti ren yuan zhuanwei guojia zhigong deng gongzuo de qingshi, baogao* (本局关于加强农村卫生工作、农村医疗队集体人员转为国家职工等工作的请示报告)[*The Bureau's Proposal and Report on Reinforcing Rural Health Work and Converting Collective Members of Rural Medical Team into State Employees*].
- Berridge, Virginia. (2011). Using history in policy and practice. In V. Berridge (Ed.), *Public health in history*: McGraw-Hill Education (UK).
- Bhattacharyya, Onil, Yin, Delu, Wong, Sabrina T, & Chen, Bowen. (2011). Evolution of primary care in China 1997 – 2009. *Health Policy*, 100(2-3), 174-180. doi: 10.1016/j.healthpol.2010.11.005
- Biro, A. (2013). *Copayments, gatekeeping, and the utilization of outpatient public and private care at age 50 and above in Europe*: Health Policy. 111 (1) (pp 24-33), 2013. Date of Publication: 2013.
- Bjornsson, S., Sigurdsson, J. A., Svavarsdottir, A. E., & Gudmundsson, G. H. (2010). Gatekeeping and referrals from GPs to cardiologists: patients' opinions and registration of information flow. [Icelandic].
- Bloom, G., & Standing, H. (2008). Future health systems: Why future? Why now? *Soc Sci Med*, 66(10), 2067-2075. doi: 10.1016/j.socscimed.2008.01.032

- Bloom, G., & Wolcott, S. (2013). Building institutions for health and health systems in contexts of rapid change. *Soc Sci Med*, *96*, 216-222. doi: 10.1016/j.socscimed.2012.12.014
- Bloom, Gerald. (1985). Two Models for Change in the Health Services in Zimbabwe. *International Journal of Health Services*, *15*(3), 451-468. doi: 10.2190/kv70-akeg-y1je-klne
- Bloom, Gerald. (2014). History, complexity and health systems research. *Social science & medicine*, *117*, 160-161.
- Bloom, Gerald, & Tang, Shenglan. (2004). *Health care transition in urban China*: Gower Publishing, Ltd.
- Borowy, Iris. (2009). Thinking Big--League of Nation's Efforts towards a Reformed National Health System in China. In I. Borowy (Ed.), *Uneasy Encounters: The Politics of Medicine and Health in China 1900-1937* (pp. 205-228). Frankfurt am Main, Germany: Peter Lang.
- Bullock, Mary Brown. (1980a). *An American Transplant: The Rockefeller Foundation and Peking Union Medical College*: Univ of California Press.
- Bullock, Mary Brown. (1980b). John Grant, Medical Bolshevik *An American Transplant: The Rockefeller Foundation and Peking Union Medical College* (pp. 280). Berkley and Los Angeles, California: University of California Press.
- Caminal, Josefina, Starfield, Barbara, Sánchez, Emília, Casanova, Carmen, & Morales, Marianela. (2004). The role of primary care in preventing ambulatory care sensitive conditions. *The European Journal of Public Health*, *14*(3), 246-251. doi: 10.1093/eurpub/14.3.246
- Cao, Wenyong. (1955, 10 September, 1955). Zhizhi gongfeiyiliao zhong de yanzhong langfei (制止公费医疗中的严重浪费) [Stopping the serious waste in publicly funded medical care], *People's Daily (renmin ribao)*.
- Capoccia, Giovanni. (2015). Critical Junctures and Institutional Change. *Advances in Comparative Historical Analysis*, 147-179.
- Capoccia, Giovanni, & Kelemen, R Daniel. (2007). The study of critical junctures: Theory, narrative, and counterfactuals in historical institutionalism. *World politics*, *59*(03), 341-369.
- Carey, Gemma, Malbon, Eleanor, Carey, Nicole, Joyce, Andrew, Crammond, Brad, & Carey, Alan. (2015). Systems science and systems thinking for public health: a systematic review of the field. *BMJ Open*, *5*(12). doi: 10.1136/bmjopen-2015-009002
- Central Committee of Communist Party of China, & State Council. (2009). *Guanyu shenhua yiyao weisheng tizhi gaige de yijian* [关于深化医药卫生体制改革的意见](Opinions on Deepening Health System Reform). Retrieved February 6, 2013

- Centre for Reviews and Dissemination. (2011). The effects of gatekeeping: a systematic review of the literature--Velasco Garrido M, Zentner A, Busse R. from <http://www.crd.york.ac.uk/CRDWeb/ShowRecord.asp?ID=12011002057>
- Ch'en, C. C. (1933). Scientific Medicine as Applied in Ting Hsien: Third Annual Report of the Rural Public Health Experiment in China. *The Milbank Memorial Fund Quarterly Bulletin*, 11(2), 97-129. doi: 10.2307/3347604
- Ch'en, C. C. (1934). Public Health in Rural Reconstruction at Ting Hsien: Fourth Annual Report of the Rural Public Health Experiment in China. *The Milbank Memorial Fund Quarterly*, 12(4), 370-378. doi: 10.2307/3347713
- Ch'en, C. C. (1937a). Some problems of medical organisation in rural China. *Chinese Medical Journal*, 1937(6), 803-814.
- Ch'en, C. C. (1937b). Ting Hsien and the Public Health Movement in China. *The Milbank Memorial Fund Quarterly*, 15(4), 380-390. doi: 10.2307/3347764
- Chan, Margaret. (2011). Historical agreements at the Sixty-fourth World Health Assembly. from http://www.who.int/dg/speeches/2011/wha_20110524/en/
- Chatterton, Jocelyn. (2010). *Protestant Medical Missionary Experience during the War in China 1937-1945: The Case of Hubei Province*. SOAS, University of London.
- Chen, Bangxian (Ed.). (1991). *History of Medicine in China (Zhongguo Yixue Shi)*: Shanghai Shudian Publishing House.
- Chen, C. C. (1935). State medicine and medical education. *Chinese Medical Journal*, 49(9), 951-954.
- Chen, C. C. (1989). *Medicine in rural China: a personal account*. London: Univ of California Press.
- Chen, Haifeng, & Xu, Shijin. (1984). *Zhongguo Yixue Baikequanshu--Shehui Yixue yu Weisheng Guanlixue [中国医学百科全书--社会医学与卫生管理学] (Chinese Medical Encyclopedia--Social Medicine and Health Management Sciences)*. Shanghai: Shanghai Science and Technology Press.
- Chen, Yuyu, & Zhou, Li-An. (2007). The long-term health and economic consequences of the 1959-1961 famine in China. *Journal of Health Economics*, 26(4), 659-681. doi: 10.1016/j.jhealeco.2006.12.006
- China Medical Commission of the Rockefeller Foundation. (1914). *Medicine in China*: The University of Chicago Press, Chicago, Illinois, U.S.A.

- The Chinese Repository*. (1834). Kraus Reprint.
- Colledge, T. R., Parker, Peter, & Bridgman, E. C. (1835). Suggestions for the formation of a Medical Missionary Society offered to the consideration of all Christian nations, more especially to the kindred nations of England and the United States of America *The Medical Missionary Society in China: Address with Minutes of Proceedings, etc., etc.* Canton, 1838.
- Commentator. (1965, 24 August 1965). *Puji nongcun yiliao weisheng gongzuo de zhongyao tujing*[普及农村医疗卫生工作的重要途径](An important approach of popularizing rural medical and health services), *People's Daily (Renmin ribao)*.
- Compiling Committee. (1984). *China's Medical Encyclopedia--social medicine and health management science (Zhongguo yixue baike quanshu--shehui yixue yu weisheng shiye guanlixue) (in Chinese)*. Shanghai: Shanghai Science and Technology Press.
- Conference Committee. (1907). *China Centenary Missionary Conference Records--Report of the Great Conference Hold at Shanghai, April 5th to May 8th, 1907*. New York: American Tract Society.
- Cooke, Molly, Irby, David M., Sullivan, William, & Ludmerer, Kenneth M. (2006). American Medical Education 100 Years after the Flexner Report. *New England Journal of Medicine*, 355(13), 1339-1344. doi: doi:10.1056/NEJMra055445
- Crawford, S. M. (2014). The importance of primary care for cancer diagnoses. *Lancet Oncology*, 15(2), 136-137.
- David, Paul A. (1994). Why are institutions the 'carriers of history'? Path dependence and the evolution of conventions, organizations and institutions. *Structural change and economic dynamics*, 5(2), 205-220.
- David, Paul A. (2001). Path dependence, its critics and the quest for 'historical economics' *Evolution and path dependence in economic ideas: Past and present* (Vol. 15, pp. 40). Cheltenham, U.K.: Edward Elgar.
- de Savigny, Donald, & Adam, Taghreed. (2009). *Systems thinking for health systems strengthening* (D. d. Savigny; & T. Adam Eds.): World Health Organization.
- Deng, Tietao, & Cheng, Zhifan. (2000). *Jindai juan (Modern times volume) (in Chinese)*: People's Medical Publishing House.
- Di Tella, Rafael, & Schargrotsky, Ernesto. (2004). Do police reduce crime? Estimates using the allocation of police forces after a terrorist attack. *The American Economic Review*, 94(1), 115-133.
- Dourgnon, P., & Naiditch, M. (2010). The preferred doctor scheme: a political reading of a French

- experiment of gate-keeping. *Health Policy*, 94(2), 129-134. doi: 10.1016/j.healthpol.2009.09.001
- Duckett, Jane. (2012). *The Chinese state's retreat from health: Policy and the politics of retrenchment*: Routledge.
- Duflo, Esther. (2001). Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. *American Economic Review*, 91(4), 795-813. doi: doi: 10.1257/aer.91.4.795
- Dusheiko, Mark, Gravelle, Hugh, Yu, Ning, & Campbell, Stephen. (2007). The impact of budgets for gatekeeping physicians on patient satisfaction: evidence from fundholding. *Journal of health economics*, 26(4), 742-762.
- Dyer, Brian R. (1936). Methods Developed at the Central Field Health Station for the Training of Sanitation Personnel. *Chinese Medical Journal*(50), 76–81.
- Editorial Office of the Red Flag Magazine. (1968). Seeing the direction of medical education reform from the growth of "barefoot doctors" in Jiangzhen Commune (*Cong Jiangzhen gongshe "chijiao yishen" de chengzhang kan yixue jiaoyu geming de fangxiang*) (in Chinese). *Red Flag (Hongqi)*.
- Ellis, Harold. (2002). *A history of surgery*: Cambridge University Press.
- Escarce, Jose J, Kapur, Kanika, Joyce, Geoffrey F, & Van Vorst, Krista A. (2001). Medical care expenditures under gatekeeper and point-of-service arrangements. *Health Services Research*, 36(6 Pt 1), 1037.
- Fairbank, John K. (1978). *The Cambridge History of China: Volume 10, Late Ch'ing 1800-1911*: Cambridge University Press.
- Fairbank, John K., & Feuerwerker, Albert. (1986). *The Cambridge History of China: Republican China 1912-1949, Part 2* (Vol. 13): Cambridge University Press.
- Fan, Xingzhun. (1951). Zhongguo yufang yixue shi (中国预防医学史) [History of preventive medicine in China]. *Yishi zazhi (医史杂志) [Journal of Medical History]*, 4(3), 144.
- Fang, Xiaoping. (2012). *Barefoot Doctors and Western Medicine in China*: University of Rochester Press.
- Ferris, T. G., Chang, Y., Blumenthal, D., & Pearson, S. D. (2001). Leaving gatekeeping behind--effects of opening access to specialists for adults in a health maintenance organization. *N Engl J Med*, 345(18), 1312-1317. doi: 10.1056/NEJMsa010097

- Ferris, T. G., Perrin, J. M., Manganello, J. A., Chang, Y., Causino, N., & Blumenthal, D. (2001). Switching to gatekeeping: changes in expenditures and utilization for children. *Pediatrics*, *108*(2), 283-290.
- Ferver, Kari, Burton, Bryan, & Jesilow, Paul. (2009). The Use of Claims Data in Healthcare Research. *Open Public Health Journal*, *2*.
- Fioretos, Orfeo, Falleti, Tulia G, & Sheingate, Adam. (2016). *The Oxford Handbook of Historical Institutionalism*: Oxford University Press.
- Flexner, Abraham. (1910). *Medical education in the United States and Canada : a report to the Carnegie foundation for the advancement of teaching*. New York City: Daniel Berkeley Updike, The Merrymount Press.
- Freidson, Eliot. (2001). *Professionalism, the third logic: on the practice of knowledge*: University of Chicago press.
- Frenk, Julio. (2009). Reinventing primary health care: the need for systems integration. *The Lancet*, *374*(9684), 170-173. doi: [http://dx.doi.org/10.1016/S0140-6736\(09\)60693-0](http://dx.doi.org/10.1016/S0140-6736(09)60693-0)
- Fu, Louis. (2011a). Medical missionaries to China: the Jesuits. *Journal of medical biography*, *19*(2), 73-79.
- Fu, Yi. (2011b). Analysis of situation and influencing factors regarding employment of medical students from Zhongshan Hospital, Fudan University. *Chinese Medical Education Journal*, *31*(2).
- Gan, Y., Li, W., Cao, S., Dong, X., Li, L., Mkandawire, N., . . . Lu, Z. (2016). *Patients' Willingness on Community Health Centers as Gatekeepers and Associated Factors in Shenzhen, China: A Cross-sectional Study*: *Medicine*. *95* (14) (pp e3261), 2016. Date of Publication: 01 Apr 2016.
- Gao, Hua. (2010). Dangdai zhongguoshi shiliao de ruogan wenti (当代中国史史料的若干问题) [Issues about historical sources in contemporary Chinese history]. In H. Gao (Ed.), *Geming niandai (革命年代) [Revolutionary Years]* (pp. 324-331). Guangzhou: Guangdong People's Publishing House.
- Gao, Xi. (2012). Between the State and the Private Sphere: Chinese State Medicine Movement 1930-1949. In L. Bu, D. H. Stapleton & K.-C. Yip (Eds.), *Science, Public Health and the State in Modern Asia*. New York: Routledge.
- Garrido, M. V., Zentner, A., & Busse, R. (2011). The effects of gatekeeping: a systematic review of the literature. *Scandinavian journal of primary health care*, *29*(1), 28-38.
- Ge, Y., & Gong, S. (2007). *Zhongguo yigai: wenti, genyuan, chulu [中国医改: 问题、根源、出路]*

- (*Chinese health care reform: problems, reasons and solutions*): Zhongguo Fazhan Chubanshe (China Development Press).
- General Office of Central Committee of Communist Party of China, & General Office of State Council. (1993). *Guanyu sheji nongmin fudan xiangmu shenhe chuli yijian de tongzhi* (关于涉及农民负担项目审核处理意见的通知) [Circular on opinions regarding approval and handling of programmes involving burden on peasants].
- Gertler, P.J., Martinez, S., Premand, P., Rawlings, L.B., & Vermeersch, C.M.J. (2011). *Impact Evaluation in Practice*. Washington DC: World Bank Publications.
- Gertman, P. M., & Restuccia, J. D. (1981). The appropriateness evaluation protocol: a technique for assessing unnecessary days of hospital care. *Med Care*, 19(8), 855-871.
- Gervas, Juan, Ferna, Merceces Pérez, & Starfield, Barbara H. (1994). Primary care, financing and gatekeeping in western Europe. *Family practice*, 11(3), 307-317.
- Gibson, Odette R, Segal, Leonie, & McDermott, Robyn A. (2013). A systematic review of evidence on the association between hospitalisation for chronic disease related ambulatory care sensitive conditions and primary health care resourcing. *BMC health services research*, 13(1), 336.
- Gilson, Lucy. (2012). *Health Policy and Systems Research: A Methodology Reader*.
- Godber, Edward, Robinson, Ray, & Steiner, Andrea. (1997). Economic evaluation and the shifting balance towards primary care: definitions, evidence and methodological issues. *Health economics*, 6(3), 275-294.
- Gould, Stephen Jay, & Eldredge, Niles. (1972). Punctuated equilibria: an alternative to phyletic gradualism.
- Government Administration Council of the Central People's Government. (1951). *Zhonghua renmin gongheguo laodong baoxian tiaoli* (中华人民共和国劳动保险条例) [People's Republic of China's Regulations on Labour Insurance].
- Government Administration Council of the Central People's Government. (1952a). *Guanyu quanguo geji renmin zhengfu, dangpai, tuanti ji suoshu shiye danwei de guojia gongzuo renyuan shixing gongfei yiliao yufang de zhishi* [关于全国各级人民政府、党派、团体及所属事业单位的国家工作人员实行公费医疗预防的指示] (Directive on Implementing Publicly Funded Medical Care and Prevention for Nationwide State Employees in All Levels of People's Government, Parties, Groups and Affiliated Public Institutions).
- Government Administration Council of the Central People's Government. (1952b). *Guojia gongzuo renyuan gongfei yiliao shishi banfa* [国家工作人员公费医疗预防实施办法] (Implementation

Measures of Publicly Funded Medical Care and Prevention for State Employees).

- Grant, John B. (1963 (1934)). Principles for medicine and public health in the China experiment. In C. Seipp (Ed.), *Health care for the community--selected papers of Dr. John B. Grant*. Baltimore: The Johns Hopkins Press.
- Green, Judith, & Thorogood, Nicki. (2004). *Qualitative methods for health research*: Sage.
- Greenfield, Geva, Foley, Kimberley, & Majeed, Azeem. (2016). Rethinking primary care's gatekeeper role.
- Greenfield, Geva, Pliskin, Joseph S, Feder-Bubis, Paula, Wientroub, Shlomo, & Davidovitch, Nadav. (2012). Patient-physician relationships in second opinion encounters-The physicians' perspective. *Social science & medicine*, 75(7), 1202-1212.
- Grmek, Mirko Dražen. (1966). *Serving the Cause of Public Health. Selected Papers of Andrija Stampar*.
- Guangxi Provincial Health Gazette Compiling Committee. (1998). *Guangxi tongzhi · Yiliao weisheng zhi* (广西通志·医疗卫生志) [*Guangxi general gazette--Health gazette*]: Guangxi People's Publishing House.
- Guo, Hong, Liang, Xiaoyun, & Jin, Chenggang. (2012). Revenue and Expenditure Separation Increases Rate of First Contact Care in Community Health Centers among Community Residents (社区卫生收支两条线改革对社区首诊的作用研究). *Chinese General Practice*, 15(31), 3578-3580. doi: 10.3969/j.issn.1007-9572.2012.31.005
- Gusmano, Michael K. (2015). China. In K. Fierlbeck; & H. A. Palley (Eds.), *Comparative Health Care Federalism*. Farnham, England: Ashgate Publishing Limited.
- Hacker, Jacob S. (1998). The historical logic of national health insurance: Structure and sequence in the development of British, Canadian, and US medical policy. *Studies in American Political Development*, 12(01), 57-130.
- Hall, Peter A, & Taylor, Rosemary CR. (1996). Political science and the three new institutionalisms*. *Political studies*, 44(5), 936-957.
- Halstead, SB, Walsh, JA, & Warren, KS. (1985). Good health at low cost: Rockefeller Foundation conference report. *New York: Rockefeller Foundation*.
- Hardiman, David. (2006). *Healing Bodies, Saving Souls: Medical Missions in Asia and Africa*: Rodopi.
- Hardin, Russell. (1989). Why a Constitution? The Federalist Papers and the New Institutionalism. Ed. Bernard Grofman, Donald Wittman: Agathon Press.

- Hayford, C.W. (1990). *To the People: James Yen and Village China*: Columbia University Press.
- He, Alex Jingwei. (2012). Is the Chinese Health Bureaucracy Incapable of Leading Healthcare Reforms?: The Case of Fujian Province. *China: An International Journal*, 10(1), 93-112.
- He, Biao. (2001). *Autobiography of He Biao (He Biao Huiyilu) (in Chinese)*. Beijing: Chinese People's Liberation Army Publishing House (*Jiefangjun Chubanshe*).
- Henderson, Gail. (1989). Issues in the modernisation of medicine in China. In D. F. Simon & M. Goldman (Eds.), *Science and Technology in Post-Mao China* (pp. 199-221). Harvard: Harvard University Press.
- Hillier, S.M., & Jewell, T. (1983). *Health Care and Traditional Medicine in China 1800-1982*. Boston: Routeledge & Kegan Paul plc.
- Ho, FC. (2012). The beginning of medical education in Hong Kong 125 years ago... its unique features in comparison with similar efforts in China. *Hong Kong medical journal= Xianggang yi xue za zhi/Hong Kong Academy of Medicine*, 18(6), 544-550.
- Honigsbaum, Frank. (1979). *The division in British medicine: A history of the separation of general practice from hospital care, 1911-1968*. London School of Economics and Political Science (University of London).
- Hou, Jianlin, Michaud, Catherine, Li, Zhihui, Dong, Zhe, Sun, Baozhi, Zhang, Junhua, . . . Wei, Bo. (2014). Transformation of the education of health professionals in China: progress and challenges. *The Lancet*, 384(9945), 819-827.
- Hu, Jintao. (2007). *Hujintao's Report at the 17th National Congress of the Chinese Communist Party (Hujintao zai zhongguo gongchandang dishiqici quanguo dabiao dahui shang de baogao) (in Chinese)*.
- Hu, Wei. (1998). *Zhengfu guocheng (政府过程) [Government process]*: Zhejiang People's Publishing House.
- Huang, Qingchang. (2013, 15 July 2013). Gelansushike (zhongguo) bufen gaoguan bei lian zhencha: jiekai kuaguo yaoqi shangye huilu liyilian (葛兰素史克(中国)部分高管被立案侦查 揭开跨国药企商业贿赂利益链) [High-level Executives of GlaxoSmithKline (China) prosecuted, revealing interest chain of commercial bribery of cross-national pharma], *Renmin Ribao [People's Daily]*. Retrieved from <http://finance.people.com.cn/n/2013/0715/c1004-22194208.html>
- Huang, Yanzhong. (2000). *The paradox of state power: Political institutions, policy process, and public health in post-Mao China*. (9990562), The University of Chicago, Ann Arbor.

- Huang, Yun. (1957). Dangqian gongfei yiliao jingfei guanli zhong de wenti he gaijin yijian (当前公费医疗管理中的问题和改进意见) [Issues and Opinions on Improvement of the Current Fund Management of Publicly Funded Medical Care]. *Caizheng [Finance]*(3), 15-17.
- Huangsongyu Commune Health Centre. (1989). Report on Integration of Village Doctors into Township Health Centres.
- Hughes, Jane, & Challis, David. (2004). Frail older people – margins of care. *Reviews in Clinical Gerontology*, 14(02), 155-164. doi: doi:10.1017/S0959259804001364
- Hume, Edward H. (1922). Medical Education in China: A Survey and Forecast. *Addresses & papers, dedication ceremonies and Medical conference, Peking Union Medical College, September 15-22, 1921* (pp. 71-93). Peking: Peking Union Medical College.
- Hume, Edward H. (1935). The place of mission hospitals and medical schools in the national health programme of China. *Chinese Medical Journal*, 49(9), 946-950.
- Hymes, R. P. (1987). Not quite gentlemen? Doctors in Sung and Yuan. *Chin Sci*, 8, 9-76.
- Imas, Linda G. Morra, & Rist, Ray C. (2009). *The road to results: Designing and conducting effective development evaluations*: World Bank Publications.
- Immergut, Ellen M. (1990). Institutions, veto points, and policy results: A comparative analysis of health care. *Journal of Public policy*, 10(04), 391-416.
- Inglis, A. L., Coast, J., Gray, S. F., Peters, T. J., & Frankel, S. J. (1995). Appropriateness of hospital utilization. The validity and reliability of the Intensity-Severity-Discharge Review System in a United Kingdom acute hospital setting. *Med Care*, 33(9), 952-957.
- Institute of Medicine. (1978). A Manpower Policy for Primary Health Care: Report of a Study. Washington, D.C.
- Institute of Medicine. (1996). *Primary care: America's health in a new era*: National Academies Press.
- Jiang, Zhongzheng. (1943). Zhongguo zhi mingyun (中国之命运) [China's destiny].
- Jiangxi Provincial Health Gazette Compiling Committee. (1997). *Jiangxi Provincial Health Gazette*. Hefei, Jiangxi: Huangshan Shushe.
- Jiangxi Provincial Health Gazette Compiling Committee. (1997). *Jiangxi Provincial Health Gazette*. Hefei, Jiangxi: Huangshan Shushe.
- Jin, Baoshan. (1946). *Sanshi nian lai zhongguo gonggong weishegn zhi huigui yu qianzhan* [三十年来

- 中国公共卫生之回顾与前瞻] (Public health in China over the past 30 years and its prospects). *National Medical Journal of China*, 32(1).
- Johnston, Lucy, Lardner, Clare, & Jepson, Ruth. (2008). *Overview of evidence relating to shifting the balance of care: a contribution to the knowledge base*: Scottish Government Social Research.
- Joskow, Paul L. (1988). Asset specificity and the structure of vertical relationships: empirical evidence. *Journal of Law, Economics, & Organization*, 4(1), 95-117.
- Kmietowicz, Zosia. (2006). A century of general practice. *BMJ : British Medical Journal*, 332(7532), 39-40.
- Kohn, Robert, & White, Kerr L. (1976). *Health care: an international study*: Oxford University Press, 37 Dover Street, London W1X 4AH.
- Krasner, Stephen D, Nordlinger, Eric, Geertz, Clifford, Skowronek, Stephen, Tilly, Charles, Grew, Raymond, & Trimberger, Ellen Kay. (1984). Approaches to the state: Alternative conceptions and historical dynamics: JSTOR.
- Kringos, D. S., Boerma, W. G., Hutchinson, A., van der Zee, J., & Groenewegen, P. P. (2010). The breadth of primary care: a systematic literature review of its core dimensions. *BMC Health Serv Res*, 10, 65. doi: 10.1186/1472-6963-10-65
- Kristensen, Søren Rud, Meacock, Rachel, Turner, Alex J., Boaden, Ruth, McDonald, Ruth, Roland, Martin, & Sutton, Matthew. (2014). Long-Term Effect of Hospital Pay for Performance on Mortality in England. *New England Journal of Medicine*, 371(6), 540-548. doi: doi:10.1056/NEJMoa1400962
- Kuhn, Philip A. (2002). *Origins of the modern Chinese state*: Stanford University Press.
- Lampton, D.M. (1977). *The Politics of Medicine in China: The Policy Process, 1949-1977*: Westview Press.
- Lauridsen, Sigurd. (2009). Administrative gatekeeping—a third way between unrestricted patient advocacy and bedside rationing. *Bioethics*, 23(5), 311-320.
- Lavis, John N, & Anderson, Geoffrey M. (1996). Appropriateness in health care delivery: definitions, measurement and policy implications. *CMAJ: Canadian Medical Association Journal*, 154(3), 321.
- Lazich, Michael C. (2006). Seeking Souls through the Eyes of the Blinds: The Birth of the Medical Missionary Society in Nineteenth-Century China. In D. Hardiman (Ed.), *Healing Bodies, Saving Souls: Medical Missions in Asia and Africa* (pp. p59-86). Amsterdam; New York: Rodopi.

- League of Nations Health Organization. (1937). Report of the Intergovernmental Conference of Far-Eastern Countries on Rural Hygiene (held in Bandoeng, Java). Geneva, Switzerland.
- Lei, Haichao, Bian, Ying, Liu, Xingzhu, & Li, Shixue. (1996). *Sanji zonghe yiyuan menzhen yu zhuyuan huanzhe de ke fenliu xing diaocha ji jingjixue yiyi* [三级综合医院门诊与住院患者的可分流动性调查及经济学意义] (Study on of the feasibility to shift of outpatients and inpatients of tertiary comprehensive hospitals to other facilities and economic significance). *Weisheng ruankexue* [卫生软科学] (*Soft sciences for health*), 3, 34-36.
- Lei, Sean Hsiang-lin. (2014). *Neither donkey nor horse: medicine in the struggle over China's modernity*: University of Chicago Press.
- Leung, Angela Ki Che. (1987). Organized medicine in Ming-Qing China: State and private medical institutions in the lower Yangzi region. *Late Imperial China*, 8(1), 134-166.
- Levi, Margaret. (1997). A model, a method, and a map: Rational choice in comparative and historical analysis. *Comparative politics: Rationality, culture, and structure*, 28, 78.
- Li, Jingwei. (1990). *Xixue dongjian yu zhongguo jindai yixue sichao* (西学东渐与中国近代医学思潮) [*Influence of Western studies on the east, and thoughts on medicine in modern China*]. Wuhan: Hubei Science and Technology Publishing Co., Ltd.
- Li, Mubai, & Jin, Xuechen. (1964). *Tiantou weishengyuan* (田头卫生员) [The health worker at the head of the field]: Shanghai renmin meishu chubanshe (上海人民美术出版社).
- Li, Shixue. (1997). The problems of hospital reimbursement and related suggestions. Jinan.
- Li, Shiyun, & Li, Qin. (2008). *Woguo shequ shouzhen he shuangxiang zhuanzhen de wenxian fenxi* [我国社区首诊和双向转诊研究的文献分析] (Analysis of the literature on first-contact care and two-way referrals in community health facilities). *Zhongguo quanke yixue* [中国全科医学] (*Chinese General Medicine*), 11(10), 1738-1738.
- Liang, Qizi. (2001). *Shishan yu jiaohua* (施善与教化: 明清的慈善组织) [*Almsgiving and culturalizing: benevolent societies in Ming and Qing*]: Hebei Education Publishing House.
- Liang, Qizi. (2012). *Miandui jibing: chuantong Zhongguo shehui de yiliao guannian yu zuzhi* (面对疾病: 传统中国社会的医疗观念与组织) [*Facing diseases: medical notions and organisations in traditional Chinese society*]: China Renmin University Press.
- Liang, Qizi. (2013). *Jindai zhongguo yiyuan de dansheng* (近代中国医院的诞生) [The birth of modern Chinese hospitals]. In Y. Zhu (Ed.), *Jiankang yu shehui: huaren weisheng xinshi* (健康與社會: 華人衛生新史) [*Health and society: new history of Chinese health*] (pp. 41-68). Taiwan: Linking Publishing.

- Lieberthal, K. (2004). *Governing China: From Revolution Through Reform*: W. W. Norton.
- Lim, R. K. S., & Chen, C. C. (1937). State medicine. *Chinese Medical Journal*, 51(6), 781-796.
- Lipset, Seymour M, & Rokkan, Stein. (1967). Cleavage structures, party systems, and voter alignments: an introduction.
- Litsios, Socrates. (2015). On the Origins of Primary Health Care. In A. Medcalf, S. Bhattacharya, H. Momen, M. Saavedra & M. Jones (Eds.), *Health for All--The Journey to Universal Health Coverage*. Hyderabad: Orient Blackswan Pvt. Ltd.
- Liu, Hongqing. (2009). Gongfei laobao yiliao--jianxing jianyuan de jiyi [公费劳保医疗--渐行渐远的记忆] (Government Insurance medical Scheme and Labour Insurance medical Scheme--increasingly distant memories). *Zhongguo Shehui Baozhang [Chinese Social Security]*, 11(1), 121.
- Liu, J. Heng. (1936). *China Year Book 1935-1936*.
- Liu, Ruiqing. (1957). Gongfei yiliao zenyang cong chaozhi bianwei youyu (公费医疗怎样从超值变为有余) [How to turn deficit of publicly funded medical care into surplus]. *Caizheng [Finance]*(10), 32.
- Liu, X., Liu, Y., & Chen, N. (2000). The Chinese experience of hospital price regulation. *Health Policy Plan*, 15(2), 157-163.
- Liu, Xingzhu, & Cao, Huaijie. (1992). China's Cooperative Medical System: Its Historical Transformations and the Trend of Development. *Journal of Public Health Policy*, 13(4), 501 - 511.
- Liu, Xingzhu, & Mills, Anne. (2005). The effect of performance-related pay of hospital doctors on hospital behaviour: a case study from Shandong, China. *Human Resources for Health*, 3(11).
- Liu, Xingzhu, Xu, Lingzhong, & Wang, Shuhong. (1996). Reforming China's 50 000 township hospitals—effectiveness, challenges and opportunities. *Health Policy*, 38(1), 13-29.
- Liu, Yuanli, & Rao, Keqin. (2006). Providing Health Insurance in Rural China: From Research to Policy. *Journal of Health Politics, Policy and Law*, 31(1), 71-92. doi: 10.1215/03616878-31-1-71
- Lockhart, William. (1861). *The medical missionary in China: A narrative of twenty years' experience*: Hurst and Blackett.
- Lord Dawson. (1920). Interim Report on the Future Provision of Medical and Allied Services 1920. <http://www.sochealth.co.uk/healthcare-generally/history-of-healthcare/interim-report-on-the-future-provision-of-medical-and-allied-services-1920-lord-dawson-of-penn/>

- Lucas, AnElissa. (1980). Changing medical models in China: organizational options or obstacles? *China Quarterly*, 461-489.
- Luk, Sabrina. (2014). *Health Insurance Reforms in Asia*: Taylor & Francis.
- MacFarquhar, Roderick, & Fairbank, JK. (1987). *The Cambridge History of China: Volume 14, The People's Republic, Part 1, The Emergence of Revolutionary China, 1949-1965*. New York: Cambridge University Press.
- MacFarquhar, Roderick, & Fairbank, JK. (1991). *The Cambridge History of China: The People's Republic, Pt. 2: Revolutions Within the Chinese Revolution, 1966-1982*: New York: Cambridge University Press.
- Macinko, James, Starfield, Barbara, & Erinosh, Temitope. (2009). The Impact of Primary Healthcare on Population Health in Low-and Middle-Income Countries. *The Journal of ambulatory care management*, 32(2), 150-171.
- Macinko, James, Starfield, Barbara, & Shi, Leiyu. (2003). The Contribution of Primary Care Systems to Health Outcomes within Organization for Economic Cooperation and Development (OECD) Countries, 1970-1998. *Health Services Research*, 38(3), 831-865. doi: 10.1111/1475-6773.00149
- Mackie, J. L. (1965). Causes and Conditions. *American Philosophical Quarterly*, 2(4), 245-264.
- Mahoney, James. (2000). Path dependence in historical sociology. *Theory and society*, 29(4), 507-548.
- Mao, Zedong. (1949). Zai zhongguo gongchandang diqijie zhongyang weiyuanhui di'erci quanti huiyi shang de baogao (在中国共产党第七届中央委员会报告第二次全体会议上的报告) [Report to the Second Plenary of the Seventh Central Committee of the Chinese Communist Party].
- Mao, Zedong. (1991). *Mao Zedong xuanji (毛泽东选集) [Selected Works of Mao Zedong]*: People's Publishing House.
- Matin, W. A. P. (1896). *A Cycle of Cathay or China, South and North, with Personal Reminiscences*: New York, Chicago, Toronto: Fleming H. Revell ComPany.
- Matthews-King, Alex. (2016, 2016 Jan 8). GPs continue to be offered "questionable" incentives to reduce cancer referrals. from <http://www.pulsetoday.co.uk/news/commissioning/commissioning-topics/referrals/gps-continue-to-be-offered-questionable-incentives-to-reduce-cancer-referrals/20030659.fullarticle>

- McDonald, A. G., Cudderford, G. C., & Beale, E. M. L. (1974). Balance of care: some mathematical models of the National Health Service. *British Medical Bulletin*, 30(3), 262-270.
- McMillan, A. W., Gornick, M. E., Rogers, R. R., & Gotten, M. K. (1970). *Assessing the balance of physician manpower in a metropolitan area*: Public Health Reports. 85 (1) (pp 1001-1011), 1970. Date of Publication: 1970.
- McPake, Barbara, & Hanson, Kara. (2016). Managing the public–private mix to achieve universal health coverage. *The Lancet*, 388(10044), 622-630.
- Medical Research Council. (2008). Developing and evaluating complex interventions: new guidance.
- Meng, Qingyue, Fang, Hai, Liu, Xiaoyun, Yuan, Beibei, & Xu, Jin. (2015). Consolidating the social health insurance schemes in China: towards an equitable and efficient health system. *The Lancet*, 386(10002), 1484-1492.
- Meng, Qingyue, Sun, Xiaojie, Li, Hui, Song, Kuimeng, Yuan, Jing, Wang, Haipeng, & Chang, Jie. (2011). Functions and Human Resources of China's Primary Health Care System in the Early Stage of Health Systems Reform. Beijing: Peking University China Center for Health Development Studies.
- Meng, Qingyue, & Xu, Ke. (2014a). Progress and challenges of the rural cooperative medical scheme in China. *Bulletin of the World Health Organization*, 92(6), 447-451.
- Meng, Qingyue, & Xu, Ke. (2014b). Progress and challenges of the rural cooperative medical scheme in China. *Bull World Health Organ*, 92(6), 447-451.
- Meng, Qun, Xu, Ling, Zhang, Yaoguang, Qian, Juncheng, Cai, Min, Xin, Ying, . . . Barber, Sarah L. (2012). Trends in access to health services and financial protection in China between 2003 and 2011: a cross-sectional study. *The Lancet*, 379(9818), 805-814.
- Mills, Anne. (1990). The economics of hospitals in developing countries. Part I: expenditure patterns. *Health policy and planning*, 5(2), 107-117.
- Mills, Anne. (2012). Health policy and systems research: defining the terrain; identifying the methods. *Health Policy and Planning*, 27(1), 1-7. doi: 10.1093/heapol/czr006
- Ministry of Health. (1950). Directives on Medical Administration Work This Year. *Shandong Policy Paper (Shandong zhengbao)*(05), 81.
- Ministry of Health. (1951a). Detailed Implementation Plan of Work on Rural Health Organization (Draft) (*Nongcun weisheng jiceng zuzhi gongzuo juti shishi banfa (cao'an)*). In K. L. Changming Li, Zhaoyang Zhang, Chunlei Nie, Wei Fu, Hongming Zhu, Bin Wang (Ed.), *Collection of Rural*

- Health Policy Documents (1951-2000)* (pp. 241-251): Department Grassroot Health and Marternal and Child Health, Ministry of Health.
- Ministry of Health. (1951b). *Zhongyang renmin zhengfu weishengbu gongbu quanguo weishegn dahui guanyu yiyao jiaoyu deng sixiang jue ding* [中央人民政府卫生部公布全国卫生大会关于医药教育等四项决定] (Central People's Government Issued Four Decisions of the National Health Congress Regarding Medicines, Education, etc.). *Traditional Chinese Medicine Journal*(01), 1-5.
- Ministry of Health. (1956a). *Guanche gaibian xian weishegn yuan zuzhi jigou jiaqiang fangyi gongzuo de tongzhi* (关于贯彻改变县卫生院组织机构, 加强防疫工作的通知) [Circular on implementing the change of organisational institutions of county health centres to reinforce anti-epidemic work].
- Ministry of Health. (1956b). *Guojia weisheng jishu ren yuan zhiwu mingcheng he zhiwu jinsheng tiaoli* (国家卫生技术人员职务名称和职务晋升条例) [National regulations for professional titles and professional promotions of health professionals].
- Ministry of Health. (1957a). Directive on Strengthening Leadership of Grassroot Health Organisations. In K. L. Changming Li, Zhaoyang Zhang, Chunlei Nie, Wei Fu, Hongming Zhu, Bin Wang (Ed.), *Collection of Rural Health Policy Documents (1951-2000)*: Department Grassroot Health and Marternal and Child Health, Ministry of Health.
- Ministry of Health. (1957b). *Guanyu gaijin huaqu yiliao fuwu gongzuo de zhishi* [关于改进划区医疗服务工作的指示] (Directives on improving sectional medical services).
- Ministry of Health. (1959). *Guanyu renmin gongshe weisheng gongzuo jige wenti de yijian* (关于人民公社卫生工作几个问题的意见) [Opinions on several issues regarding health work in people's communes].
- Ministry of Health. (1960). *Guanyu quanguo nongcun weisheng gongzuo Shanxi Qishan xianchang huiyi qingkuang de baogao* (关于全国农村卫生工作山西稷山现场会议情况的报告) [Report on the National Rural Health Field Conference in Qishan, Shanxi].
- Ministry of Health. (1962). *Opinions on improving several issues related to hospital work (Guanyu gaijin yiyuan gongzuo ruogan wenti de yijian)* (in Chinese).
- Ministry of Health. (1965). *Report on putting the stress of health work in rural areas (Guanyu ba weisheng gongzuo zhongdian fangdao nongcun de baogao)* (in Chinese).
- Ministry of Health. (1974). *1973 nian quanguo jiaoyu weishegn he xingzheng zuotanhui fujian san ·Guanyu weisheng shiye jihua caiwu gongzuo zhong ruogan wenti de yijian (zhaiyao)* (1973 年全国教育卫生和行政座谈会附件三·关于卫生事业计划财务工作中若干问题的的

意见 (摘录) [1973 National Administrative Group Discussion on Education and Health, third attachment--Opinions regarding issues in health affairs planning and financial work].

Ministry of Health. (1980a). *Guanyu gaohao sanfen zhiyi zuoyou xian de weisheng shiye zhengdun jianshe de jidian yijian* (关于搞好三分之一左右县的卫生事业整顿建设的几点意见)[Opinions regarding carrying out rectification and construction of health affairs in about one third of counties].

Ministry of Health. (1980b). *Opinions regarding strengthening work of county hospitals (Guanyu jiaqiang xianyiyuan gongzuo de jidian yijian)* (in Chinese).

Ministry of Health. (1985a). *China Health Yearbook 1985*. Beijing: People's Medical Publishing House.

Ministry of Health. (1985b). *Weishengbu guanyu kaizhan weisheng gaige zhong xuyao huaqing de jitiao zhengce jixian* (卫生部关于开展卫生改革中需要划清的几条政策界限) [On Several Policy Boundaries that Need Clarification in Implementing Health Reform].

Ministry of Health. (1986a). *China Health Yearbook 1986*. Beijing: People's Medical Publishing House.

Ministry of Health. (1986b). *Weisheng jishu ren yuan zhiwu shixing tiaoli* (卫生技术人员职务试行条例) [Interim regulations on professional posts of health professionals].

Ministry of Health. (1987). *China Health Yearbook 1987*. Beijing: People's Medical Publishing House.

Ministry of Health. (1988). *China Health Yearbook 1988*. Beijing: People's Medical Publishing House.

Ministry of Health. (1989). *Jianguo sishinian--quanguo weisheng tongji ziliao* (建国四十年--全国卫生统计资料) [Health Statistics Information in China 1949-1988].

Ministry of Health. (1990a). *China Health Yearbook 1990*. People's Medical Publishing House.

Ministry of Health. (1990b). *The Eighth Five-year Plan for Health Work and Planning Proposal for 2000* (*Weisheng shiye dibage wunian jihua ji erlinglingling nian gui hua shexiang*).

Ministry of Health. (1991). *China Health Yearbook 1991*. Beijing: People's Medical Publishing House.

Ministry of Health. (1992). *China Health Yearbook 1992*. Beijing: People's Medical Publishing House.

Ministry of Health. (1993). *China Health Yearbook 1993*. Beijing: People's Medical Publishing House.

Ministry of Health. (1994a). *China Health Yearbook 1994*. Beijing: People's Medical Publishing House.

Ministry of Health. (1994b). *Yiliao jigou guanli tiaoli shishi xize* (医疗机构管理条例实施细则)

[Implementation details on regulations for medical facilities]. Retrieved from <http://www.moh.gov.cn/mohzcfgs/pgz/200804/18303.shtml>.

- Ministry of Health. (1995). *China Health Yearbook 1995*. Beijing: People's Medical Publishing House.
- Ministry of Health. (1996). *China Health Yearbook 1996*. Beijing: People's Medical Publishing House.
- Ministry of Health. (1997). *China Health Yearbook 1997*. Beijing: People's Medical Publishing House.
- Ministry of Health. (1998). *China Health Yearbook 1998*. Beijing: People's Medical Publishing House.
- Ministry of Health. (1999). *China Health Yearbook 1999*. Beijing: People's Medical Publishing House.
- Ministry of Health. (2000). *China Health Yearbook 2000*. Beijing: People's Medical Publishing House.
- Ministry of Health. (2001). *China Health Yearbook 2001*. Beijing: People's Medical Publishing House.
- Ministry of Health. (2002). *China Health Yearbook 2002*. Beijing: People's Medical Publishing House.
- Ministry of Health. (2003). *China Health Statistical Yearbook 2003*. Beijing: Peking Union Medical College Publishing House.
- Ministry of Health. (2004). *China Health Statistical Yearbook 2004*. Beijing: Peking Union Medical College Publishing House.
- Ministry of Health. (2005). *China Health Statistical Yearbook 2005*. Beijing: Peking Union Medical College Publishing House.
- Ministry of Health. (2006). *China Health Statistical Yearbook 2006*. Beijing: Peking Union Medical College Publishing House.
- Ministry of Health. (2007). *China Health Statistical Yearbook 2007*. Beijing: Peking Union Medical College Publishing House.
- Ministry of Health. (2008). *China Health Statistical Yearbook 2008*. Beijing: Peking Union Medical College Publishing House.
- Ministry of Health. (2009). *China Health Statistical Yearbook 2009*. Beijing: Peking Union Medical College Publishing House.
- Ministry of Health. (2010). *China Health Statistical Yearbook 2010*. Beijing: Peking Union Medical College Publishing House.

- Ministry of Health. (2011). *China Health Statistical Yearbook 2011*. Beijing: Peking Union Medical College Publishing House.
- Ministry of Health. (2012). *China Health Statistics Yearbook 2012*.
- Ministry of Health, & Ministry of Finance. (1960). Joint circular regarding that expenditures on hospital staff wages becomes full-budgeted by the state (*Guanyu yiyuan gongzuo renyuán de gongzi quanbu you guojia yusuan kaizhi de lianhe tongzhi*) (in Chinese).
- Ministry of Health, Ministry of Finance, & State Administration of Traditional Chinese Medicine. (2007). *Guiding opinions regarding perfecting measures of pooling and reimbursement of the New Rural Cooperative Medical Scheme (Guanyu wanshan xinxing nongcun hezuoyiliao tongchou buchang fang'an de zhidao yijian)* (in Chinese).
- Ministry of Health, National Development and Reform Commission, Ministry of Industry and Information Technology, Ministry of Supervision, Ministry of Finance, Ministry of Human Resources and Social Security, . . . State Food and Drug Administration. (2009). *Guanyu yinfa guanyu jianli guojia jiben yaowu zhidu de shishi yijian de tongzhi* [关于印发《关于建立国家基本药物制度的实施意见》的通知] (Circular regarding the implementation opinions regarding establishing the National Essential Medicines System).
- Ministry of Human Resources and Social Security, Ministry of Finance, Ministry of Health. (2009). *Guanyu yinfa gonggong weisheng yu jiceng yiliao weisheng shiye danwei shishi jixiao gongzhi de zhidao yijian de tongzhi* [关于印发公共卫生与基层医疗卫生事业单位实施绩效工资的指导意的通知] (Circular regarding issuing of guiding opinions regarding implementing performance-based salary in public institutions on public health and primary care).
- Ministry of Health, Ministry of Finance, & State General Administration of Labour. (1979). *Guanyu jiaqiang yiyuan jingji guanli shidian gongzuo de tongzhi* (关于加强医院经济管理试点工作的通知) [Circular on Strengthening Hospital Economic Management Piloting Work].
- Mooney, G. H., & Drummond, M. F. (1982). Essentials of health economics: Part III (continued)--developing health care policies. *Br Med J (Clin Res Ed)*, 285(6351), 1329-1331.
- Mooney, G.H., Russell, E.M., & Weir, R.D. (1986). *Choices for Health Care: A Practical Introduction to the Economics of Health Provision*: Macmillan.
- Mugisha, F., & Nabyonga-Orem, J. (2010). *To what extent does recurrent government health expenditure in Uganda reflect its policy priorities? : Cost Effectiveness and Resource Allocation*. 8 (no pagination), 2010. Article Number: 19. Date of Publication: 20 Oct 2010.
- Muldoon, L. K., Hogg, W. E., & Levitt, M. (2006). Primary care (PC) and primary health care (PHC). What is the difference? *Can J Public Health*, 97(5), 409-411.

- Nabyonga, J, Munguti, N, & Thomas, S. (2002). National Health Accounts (NHA) in East and Southern Africa (ESA): a comparative analysis. *Cape Town: East and Southern African National Health Accounts Network*.
- National Bureau of Statistics of China. (2015). *China Statistical Yearbook 2015*. Beijing: China Statistics Press.
- National Congress (Cartographer). (1947). *Zhonghuaminguo xianfa* [中华民国宪法] (Constitution of the Republic of China).
- National Health and Family Planning Commission. (2013). *China Health and Family Planning Statistical Yearbook 2013*.
- National Health and Family Planning Commission. (2014). *China Health and Family Planning Statistical Yearbook 2014* (P. U. M. C. Press Ed.). Beijing.
- Newell, Kenneth W. (1975). *Health By The People*. Geneva.
- Oldt, F. (1937). State medicine problems. *Chinese Medical Journal*, 51(6), 797-802.
- Organization for Economic Co-operation and Development. (2013). Organisation of health care delivery--OECD Health system characteristics survey 2012. Retrieved 11 May 2017, from <http://www.oecd.org/els/health-systems/organisation-health-care-delivery.htm>
- Ose, D, Broge, B, Riens, B, & Szecsenyi, J. (2008). Contacts to Specialists with Referrals by GP-Have GP Centred Health Care (HZV) Contracts an Impact? *Zeitschrift fur Allgemeinmedizin*, 84(08), 321-326.
- Ouyang, Jing. (1984). *Huiyi shanganning bianqu de weisheng gongzuo (xia)*[回忆陕甘宁边区的卫生工作] (Recollections on health work in Shaanxi-Gansu-Ningxia border areas (2nd part)). *Hospital Management (Yiyuan guanli)*, 47(2), 57-63.
- Paina, Ligia, Bennett, Sara, Ssengooba, Freddie, & Peters, David H. (2014). Advancing the application of systems thinking in health: exploring dual practice and its management in Kampala, Uganda. *Health Research Policy and Systems*, 12(1), 41. doi: 10.1186/1478-4505-12-41
- Parker, Peter. (1836). Second Quarterly Report *The Chinese Repository 1836* (pp. 34).
- Parker, Peter. (1842). *Statements Respecting Hospitals in China--preceded by a Letter to John Abercrombie, M.D., V.P.R.S.E.*, . Glasgow: Medical Missionary of the American Board of Foreign Missions in China.
- Party Branch of Heidouyu Brigade at Huangsongyu Commune. (1977). *Jianchi sixiang zhengdun*,

jianshe geminghua weishengyuan (坚持思想整顿 建设革命化卫生院) [holding on to thought correction, building revolutionalized a health centre]

Pawson, Ray. (2013). *The science of evaluation: a realist manifesto*: Sage.

People's Daily. (1951, 3 November 1951). *Quanguo tuixing xin yixue jiaoyu zhidu huode chengji, jinnian quanguo yixuexiao jixuezhe chaoguo yiwang 69 nian suo xunlian de yisheng zongshu* [全国推行新医学教育制度获得成绩 今年全国医学校就学者超过已往六十九年所训练的医生总数] (Nationwide Implementation of New Medical Educaiotn System Achieve Good Results, Students in Medical Schools Across Nation Has Exceeded Total Number of Doctor Trained over Past 69 Years), *People's Daily*

Perneger, T. V., Etter, J. F., & Rougemont, A. (1996). Switching Swiss enrollees from indemnity health insurance to managed care: the effect on health status and stisfaction with care. *Am J Public Health, 86*(3), 388-393.

Peters, David H. (2014). The application of systems thinking in health: why use systems thinking? *Health Research Policy and Systems, 12*(1), 1-6. doi: 10.1186/1478-4505-12-51

Pierson, Paul. (2000). Not Just What, but When: Timing and Sequence in Political Processes. *Studies in American Political Development, 14*(01), 72-92. doi: doi:null

Pierson, Paul. (2004). *Politics in time: History, institutions, and social analysis*: Princeton University Press.

Pinggu Health Bureau. (1973). Health Statistics Report 1973.

Pinggu Health Bureau. (1975). Statistics of Barefoot Doctors Trained between 1965 and 1975.

Pinggu Health Bureau. (1979). Health Statistics Report 1979.

Pinggu Health Bureau. (1986a). Medical Education in Pinggu (*Pinggu yixue jiaoyu*) (in Chinese).

Pinggu Health Bureau. (1986b). *Pinggu Health Gazette*.

Primary Health Care Performance Initiative. (2015). Methodology Note.

Qian, Xinzhong. (1988). Historical Experience in Health Management in China (*Zhongguo weisheng guanli de lishi jingyan*) (in Chinese). *Chinese Social Medicine*.

QSR International. (2013). NVivo 11 for Windows.

Ramsey, Matthew. (1992). The popularization of medicine in France, 1650-1900. *The Popularization of*

Medicine, 1650-1850.

- Reibling, Nadine, & Wendt, Claus. (2013). Regulating patients' access to healthcare services. *Healthcare Management and Economics: Perspectives on Public and Private Administration: Perspectives on Public and Private Administration*, 53.
- Renmin Ribao. (1960). Stick to politics at the commanding height, closely integrate with production, scale up mass mobilisation (*Jianchi zhengzhi guashuai, miqie jiehe shengchan, dagao qunzhong yundong*) (in Chinese).
- Renshaw, Michelle Campbell. (2003). Accommodating the Chinese: the American hospital in China, 1880-1920.
- Revolutionary Committee of Leyuan Commune. (1968, 11 December 1968). We get firm hold of three major issues--experiences on implementing cooperative medical scheme in Leyuan Commune (*Women henzhua le sanjian dashi--leyuan gongshe shixing hezuo yiliao zhidu de jingyan*) (in Chinese), *People's Daily (Renmin ribao)*.
- Richardson, George P. (1986). Problems with causal-loop diagrams. *System dynamics review*, 2(2), 158-170.
- Rico, Ana, Saltman, Richard B, & Boerma, Wienke GW. (2003). Organizational restructuring in European health systems: the role of primary care. *Social Policy & Administration*, 37(6), 592-608.
- Risse, Guenter B. (1999). *Mending bodies, saving souls: a history of hospitals*: Oxford University Press, USA.
- Ritchie, Jane, & Spencer, Liz. (2002). Qualitative data analysis for applied policy research. *The qualitative researcher's companion*, 305-329.
- Rockefeller Foundation. (1922). *The Rockefeller Foundation Annual Report 1922*. New York.
- Rosano, A., Loha, C. A., Falvo, R., van der Zee, J., Ricciardi, W., Guasticchi, G., & de Belvis, A. G. (2013). The relationship between avoidable hospitalization and accessibility to primary care: a systematic review. *Eur J Public Health*, 23(3), 356-360. doi: 10.1093/eurpub/cks053
- Roter, Debra, & Hall, Judith A. (2006). *Doctors talking with patients/patients talking with doctors: improving communication in medical visits* (Vol. 2): Praeger Westport, CT.
- Royal College of General Practitioners. (1972). *The Future GP: Learning and Teaching*. London.
- Royston, Geoff. (1998). Shifting the balance of health care into the 21st century. *European Journal of Operational Research*, 105(2), 267-276. doi: <http://dx.doi.org/10.1016/S0377-349>

[2217\(97\)00234-8](#)

- Rwashana, A. S., Williams, D. W., & Neema, S. (2009). System dynamics approach to immunization healthcare issues in developing countries: a case study of Uganda. *Health Informatics J*, 15(2), 95-107. doi: 10.1177/1460458209102971
- Rwashana, Agnes Semwanga, Nakubulwa, Sarah, Nakakeeto-Kijjambu, Margaret, & Adam, Taghreed. (2014). Advancing the application of systems thinking in health: understanding the dynamics of neonatal mortality in Uganda. *Health Research Policy and Systems*, 12(1), 36.
- Rwashana, Agnes Semwanga, & Williams, Ddembe Wileese. (2007). Enhancing immunization coverage through health information systems: a system dynamics approach. *Studies in health technology and informatics*, 130, 247.
- Saltman, Richard, Rico, Ana, & Boerma, Wienke. (2005). *Primary Care in the Driver's Seat: Organizational Reform in European Primary Care*: Open University Press.
- Sanders, D, Kravitz, J, Lewin, S, & McKee, M. (1998). Zimbabwe's Hospital Referral System: Does It Work? *Health Policy and Planning*, 13(4), 359-370.
- Sarriot, Eric, Morrow, Melanie, Langston, Anne, Weiss, Jennifer, Landegger, Justine, & Tsuma, Laban. (2015). A causal loop analysis of the sustainability of integrated community case management in Rwanda. *Social Science & Medicine*, 131, 147-155. doi: <http://dx.doi.org/10.1016/j.socscimed.2015.03.014>
- Schnitzer, S., Balke, K., Walter, A., Litschel, A., & Kuhlmeier, A. (2011). *Do gatekeeping programs increase equality of health care in Germany? A comparison of the health care situation of participants and nonparticipants. [German]: Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*. 54 (8) (pp 942-950), 2011. Date of Publication: August 2011.
- Schwenkglenks, Matthias, Preiswerk, Georges, Lehner, Roman, Weber, Fritz, & Szucs, Thomas D. (2006). Economic efficiency of gatekeeping compared with fee for service plans: a Swiss example. *Journal of epidemiology and community health*, 60(1), 24-30.
- Scott, John. (1990). *A matter of record*: Cambridge: Polity Press.
- Scottish Executive. (2004). *Delivering for Health*. Edinburgh, UK.
- Sewell, William H. (1996). Three temporalities: Toward an eventful sociology. *The historic turn in the human sciences*, 245-280.
- Shandong Provincial Health Historical Gazette Compiling Committee. (1992). *Shandong sheng weishengzhi (山东省卫生志) [Shandong Provincial Health Gazette]*. Shandong: Shandong

People's Publishing House.

- Shanghai renmin chubanshe, Wang Linkun, collective work. (1974). Chijiao yisheng bian shancun hezuo yiliao qixiang xin [赤脚医生遍山村合作医疗气象新] (Barefoot doctors are all over the mountain villages, cooperation creates a new atmosphere of medical treatment): Shanghai renmin chubanshe (上海人民出版社).
- Sheikh, Kabir, Gilson, Lucy, Agyepong, Irene Akua, Hanson, Kara, Ssengooba, Freddie, & Bennett, Sara. (2011). Building the field of health policy and systems research: framing the questions. *PLoS Med*, 8(8), e1001073.
- Shi, Leiyu. (1995). Balancing primary versus specialty care. *Journal of the Royal Society of Medicine*, 88(8), 428.
- Shi, Leiyu. (2012). The Impact of Primary Care: A Focused Review. *Scientifica*, 2012.
- Shi, Leiyu, Starfield, Barbara, Politzer, Robert, & Regan, Jerri. (2002). Primary care, self-rated health, and reductions in social disparities in health. *Health services research*, 37(3), 529-550.
- Skocpol, Theda. (1985). Bringing the state back in: Strategies of analysis in current research. In P. B. Evans, D. Rueschemeyer & T. Skocpol (Eds.), *Bringing the state back in*. Cambridge, UK: Cambridge University Press.
- Skocpol, Theda, & Amenta, Edwin. (1986). States and social policies. *Annual Review of Sociology*, 131-157.
- Song, Chunyan, & Yin, Aitian. (2010). Guanyu wanshan xinxing nongcun hezuoyiliao tongchou buchang fang'an de zhidao yijian [关于完善新型农村合作医疗统筹补偿方案的指导意见] (Guiding opinions regarding perfecting measures of pooling and reimbursement of the New Rural Cooperative Medical Scheme). *Chinese Health Economics*, 29(3), 57-59.
- Spence, J.D. (1969). *To Change China: Western Advisers in China, 1620-1960*: Little, Brown.
- Štampar, Andrija. (1936). Health and Social Conditions in China. *Quarterly Bulletin of the Health Organization of the League of Nations*, 5, 1090-1126.
- Stanley, John R. (2006). Professionalising the Rural Medical Mission in Weixian, 1890–1925. In D. Hardiman (Ed.), *Healing Bodies, Saving Souls: Medical Missions in Asia and Africa* (pp. 115-136): Rodopi.
- Starfield, Barbara. (1992). *Primary Care: Concept, Evaluation, and Policy*. New York, NY: Oxford University Press.

- Starfield, Barbara. (1994). Is primary care essential? *Lancet*, 344(8930), 1129-1133.
- Starfield, Barbara. (1998). *Primary care: balancing health needs, services, and technology*: Oxford University Press, USA.
- State Council. (1963). *Guowuyuan pifu weishegnbu guanyu nongcun yiyao jiuji zhuankuan wenti de baogao* [国务院批复卫生部《关于农村医疗救济转款问题的报告》] (*State Council's Approval and Reply to Ministry of Health Report on Issues Concerning Special Fund for Medical Relief in Rural Areas*).
- State Council. (1978a). *Guanyu anzhi lao ruo bing can ganbu de zanxing banfa* [关于安置老弱病残干部的暂行办法] (*Temporary measures on settling old, weak, sick and handicapped cadres*).
- State Council. (1978b). *Guanyu gongren tuixiu tuizhi de zanxing banfa* [关于工人退休退职的暂行办法] (*Temporary Measures on Retirement and Resignment of Workers*).
- State Council. (1981a). *Guowuyuan pizhuan weisheng bujiejue yiyuan peiben wenti de baogao* [国务院批转卫生部《解决医院赔本问题的报告》] (*State Council's circular on approving and Ministry of Health's Report Regarding Solving the Problem of Economic Loss of Hospitals*).
- State Council. (1981b). *Pizhuan weishengbu guanyu heli jiejie chijiao yisheng butie wenti de baogao de tongzhi* [批转卫生部《关于合理解决赤脚医生补贴问题的报告》的通知] (*Circular on approving and forwarding Ministry of Health's Report on Appropriately Solving the Subsidy for Barefoot Doctors*).
- State Council. (1985). *Guowuyuan pizhuan weishengbu guanyu weisheng gongzuo gaige ruogan zhegnce wenti baogao de tongzhi* [国务院批转卫生部关于卫生工作改革若干政策问题报告的通知] (*Circular on Approving and Forwarding Ministry of Health's Report on Several Policy Issues Related to Health Work Reform*).
- State Council. (1989). *Guanyu kuoda yiliao weisheng fuwu youguan wenti de yijian* [关于扩大医疗卫生服务有关问题的意见] (*Opinions on Issues Related to Expanding Medical and Health Service*).
- State Council. (1992). *Guanyu shenhua weisheng gaige de jidian yijian* [关于深化卫生改革的几点意见] (*Some Opinions Regarding Deepening Health Reform*).
- State Council. (1998). *Guanyu jianli chengzhen zhigong jiben yiliao baoxian zhidu de jue ding* [关于建立城镇职工医疗保险制度的决定] (*Decision regarding establishing urban employees basic medical insurance scheme*).
- State Council. (2000). *Guanyu chengzhen yiyao weisheng tizhi gaige de zhidao yijian* [关于城镇医药卫生体制改革的指导意见] (*Guiding Opinions on Urban Health System Reform*).

- State Council. (2009). *Guowuyuan guanyu yinfa yiyao weishegn tizhi gaige jinqi zhongdian shishifangan (2009-2011nian) de tongzhi* [国务院关于印发医药卫生体制改革近期重点实施方案 (2009—2011 年)] (*The State Council's circular regarding the implementation plan for key points of health system reform in the near future (2009-2011)*).
- Steinmo, Sven, & Watts, Jon. (1995). It's the institutions, stupid! Why comprehensive national health insurance always fails in America. *Journal of Health Politics, Policy and Law*, 20(2), 329-372.
- Sterman, John. (2000). *Business dynamics: Systems Thinking and Modeling for a Complex World*: Irwin-McGraw-Hill.
- Stone, D. A. (1978). Physicians as gatekeepers: illness certification as a rationing device. *Public Policy*, 27(2), 227-254.
- Sun, Zesheng, Wang, Shuhong, & Barnes, Stephen R. (2015). Understanding congestion in China's medical market: an incentive structure perspective. *Health Policy and Planning*. doi: 10.1093/heapol/czv062
- Sutton, Matt, Nikolova, Silviya, Boaden, Ruth, Lester, Helen, McDonald, Ruth, & Roland, Martin. (2012). Reduced mortality with hospital pay for performance in England. *New England Journal of Medicine*, 367(19), 1821-1828.
- Sze, Szeming. (1943). *China's Health Problems*. Washington, D.C.: Chinese Medical Association.
- Tucker, Sue, Brand, Christian, Wilberforce, Mark, & Challis, David. (2013). The balance of care approach to health and social care planning: Lessons from a systematic literature review. *Health Services Management Research*, 26(1), 18-28. doi: 10.1177/0951484813481966
- Tucker, Sue, Hughes, Jane, Brand, Christian, Buck, Deborah, & Challis, David. (2015). The quality and implications of Balance of Care studies: Lessons from a systematic literature review. *Health Services Management Research*. doi: 10.1177/0951484815607548
- Tuohy, C.H. (1999). *Accidental Logics: The Dynamics of Change in the Health Care Arena in the United States, Britain, and Canada*: Oxford University Press.
- Unger, Jonathan, & Chan, Anita. (1995). China, corporatism, and the East Asian model. *The Australian Journal of Chinese Affairs*, 29-53.
- United Nations. (2015). *Transforming our World: the 2030 agenda for sustainable development*: A/RES/70/1: United Nations New York.
- Varghese, J., Kutty, R., Paina, L., & Adam, T. (2014). Advancing the application of systems thinking in health: understanding the growing complexity governing immunization services in Kerala,

- India. *Health Res Policy Syst*, 12. doi: 10.1186/1478-4505-12-47
- Ventana Systems Inc. (2016). Vensim® Personal Learning Edition. Retrieved from <http://vensim.com/vensim-personal-learning-edition/>
- Vuori, H. (1985). The role of the schools of public health in the development of primary health care. *Health Policy*, 4(3), 221-230.
- Walker, Simon, Mason, Anne R, Claxton, Karl, Cookson, Richard, Fenwick, Elisabeth, Fleetcroft, Robert, & Sculpher, Mark. (2010). Value for money and the Quality and Outcomes Framework in primary care in the UK NHS. *Br J Gen Pract*, 60(574), e213-e220.
- Wammes, Joost Johan Godert, Jeurissen, Patrick Paulus Theodoor, Verhoef, Lise Maria, Assendelft, Willem JJ, Westert, Gert P, & Faber, Marjan J. (2014). Is the role as gatekeeper still feasible? A survey among Dutch general practitioners. *Family practice*, 31(5), 538-544.
- Wang, Gusmano, MK, & Cao, Q. (2011a). An evaluation of the policy on community health organizations in China: will the priority of new healthcare reform in China be a success? *Health Policy*, 99(1), 37-43. doi: 10.1016/j.healthpol.2010.07.003
- Wang, Guixin. (1957). Over thirty cities have implemented sectional medicine--urban medical and preventive care is gradually improving (*sanshi duo chengshi shixing huaqu yiliao--chengshi yiliao he yufang gongzuo de mianmao zhengzai zhubu gaishan*) (in Chinese), *People's Daily (Renmin ribao)*.
- Wang, Hufeng, Gusmano, Michael K, & Cao, Qi. (2011b). An evaluation of the policy on community health organizations in China: will the priority of new healthcare reform in China be a success? *Health policy*, 99(1), 37-43.
- Wang, Wenhua, Yin, Aitian, & Song, Chunyan. (2011). *Shandong nongcun diqu menzhen chagnjianbing ke fenliu bingzhong yanjiu* [山东省农村地区门诊常见病可分流病种研究] (On the cases of common illnesses of which patient utilisation pament could be rationalized: a study based on the rural ambulatory care settings in Shandong Province). *Zhongguo weisheng shiye guanli [中国卫生事业管理] (Chinese Health Services Management)*, 28(3), 212-214. doi: 10.3969/j.issn.1004-4663.2011.03.018
- Ward, TR. (2010). Implementing a gatekeeper system to strengthen primary care in Egypt: pilot study. *EMHJ*, 16(6).
- Watt, J.R. (2013). *Saving Lives in Wartime China: How Medical Reformers Built Modern Healthcare Systems Amid War and Epidemics, 1928-1945*: Brill.
- Wear, Andrew. (1992). The popularization of medicine in early modern England. *The Popularization of*

Medicine, 1650-1850.

Webster, Charles. (1993). The metamorphosis of Dawson of Penn. In D. Porter & R. Porter (Eds.), *Doctors, politics and society: historical essay* (pp. 212-228). Amsterdam: Rodopi.

White, Kerr. (1998). History of Health Services Research Project. In E. Berkowitz (Ed.).

White, Kerr L, Williams, T Franklin, & Greenberg, Bernard G. (1961). The ecology of medical care. *The New England journal of medicine*, 265, 885-892.

Williams, B., & Hummelbrunner, R. (2010a). *Systems Concepts in Action: A Practitioner's Toolkit*. Stanford, CA: Stanford University Press.

Williams, Bob. (2015). Prosaic or Profound? The Adoption of Systems Ideas by Impact Evaluation. *IDS Bulletin*, 46(1), 7-16. doi: 10.1111/1759-5436.12117

Williams, Bob, & Hummelbrunner, Richard. (2010b). *Systems concepts in action: A practitioner's toolkit*. Stanford, CA: Stanford University Press.

Wilsford, David. (1994). Path dependency, or why history makes it difficult but not impossible to reform health care systems in a big way. *Journal of public policy*, 251-283.

Wong, K. Chimin, & Wu, Lien-teh. (1936). *History of Chinese Medicine--Being a chronicle of medical happenings in China from ancient times to the present period*. Shanghai, China: National Quarantine Service.

World Bank. (2016a). Gross domestic product 2015.

World Bank. (2016b). Population 2015.

World Health Organization. (1978). *Alma-Ata 1978: primary health care*. Paper presented at the Report of the International Conference on Primary Health Care.

World Health Organization. (2000). *The world health report 2000: health systems: improving performance*: WHO.

World Health Organization. (2003). *The world health report 2003: shaping the future*: World Health Organization.

World Health Organization. (2007). *Everybody business : strengthening health systems to improve health outcomes*. Geneva.

World Health Organization. (2008). *The world health report 2008: primary health care: now more than*

- ever. Geneva: World Health Organization.
- World Health Organization. (2014). WHO Global Health Expenditure Atlas. Geneva, Switzerland.
- World Health Organization Regional Office for Europe. (2016). Health topics > Health systems > Primary health care > Main terminology. Retrieved 23 June 2016, from <http://www.euro.who.int/en/health-topics/Health-systems/primary-health-care/main-terminology>
- World Trade Organization. (2015). International Trade Statistics 2015.
- Wu, Lien-teh. (1931). Altruism in the medical profession. *The National Medical Journal of China*, 47(2), 268-282.
- Wu, Liende. (1937). *Gongyi zhidu zhi gaiyao* [公医制度之概要] (Brief introduction to State Medicine). *National Medical Journal of China*, 23(5), 573-577.
- Wu, Wenjun. (2014). The Study of Construction of the Public Health System in the Early New China—Based on the perspective of Anti-waste (*xinzhongguo chuqi gongfei yiliao zhidu jianshe yanjiu—jiyu fan langfei de shijiao*) (in Chinese). *Technological Economy and Management Studies (jingji jishu yu guanli yanjiu)*(7), 86-90.
- Wu, Yixiong. (2000). *Zai zongjiao yu shisu zhijian—jidujiao xinjiao zai hua'nan yanhai de zaoqi huodong yanjiu* [在宗教与世俗之间——基督教新教在华南沿海的早期活动研究] (*Between Religion and Secularism: Study on Early Activities of Christian Puritants Missionaries in Southern China*). Guangdong: Guangdong Education Press.
- Xinhua. (2016a, 20 August 2016). Quanguo weisheng yu jiankang dahui 19 ri zhi 20 ri zai jing zhaokai [全国卫生与健康大会 19 日至 20 日在京召开] (National Congress on Health held in Beijing between 19th and 20th). Retrieved 1 October 2016, from http://www.gov.cn/guowuyuan/2016-08/20/content_5101024.htm
- Xinhua. (2016b). Zhonggong zhongyang guowuyuan yinfa "jiankang zhongguo 2030" guihua gangyao [中共中央 国务院印发《“健康中国 2030”规划纲要》] (Central Committee of Communist Party of China and State Council issued Plan for "Healthy China 2030"). from http://news.xinhuanet.com/politics/2016-10/25/c_1119785867.htm
- Xinhua Agency. (1951). Rapid expansion of rural health organisations across the country (*Quan'guo nongcun weisheng jiceng zuzhi xunsu kuozhan*) (in Chinese), *People's Daily (Renmin ribao)*.
- Xinhua Agency. (1961, 30 November 1961). *Chongshi jiceng liliang, jiaqiang jishu fudao, shandong nongcun weishegn zuzhi gonggu fazhan* [充实基层力量 加强技术辅导——山东农村卫生组织巩固发展] (Substantiating grassroots workforce, enhancing technical guidance--

- consolidation and development of rural health organizations in Shandong), *Renmin ribao*.
- Xu, Biao, Diwan, Vinod K., & Bogg, Lennart. (2007). Access to tuberculosis care: What did chronic cough patients experience in the way of healthcare-seeking? *SCANDINAVIAN JOURNAL OF PUBLIC HEALTH*, 35(4), 396-402.
- Xu, Ligang, Yin, Aitian, & Jing, Shanshan. (2012). Mechanism of influence of New Rural Cooperative Medical Scheme on rationalised patient pathway. *China Health Service Management*, 3, 021.
- Xu, Yunbei. (1960). *Zuohao renmin gongshe de weisheng gongzuo* [做好人民公社的卫生工作] (Doing well health work in rural people's communes) *People's Daily (Renmin ribao)*.
- Xu, Zongyu, Wang, Jian, Meng, Qingyue, Jia, Liying, & Qu, Jiangbin. (2007). *Yiyuan menzhen huanzhe xiang shequ weisheng fuwu jigou fenliu de jingji xiaoyi* [医院门诊患者向社区卫生服务机构分流的经济效益] (Economic benefits of shifting hospital ambulatory patients to community health services institutions). *Zhongguo weisheng jingji [中国卫生经济] (Chinese Health Economics)*, 26(5), 26-28.
- Yang, Gonghuan, Wang, Yu, Zeng, Yixin, Gao, George F, Liang, Xiaofeng, Zhou, Maigeng, . . . Naghavi, Mohsen. (2013). Rapid health transition in China, 1990–2010: findings from the Global Burden of Disease Study 2010. *The lancet*, 381(9882), 1987-2015.
- Yang, Hui, Shi, Leiyu, Lebrun, Lydie A., Zhou, Xiaofei, Liu, Jiyang, & Wang, Hao. (2013). Development of the Chinese primary care assessment tool: data quality and measurement properties. *International Journal for Quality in Health Care Journal of the International Society for Quality in Health Care*, 25(1), 92.
- Yang, Nianqun. (2006). *Remaking "Patient": Spatial Politics under the Conflicts Between Chinese and Western Medicine, 1832-1985* China Renmin University Press.
- Yi, Yaowen. (1958). *Jishanxian de nongcun weisheng baojian wang* [稷山县的农村卫生保健网] (Rural Health Care Network in Jishan County), *People's Daily (Renmin Ribao)*.
- Yin, Qian. (2013). *Minguo shiqi de yishi qunti yanjiu (1912-1937): yi Shanghai wei taolun zhongxin [民国时期的医师群体研究 (1912-1937) 以上海为讨论中心] (Study on medical doctors group in the Republic era (1912-1937): focusing on discussion about Shanghai)*. Beijing: Chinese Social Sciences Press.
- Yin, Robert K. (2009). *Case study research: Design and methods* (Vol. 5): Sage.
- Yin, Yi. (1955, 8 January, 1955). *Kefu gongfei yiliao zhong de langfei xianxiang* [克服公费医疗中的浪费现象] (Overcoming the phenomena of waste in publicly funded medical care), *People's Daily (renmin ribao)*.

- Yip, Ka-Che. (1992). Health and Nationalist Reconstruction: Rural Health in Nationalist China, 1928-1937. *Modern Asian Studies*, 26(2), 395-415.
- Yip, W, & Hsiao, WC. (2009). How effective is China's new cooperative medical scheme in reducing medical impoverishment? *Social Science & Medicine*, 68, 201 - 209.
- Yip, Winnie, & Hsiao, William C. (2015). What Drove the Cycles of Chinese Health System Reforms? *Health Systems & Reform*, 1(1), 52-61. doi: 10.4161/23288604.2014.995005
- Zahradnik, Anne G. (2008). Does providing uninsured adults with free or low-cost primary care influence their use of hospital emergency departments? *Journal of health and human services administration*, 240-258.
- Zhang, Daqing. (2006). *Zhongguo jindai jibing shehuishi (1912-1937) (中国近代疾病社会史 (1912-1937))* [Social history of diseases in Modern China: 1912–1937]. Jinan: Shandong Education Press.
- Zhang, Qin. (1950). *Guanyu nongcun weisheng jianshe wenti--ji zhongyang weishengbu nongcun weisheng zuotanhui* [关于农村卫生建设问题——记中央卫生部农村卫生座谈会] (On constructing rural health services--notes of Rural Health Forum of Central Ministry of Health), *Renmin ribao (People's Daily)*.
- Zhang, Xuewen. (1953). *Xinzhongguo de weishegnshiye [新中国的卫生事业] (Health Affairs of the New China)*. Beijing: Life-Reading-New Knowledge Three-United Bookstore (*shenghuo-dushu-xinzhi sanlian shudian*).
- Zhang, Zikuan. (1952). *Canyu dongbei diqu nongcun weishegn shiyaxian (qu) jianshe de tihui* [参与东北地区农村卫生试验县（区）建设的体会] (Experiences on participation in Constructing Rural Health Experiment County (Sub-county) Northeast Region).
- Zhang, Zikuan. (1955). *Guanyu tongxian xiangcun weisheng shiyangu qingkuang de diaocha* [关于通县乡村卫生试验区情况的调查] (Investigation on Situation of Rural Health Experiment Area in Tong County).
- Zhang, Zikuan. (1982). *Guanyu Jiaqiang Nongcun Weisheng Duiwu de Guanli he Jishu Peixun Wenti—Dongbei Sansheng Nongcun Yiliaoweisheng Jianshe Diaocha Zhi San* [关于加强农村卫生队伍的管理和技术培训问题——东北三省农村医疗卫生建设调查之三] (On Strengthening Management of Rural Health Workforce and Professional Training—Rural Health Construction Investigation in the Three Northeast Provinces, No. 3). *Nongcun Weisheng Shiye Guanli Yanjiu (Rural Health Service Management Research)* 2.
- Zhang, Zini. (1962, 1962-05-30). *Xiangcun yisheng (tupian)* [乡村医生（图片）] (Rural doctor (image)), *Renmin ribao (People's Daily)*.

- Zheng, Liang. (2013, 23 July 2013). *Zhangzhou yiliao gouxiao lingyu fubai wo'an daqidi: yaojia yiban yongyu "gongguan"* [漳州医疗购销领域腐败窝案大起底: 药价一半用于“公关”] (Unveiling the series corruption cases in pharmaceutical procurement and sales Zhangzhou: half of drug prices is used on "public relationships"), *People's Daily (Renmin Ribao)*. Retrieved from http://paper.people.com.cn/rmrb/html/2013-07/23/nw.D110000renmrb_20130723_10-04.htm
- Zheng, Yongnian. (2007). *De facto federalism in China: Reforms and dynamics of central-local relations* (Vol. 7): World Scientific Publishing Company.
- Zhi, Yi. (1928). *Xiwang yu weishengbu zhe* [希望于卫生部者] (Hopes on the Ministry of Health). *National Medical Journal of China*, 14(06), 113-115.
- Zhou, Xu Dong, Li, Lu, Hesketh, Therese, Liu, Qing, Tian, Xin, Tian, Jiang, & Zhang, Xinpeng. (2014). Health system reform in rural China: Voices of healthworkers and service-users: Evaluation of the effects of comprehensive reform on primary healthcare institutions in Anhui Province. *Social Science & Medicine*, 117, 134. doi: 10.1186/1472-6963-14-268 %/ BioMed Central
- Zhou, Xueguagn. (1999). Review of current situation of Western sociological research about organisational and institutional changes in China (*Xifang shehuixue guanyu zhongguo zuzhi yu zhidu bianqian yanjiu zhuangkuang shuping*) [西方社会学关于中国组织与制度变迁研究状况述评]. *Sociological research*, 4, 26-43.
- Zylberman, Patrick. (2004). Fewer parallels than antitheses: René Sand and Andrija Stampar on social medicine, 1919–1955. *Social history of medicine*, 17(1), 77-92.

Appendices

Appendix Table 1. Calculation of price deflator

	Nominal GDP (100 million yuan)	GDP at constant prices (100 million yuan)	Base year*	Price deflator	Base year**
2001	110270.4	108062.5	2000	1.020432	2000
2002	121002.0	117879.7	2000	1.026487	2000
2003	136564.6	129693.7	2000	1.052978	2000
2004	160714.4	142763.2	2000	1.125741	2000
2005	185895.8	158963.9	2000	1.169421	2000
2005	--	185895.8	2005	--	--
2006	217656.6	209478.4	2005	1.215077	2000
2007	268019.4	239223.2	2005	1.310189	2000
2008	316751.7	262238.8	2005	1.412515	2000
2009	345629.2	286459.4	2005	1.410972	2000
2010	408903.0	316907.9	2005	1.508892	2000
2010	--	408903.0	2010	--	--
2011	484123.5	447696.4	2010	1.631664	2000
2012	534123.0	482381.9	2010	1.670739	2000
2013	588018.8	519454.6	2010	1.708055	2000

Data sources: (National Bureau of Statistics of China, 2015)

Note:

- a) Base year* refers to base year used for GDP at constant prices.
- b) Base year** refers to base year used for calculating price deflector.
- c) Calculation of price deflator price deflator in year X:

If $2000 < X \leq 2005$, $DEFLATOR_X = GDP_X \div GDP_{X,2000price}$

If $2005 < X \leq 2010$, $DEFLATOR_X = GDP_X \div GDP_{X,2005price} \times DEFLATOR_{2005}$

If $2010 < X \leq 2013$, $DEFLATOR_X = GDP_X \div GDP_{X,2010price} \times DEFLATOR_{2010}$

Appendix Table 2. Standard Periodization of Chinese History

Dynasty/Regime	Years
Age of Five Rulers	B.C. 2953- B.C. 2208
Xia Dynasty	B.C. 2207- B.C. 1766
Shang Dynasty	B.C. 1765- B.C. 1122
Zhou Dynasty	B.C. 1121- B.C. 249
Warring States	B.C. 248- B.C. 222
Qin Dynasty	B.C. 221- B.C. 207
Han Dynasty	B.C.206 -A.D. 220
The Three Kingdoms	221-264
Jin Dynasty	265-419
Epoch of Division of North and South	420-589
Sui Dynasty	589-618
Tang Dynasty	619-907
Epoch of the Five Dynasties	907-960
Song Dynasty	960-1279
Jin Dynasty	1115-1233
Yuan Dynasty	1206-1368
Ming Dynasty	1368-1644
Qing Dynasty	1644-1911
Republic of China	1912-1949
Nationalist Administration	1927-1949
People's Republic of China	1949 onward
Cultural Revolution	1966-1976
Reform and Opening-up	1978 onward

Appendix Table 3. Regression results on visits

Weighted model at township level	Ambulatory visits at primary care facilities per enrollee per quarter	Ambulatory visits at secondary hospitals per enrollee per quarter	Ambulatory visits at tertiary hospitals per enrollee per quarter	Proportion of ambulatory visits at primary care in total ambulatory visits
Simple difference-in-differences (DD)	0.101*** (0.000)	-0.00854 (0.516)	-0.00131 (0.554)	0.0705 (0.475)
Multiple-period DD	0.101*** (0.000)	-0.00854 (0.525)	-0.00131 (0.563)	0.0705 (0.485)
Multiple-period DD with covariates	0.0984** (0.033)	-0.00796 (0.373)	-0.00114 (0.255)	0.0665 (0.235)
Pre-trend test	0.00337 (0.780)	-0.00271 (0.576)	-0.00134** (0.010)	-0.00917 (0.531)
Mean outcome in pilot townships before introduction of pilot programme	0.2856709	0.088252	0.0107624	0.6397892
Unweighted model at village level	Ambulatory visits at primary care facilities per enrollee per quarter	Ambulatory visits at secondary hospitals per enrollee per quarter	Ambulatory visits at tertiary hospitals per enrollee per quarter	Proportion of ambulatory visits at primary care in total ambulatory visits
Simple difference-in-differences (DD)	0.171** (0.003)	-0.0254 (0.133)	-0.00349 (0.266)	0.0856* (0.026)
Multiple-period DD	0.156*** (0.000)	-0.0295 (0.130)	-0.00349 (0.268)	0.0926** (0.013)
Multiple-period DD with covariates	0.157*** (0.000)	-0.0297 (0.127)	-0.00365 (0.244)	0.0926** (0.012)
Pre-trend test	0.0332	0.00725	-0.00275**	0.00792

	(0.289)	(0.488)	(0.042)	(0.388)
Mean outcome in pilot townships before introduction of pilot programme	0.3640666	0.1383391	0.0139499	0.5681763
Weighted model at village level	Ambulatory visits at primary care facilities per enrollee per quarter	Ambulatory visits at secondary hospitals per enrollee per quarter	Ambulatory visits at tertiary hospitals per enrollee per quarter	Proportion of ambulatory visits at primary care in total ambulatory visits
Simple difference-in-differences (DD)	0.105*** (0.001)	-0.00748 (0.342)	-0.00110 (0.360)	0.0566 (0.270)
Multiple-period DD	0.105*** (0.001)	-0.00759 (0.336)	-0.00110 (0.356)	0.0567 (0.269)
Multiple-period DD with covariates	0.104*** (0.001)	-0.00760 (0.328)	-0.00109 (0.343)	0.0565 (0.246)
Pre-trend test	0.00509 (0.641)	-0.00346 (0.299)	-0.00164*** (0.001)	-0.000730 (0.945)
Mean outcome in pilot townships before introduction of pilot programme	0.2846467	0.0880593	0.0107551	0.6108143

Note:

a) p values are shown in brackets.

b) *0.05≤p<0.1; ** 0.01≤p<0.05, *** p<0.01.

Appendix Table 4. Regression results on expenditures

Weighted model at township level	Total ambulatory expenditures at primary care per enrollee per quarter	Total ambulatory expenditures at secondary hospitals per enrollee per quarter	Total ambulatory expenditures at tertiary hospitals per enrollee per quarter	Total expenditures on ambulatory care	Proportion of total ambulatory expenditures at primary care facilities in all ambulatory expenditures
Simple DD	8.257* (0.067)	4.662 (0.558)	-12.76** (0.034)	0.158 (0.993)	0.0226 (0.746)
Multiple-period DD	8.257* (0.073)	4.662 (0.567)	-12.76** (0.038)	0.158 (0.993)	0.0226 (0.751)
Multiple-period DD with covariates	8.151 (0.221)	5.590** (0.037)	-12.49*** (0.002)	1.248 (0.891)	0.0187 (0.478)
Pre-trend test	-3.046 (0.263)	0.888 (0.847)	-0.328 (0.946)	-2.486 (0.736)	-0.00529 (0.672)
Mean outcome in pilot townships before introduction of pilot programme	54.32007	65.68258	29.39126	149.3939	0.3913123
Unweighted model at village level	Total ambulatory expenditures at primary care per enrollee per quarter	Total ambulatory expenditures at secondary hospitals per enrollee per quarter	Total ambulatory expenditures at tertiary hospitals per enrollee per quarter	Total expenditures on ambulatory care	Proportion of total ambulatory expenditures at primary care facilities in all ambulatory expenditures
Simple DD	30.04* (0.064)	9.734 (0.133)	-9.657* (0.070)	30.12 (0.144)	0.00888 (0.773)
Multiple-period DD	28.68** (0.040)	10.42 (0.116)	-9.352* (0.079)	29.75 (0.136)	0.0128 (0.673)
Multiple-period DD with covariates	29.32**	10.09	-9.666*	29.74	0.0141

	(0.037)	(0.122)	(0.065)	(0.131)	(0.646)
Pre-trend test	-0.443	1.336	-0.847	0.0460	-0.00273
	(0.865)	(0.649)	(0.822)	(0.995)	(0.762)
Mean outcome in pilot townships before introduction of pilot programme	62.35645	77.46709	26.71046	70467.11	0.3955969
Weighted model at village level	Total ambulatory expenditures at primary care per enrollee per quarter	Total ambulatory expenditures at secondary hospitals per enrollee per quarter	Total ambulatory expenditures at tertiary hospitals per enrollee per quarter	Total expenditures on ambulatory care	Proportion of total ambulatory expenditures at primary care facilities in all ambulatory expenditures
Simple DD	9.086***	5.451	-12.39**	2.151	-0.0160
	(0.003)	(0.269)	(0.012)	(0.834)	(0.649)
Multiple-period DD	9.025***	5.422	-12.39**	2.055	-0.0159
	(0.003)	(0.272)	(0.012)	(0.841)	(0.651)
Multiple-period DD with covariates	8.920***	5.464	-12.37**	2.012	-0.0164
	(0.003)	(0.241)	(0.011)	(0.843)	(0.625)
Pre-trend test	-2.793**	-0.158	-0.892	-3.843	-0.000557
	(0.031)	(0.953)	(0.796)	(0.438)	(0.955)
Mean outcome in pilot townships before introduction of pilot programme	54.12324	65.65453	29.3612	146797.8	0.4052415

Note:

a) p values are shown in brackets.

b) *0.05 ≤ p < 0.1; ** 0.01 ≤ p < 0.05, *** p < 0.01.

Appendix Table 5. Causes, effects and links related to gatekeeping

Category of factors	Causes	Effects (direct and indirect)	Source	Constructed links	Feedback loops
1. Governance	Gatekeeping created extra procedure, particularly as others are not doing this. Gatekeeping affected people with political influence mainly use tertiary hospitals.	Resentment of patients. Influential people in particular put pressure on the local government which reduced the political will of strict gatekeeping, this put a pressure on the management agency in implementing gatekeeping.	D06, D08, M01, A02	gatekeeping (+)→ resistance (-)→ gatekeeping	B1
	Insurance management agency needs to respond to the demand of patients.	Difficulty in extending gatekeeping policy with the weak primary care service capacity.	A02, A03	hospital visits (+)→ hospital bargaining power (-)→ gatekeeping	R3
	Insurance management agency integrated within health bureaucracy needs to respond to the demand of health facilities as a system.	Pressure to strengthen primary care facilities through gatekeeping (reducing the attractiveness of hospital care).	A02	gatekeeping (-)→ hospital care attractiveness (-)→ PC visits	R1', R1a
	Monitoring and evaluation focusing on public health services and NCD management	Shift of care from ambulatory curative care to public health services and NCDs.	D05, D02, M01	performance evaluation focusing on public health services (-)→ PC curative service quality	R1', R1a
2. Health financing	Revenue surplus or deficit from fundholding.	The intended effect is that surplus or deficit is used to stimulate performance improvement of primary care.	A02, A03	gatekeeping (+)→ performance bonus (+)→ PC curative service quality; PC visit(+)-> PC revenue (+)-> performance bonus	R1', R1a
	Gatekeeping makes hospital care less attractive by lowered reimbursement rate.	The intended effect is that patients are incentivized to use primary care with more visits.	A03, A02	gatekeeping (-)→ hospital care attractiveness (-)→ PC visits	R1', R1a

	Increased service use of primary care facilities.	The intended effects are reduced hospital visits and expenditures.	M03, M04	PC visits(-)→ hospital visits(+) → hospital revenue	B1'
	Fixed and low total amount of salary.	Relatively low work morale.	D03, M05, M02	PC salary policy (-)→ performance bonus (+) → PC curative service quality	R1', R1a
	As primary care staff members consider that total salary should be equally distributed, actual amount variations of performance-based bonus is small.	Performance-based bonus is generally unable to incentivize curative care performance (the intended but not achieved goal).	M01, M02, M03, M04, M05, A02	performance bonus (+) → PC curative service quality	R1'
	PC staff incentive related policies provides little control knob of internal management.	Relying on personal relationship for management	M01, M02	PC salary policy (-) → performance bonus (+) → PC curative service quality	R1', R1a
	High-powered incentive in hospitals.	Relatively high work morale in hospitals contributed hospital care attractiveness, to the large patient volume, and to large revenue.	M05	hospital incentive (+) → hospital performance (+) → hospital revenue	R2
	Hospitals maintained high revenue.	Relatively higher salary for hospital doctors, with higher stress related to work.	M05	hospital revenue (+) → hospital salary (+) → hospital job attractiveness	R1', R1a
	Hospitals accounting for the lion share of expenditures and patients.	Prioritizing hospital-related policies.	M01, A02	hospital visits (+) → hospital bargaining power	R3
3. Service delivery	Poor curative service performance in primary care facilities.	Low patient trust of primary care facilities and hence higher attractiveness of hospital care.	M02	PC curative service quality (-) → hospital care attractiveness	R1', R1a
	Relatively low trust of primary care doctors by patients, contributes to higher attractiveness of hospital care.	Lower primary care visits, and unsophisticated cases (patients coming to buy drugs).	P01, D03, D05, M02	hospital care attractiveness(-) → PC visits	R1', R1a
	Small volume and unsophisticated cases of patients in curative care.	Declined clinical capacity.	M05, M02	PC visits (+) → clinical experiences (+) → PC capacity	R1a, R1'
	Complicated procedures	Patient resentment against gatekeeping. Even doctors resent the policy (as they had little	P01, P06, D02, D06,	gatekeeping (+) → inconvenience(-) → gatekeeping	B1

Appendices

	interest)	D04		
Integrative care arrangements are underdeveloped (Patients referred had no advantage in accessing specialist care in hospitals. Hospitals did not know doctors and their capability at primary care level).	Ineffective referral policies, which contributed to patient's resentment to gatekeeping policies.	D02, D05, D06, D07, M01, M05, A02	lack of integrated care arrangement (+)→ inconvenience (-)→ gatekeeping	B1
Tension between doctors and patients	Patient get referrals if they insist, so the policy becomes an inconvenience in many cases.	D04, D06, D07, D08	doctor-patient tension (+)→ inconvenience	B1
Elimination of hospitalisation and surgery at primary care level.	Surgeons' skills are wasted. Professional development was hindered.	M01, M03, M04	restriction of PC function (-)→ PC visits (+)→ clinical experiences	R1a
Relatively low trust of primary care doctors by hospital doctors.	Low hospital-to-primary-care referral rates, and primary care doctors are unable to function as a coordinator of care.	D07, M05, D08	hospital doctors lack confidence in PC capacity (-)→ hospital-to-PC referrals	B1'
Small volume of patients at primary care.	Primary care doctors converted to public health services staff, making the job unappealing to medical graduates.	M03, M04, D04	PC visits (-)→ PC doctors pivoting to public health services (-)→ PC job attractiveness	R1b
Weak primary care capacity.	Hospitals flooded with patients.	M05	PC capacity (+)→ PC curative service quality (-)→ hospital care attractiveness	R1a
Increasing workload and performance pressure from public health services, and small work low from curative care.	Primary care doctors required to pivot to public health services.	D02, D03, D06, M03, M04	performance evaluation focusing on public health services (+)→ PC doctors pivoting to public health services	R1b
Doctors significantly abandoning their previous curative functions.	Patients perceived a decline of doctors' capacity	P03	PC doctors pivoting to public health services(+)->low patient perception of PC doctor capacity	R2
4. Health	Poor quality of primary care doctors.	D02, D04,	PC HR quality (+)→ PC capacity	R1a, R1b

workforce	primary care doctors by both patients and doctors.	D08, M05		
PC doctors training improves PC HR quality.	Good doctors leave after training due to lack of willingness to stay.	M02, M05	PC staff training (+)→ PC HR quality (+)→ PC brain drain	B2, R1b
Poor career prospects for primary care doctors.	Difficulty in recruiting and retain good quality doctors	D08, M02, M03, M04, M05, A02	PC job attractiveness (+)→ PC recruitment (+)→ PC HR quality; PC job attractiveness (-)→ PC brain drain (-)→ PC HR quality;	R1b, B2
Expansion of hospitals service capacity.	Recruitment of doctors trained for primary care level.	D04, A03	hospital capacity expansion (+)→ hospital job attractiveness (+)→ hospital recruitment (+)→ hospital capacity advantage; hospital job attractiveness (-)→ PC job attractiveness (+)→ PC brain drain (+)→ hospital recruitment; hospital expansion (+)→ hospital job attractiveness (-)→ PC job attractiveness (+)→PC recruitment	R1b, R2
Performance-based salary policy under which managers were not able to stimulate the entrepreneurship of staff.	Primary care doctors had low working morale, and tended to send patients away to hospitals.	M02	PC salary policy (+)→ performance bonus (+)→ PC curative service quality	R1a
Low work morale of primary care doctors perceived by hospital specialists.	Low trust in the technical competence of primary care doctors by doctors.	M05	hospital doctors lack confidence in PC capacity (-)→ hospital-to-PC referral	B1'
Low work morale of primary care doctors.	Primary care facilities recruiting people with low professional aspiration.	M05	PC salary policy* (-)→ PC job attractiveness	R1b
5. Medical technologies	Hospitals technical advantage (poor technical capacity in primary care	Unappealing primary care capacity, and therefore hospital attractiveness	P01, D02, D03, D04, hospital capacity advantage(+)→ hospital care attractiveness	R2

	facilities).		M01		
	Physician patient tension.	Hospital technology use a necessity, therefore primary care doctors would tend to reject/refer such patients and hospital care became more attractive.	D08	doctor-patient tension* (+)→ reliance on technology (+)→ hospital care attractiveness	R1a, R1'
	Hospital doctors prescribe medicines not available at primary care (“advanced medicines”).	Patients can only get the medicines prescribed from hospitals. Primary care facilities the main facilities to adopt restrictive medicines policy.	P02, P03, P04, D01, D02, D03, D04, D07, M01, M02, M03, M04	restrictive pharmaceutical policies (-)→ PC capacity	R1a
	Essential medicine policies reduced access to “advanced medicines”	Patients cannot get from primary care facilities the medicines prescribed by hospital doctors.	M01, M04	balance of equipment (+)→ PC capacity; restrictive pharmaceutical policies (-)→ PC capacity	R1a
6. Information	With limited information sharing, hospital doctors don't know primary care doctors' capability and have little trust.	Reluctance to refer discharged patients to primary care facilities.	D04, D05, D08, A02	hospital doctors lack confidence in PC capacity (-)→ hospital-to-PC referrals	B1'

Note:

- 1) The polarity of relationship is determined based on the direction of association between the two neutral factors (e.g. quality of primary care doctors instead of poor quality of primary care doctors).
- 2) In the table, “(+)->” was used as a symbol for positive link causations, i.e. all else being equal, an increase in the factor preceding the signs leads to an increase in the factor following the sign; “(-)->” was used as a symbol for negative link causations, i.e. all else being equal, an increase in the factor preceding the signs leads to an increase in the factor following the sign.
- 3) Variables with a sign * appear more than ones.

Appendix Research Tools

1. Information for participants and informed consent form
2. Recruitment procedures
3. Semi-structured question guide

Information for Participants

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



Study title: Impact Evaluation of a Gatekeeping Pilot Program and Its Role in the Dynamic Development of Health System in China

You are being invited to take part in a research study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and to talk to others about the study, if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

1. What is the purpose of the study?

Shifting patients from hospitals towards primary care facilities has been a priority in the reform and development of health system in China. A new pilot program is going on in Beijing to achieve this aim. Our purpose is to see how effects the reform is likely to have in terms of cost and quality of care and further implications for the health system. The study is a student research project.

2. Why have I been chosen?

You are chosen because you represent one of the key informants of this study. 40-50 other participants will be studied.

3. Do I have to take part?

It is up to you to decide to join the study. We will describe the study and go through this information sheet. If you agree to take part, we will then ask you to sign a consent form. You are free to withdraw at any time, without giving a reason. This will not affect the standard of care you receive.

4. What will happen to me if I take part?

In the case of semi-structured interviews: We will conduct an interview with you, which will take about 10-15 minutes. You will be asked a few questions based on a flexible interview guide. In order to make sure the fidelity of notes, and for better analysis of the interview, we will audio-record and transcribe all the interviews, if you agree.

In the case of focus group discussions: We will organize a focus group discussion with key stakeholders of gatekeeping including you, around the implications of gatekeeping for health system development. In order to make sure the fidelity of notes, and for better analysis of the discussion, we will audio-record and transcribe all the interviews, if you agree.

5. Gift

A souvenir gift from Peking University will be given to participants as a 'thank-you' for participation.

6. What do I have to do?

You are expected to share your genuine answers to the questions posed to you.

7. What are the possible benefits of taking part?

We cannot promise the study will help you directly but the information we get might help improve the health system reform in Beijing and other areas in China.

8. What are the possible disadvantages and risks of taking part? Will my taking part in the study be kept confidential?

There is not likely any risk of taking part in the research. All information collected about you during the course of the research will be kept strictly confidential.

9. What will happen to the results of the research study?

The result of the study will be part of a PhD thesis. And our intention is to publish the findings and communicate it with policy makers in the hope that policy can be improved with sound evidence. You will not be identified in any report/publication.

10. Who is organising and funding the research?

The funding for the research is jointly provided by China Scholarship Council and Peking University.

11. Who has reviewed the study?

This study was given a favourable ethical opinion by the Research Ethics Committees of London School of Hygiene & Tropical Medicine and Peking University.

12. Contact Details

Jin Xu, Peking University China Centre for Health Development Studies
Mailbox 505, No. 38, Xueyuan Road, Haidian, Beijing
Telephone: 010-82805701
Email: xujin@bjmu.edu.cn

You will be given a copy of the information sheet and a signed consent form to keep.

Thank you for considering taking the time to read this sheet.

INFORMED CONSENT FORM

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



Full Title of Project: **Impact Evaluation of a Gatekeeping Pilot Program and Its Role in the Dynamic Development of Health System in China**

Name of Principal Investigator: Jin Xu

Please initial box

1. I confirm that I have read and understand the participant information sheet dated (version) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered fully.	
2. I understand that my participation is voluntary and I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.	
3. I agree to take part in the above study.	
4. I agree for my quote/recording to be used in the publication or report released on the study.	

Name of Participant
(printed)

Signature/Thumbprint

Date

Name of Person taking
consent

Signature

Date

Principal Investigator

Signature

Date

Chinese version of information for participants and informed consent

参与访谈知情同意书

我们将要开展“基层首诊制改革试点效果评估及其卫生体系发展的促进作用研究”，您的具体情况符合访谈人条件，因此，我们想邀请您参加该项研究。本知情同意书将向您介绍该研究的目的、步骤、获益、风险、不便或不适等，请仔细阅读后慎重做出是否参加研究的决定。当研究者向您说明和讨论知情同意书时，您可以随时提问并让他/她向您解释您不明白的地方。您可以与家人、朋友以及您的主治大夫（对患者适用）讨论之后再做决定。我们期待您提供您对所提的问题给予真实无保留的回答。

本项研究的中方项目负责人是孟庆跃教授，单位是北京大学中国卫生发展研究中心，研究资助方是国家留学基金委及北京大学。我们的研究目的是了解北京市新型农村合作医疗制度正在开展基层首诊制试点的效果，以及对卫生体系发展的长期作用。

1. 为什么进行这项研究？

强化基层卫生服务体系，将患者从医院转移到基层卫生机构是我国医药卫生体制改革和发展的重点之一。为此，北京市新型农村合作医疗制度正在开展基层首诊制的试点。我们的研究目的是了解该试点政策对于医疗服务质量和费用的影响，以及对卫生体系发展的长期作用。

2. 哪些人将被邀请参加这项研究？

研究选取首诊改革前后相关的关键知情人。您被选中是因为您代表了本研究的一种关键知情人。

3. 多少人将参与这项研究？

本研究计划招募 40-50 名知情人，包含以下类别：

访谈对象列表

受访者类型	人数
门诊患者	4
有经验的医生	4
卫生机构管理者	4
新农合管理人员	2
一年总数	14
两年总数	28

关键组讨论列表

参与者类型	人数
有经验的医生	4
卫生机构管理者	2
新农合管理人员	2
基层卫生专家	2
一年总数	10
两年总数	20

4. 本研究包括哪些内容？

分别进行半结构式访谈和关键组讨论。对于半结构式访谈的参与者，调查员将对您进行访谈。调查员将灵活地依据访谈提纲向您询问几个问题。为了确保记录的完整如实，也为了更好地分析，如果您同意的话，我们将对访谈进行录音和文字转录。对于关键组讨论的参与者，我们将组织一场涉及到基层首诊制的关键人员将参与讨论。讨论将围绕基层首诊制对卫生体系发展的影响展开。

5. 这项研究会持续多久？

访谈时间大概 10-15 分钟，关键组讨论时间大概 60-90 分钟。

6. 参加本项研究的危险是什么？

参与研究不会给您带来危险。研究过程中所有关于您的信息都将严格保密。如果不慎泄露个人私密信息，可能会给您的工作和生活带来一些困扰。

风险的控制措施：我们将在得到谈话记录文件后第一时间进行匿名化处理，并采取其他妥善的保密措施（详见 12）。

7. 参加本项研究的获益是什么？

我们不保证参与研究会直接对您有帮助，但通过您所提供的信息，我们将能够为帮助改进北京基层首诊制政策以及相关政策在全国的制定和实施提出更好的建议。

8. 是否一定要参加并完成本项研究？

您有权决定是否参与本研究。我们将描述研究并读完本告知书。如果您同意参与，我们将请您签署知情同意书。如果您不愿意，可以拒绝参加，不会对受到任何负面影响（对于患者：这不会影响您所接受的医疗服务）。即使您同意参加以后，您也可以在任何时间改变主意，告诉研究者退出研究，而不会对受到任何负面影响（对于患者：这不会影响您所接受的医疗服务）。原则上，在您退出之后，研究者将严密保存您的相关信息直至最终销毁，期间不会继续使用或透露这些信息。但在以下极少数情况下，研究者将继续使用或透露您的相关信息，即使您已经退出研究或研究已经结束。

这些情况包括：

- 除去您的信息将影响研究结果的科学性的评价；
- 为研究、教学或其他活动提供一些有限的信息（这些信息不会包括您的姓名、身份证号码、或者其他能识别您身份的个人信息）；

一旦出现任何可能会影响您决定是否继续参加该项研究的信息，我们会及时告知您。

9. 关于研究费用和补偿

本研究由国家留学基金委及北京大学共同资助。

10. 参加该项研究受试者是否获得报酬？

谈没有经济报酬，您是否参与研究完全是自愿的。

11. 发生研究相关伤害的处理？

本研究为观察性研究，无可预期的相关伤害。

12. 我的信息会保密吗？

如果您决定参加本项研究，您参加研究及在研究中的个人资料均属保密。您的谈话记录（含录音）文件及在研究中的个人资料均会在存储及分析前做匿名化处理。在匿名化处理中，您的姓名及其他可识别身份的信息将替换为不可识别身份的研究编码，比如“某区某乡镇卫生院某科医生张某”将被编码为“01020103”。记录文件以匿名化的研究编码加以标识和存储。建立并保存一个数据表格，记录受访者属性与其相应编码，删除具体科室和机构的可识别信息，仅保留工作岗位的性质信息，不保留可识别信息。所有的研究人员和研究相关方都会按要求对您的身份保密。录音文件转录和翻译的有关人员不接触受访者属性表格，仅承担匿名化后的文件转录和翻译工作，并且需在签署保密协议后方可承担相关任务。您的档案将妥善保存，仅供研究人员查阅。为确保研究按照规定进行，必要时，政府管理部门、学校当局或伦理委员会的成员按规定可以在研究单位查阅您的个人资料。这项研究结果发表时，将不会披露您个人的任何资料。

13. 如果我有问题或困难，该与谁联系？

如果您有与本研究相关的任何问题，请联系徐进，010-82805701；
xujin@bjmu.edu.cn

如果您有与参与研究者自身权益相关的问题，可与北京大学生物医学伦理委员会联系，联系电话：010-82805751，电子邮件：llwyh@bjmu.edu.cn。

研究者声明

我已告知该受访者“基层首诊制改革试点效果评估及其卫生体系发展的促进作用研究”研究背景、目的、步骤、风险及获益情况，给予他/她足够的时间阅读知情同意书、与他人讨论，并解答了其有关研究的问题。我已告知受访者可以退出本研究；我已告知该受访者将得到这份知情同意书的副本，上面包含我和他/她的签名。

获得知情同意的研究者签名及其联系方式

日期

受访者声明

我已被告知“基层首诊制改革试点效果评估及其卫生体系发展的促进作用研究”研究的背景、目的、步骤、风险及获益情况。我有足够的时间和机会进行提问，问题的答复我很满意。我已经阅读这份知情同意书，并且同意参加本研究的半结构式访谈。我知道我可以在研究期间的任何时候无需任何理由退出本研究。我被告知我将得到这份知情同意书的副本，上面包含我和研究者的签名。

受访者签名

日期

研究者声明

我已告知该受访者“基层首诊制改革试点效果评估及其卫生体系发展的促进作用研究”研究背景、目的、步骤、风险及获益情况，给予他/她足够的时间阅读知情同意书、与他人讨论，并解答了其有关研究的问题。我已告知受访者可以退出本研究；我已告知该受访者将得到这份知情同意书的副本，上面包含我和他/她的签名。

获得知情同意的研究者签名及其联系方式

日期

受访者声明

我已被告知“基层首诊制改革试点效果评估及其卫生体系发展的促进作用研究”研究的背景、目的、步骤、风险及获益情况。我有足够的时间和机会进行提问，问题的答复我很满意。我已经阅读这份知情同意书，并且同意参加本研究的关键组讨论。我知道我可以在研究期间的任何时候无需任何理由退出本研究。我被告知我将得到这份知情同意书的副本，上面包含我和研究者的签名。

受访者签名

日期