# The spatial distribution and epidemiology of trachoma: application and evaluation of geographical information in defining disease burden and planning control

Jennifer Linnea Smith

Department of Disease Control

Faculty of Infectious and Tropical Diseases

London School of Hygiene & Tropical Medicine, University of London

**2014**

## Declaration by candidate

I, Jennifer Linnea Smith, 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.

Signed _____ Date: _____

Jennifer Linnea Smith

## Abstract

The last decade has seen significant progress towards the elimination of blinding trachoma as a health problem. However, gaps in our understanding of the epidemiology of trachoma at large scales are increasingly important in the context of programmatic scale up. This thesis therefore aimed to define the current distribution and burden of trachoma, in addition to investigating the spatial heterogeneity of trachoma and underlying risk factors at different scales.

A systematic review of trachoma prevalence data was used to generate the Global Atlas of Trachoma, a unique spatially-referenced global database. In addition to highlighting important regional differences in the geographic distribution of trachoma, this database was used to quantify the disease burden in Africa; estimating nearly 153,000 disability-adjusted life years (DALYs) attributed to trachomatous vision loss and 155,500 additional DALYs to trichiasis. Detailed analyses of individual and cluster-level risk factors underlying the distribution of trichiasis in Nigeria and active trachoma in Kenya identified a number of key socio-demographic and environmental factors. Both analyses suggested that spatial dependency (generated by underlying associations with shared risk factors at larger scales) may vary in endemic areas. These findings emphasise the importance of local epidemiology and the need for robust and well-designed survey methodologies to identify areas of high risk. Computerised simulations were used to evaluate the performance of Integrated Threshold Mapping (ITM) in comparison to the accepted gold standard trachoma survey design. The results found that ITM tended to underestimate the prevalence of trachoma across a range of epidemiological contexts where attendance was low and/or the risk of disease was lower in school-going children.

This thesis provides the first systematic investigation into the geography of trachoma; highlighting heterogeneities at different scales and their potential programmatic implications. In particular, the findings and methods from this thesis may help to inform future survey design.

## Acknowledgments

First and foremost, I offer my sincerest gratitude to my supervisors, Professor Simon Brooker and Dr. Anthony Solomon, for generously sharing their time and invaluable expertise in support of this thesis. Their support, encouragement and guidance have provided a constant source of personal and professional motivation. In addition, I would like to thank the members of my advisory panel, Dr. Rachel Pullan, Dr. Hugh Sturrock and Dr. Artemis Koukounari. In particular, my thanks go to Hugh, for his enthusiasm and academic curiosity, which has made our work together so enjoyable and productive.

I would also like to thank the International Trachoma Initiative, who funded my position at the London School of Hygiene & Tropical Medicine through a grant from the Bill & Melinda Gates Foundation and were key collaborators on developing the Global Atlas of Trachoma. In particular, I am grateful to Dr. Danny Haddad and Rebecca Mann for their support and technical expertise, as well as their hospitality during my visits to Atlanta. Furthermore, I am grateful to the Division of Ophthalmic Services in the Ministry of Public Health and Sanitation, Government of Kenya, and Dr. Clare Gilbert for contributing data used to inform Chapters 4-6.

I am grateful for support from my fellow PhD students and colleagues at LSHTM, in particular Kate Halliday and Ruth Ashton, who have provided advice, friendship, and solidarity during my studies.

Finally, I would like to thank my husband Rob, who has been a constant (and patient) source of encouragement throughout my studies and shared in all of my frustrations and achievements.

# Table of Contents

# List of figures

# List of tables

## List of acronyms

| | |
|---|---|
| AIC | Akaike information criterion |
| ASAL | arid and semi-arid lands |
| ASTRA | acceptance sampling trachoma rapid assessment |
| AUC | area under the curve |
| BHM | Bayesian hierarchical models |
| CAR | conditional autoregressive |
| CBS | Central Bureau of Statistics |
| CI | confidence interval |
| CO | corneal opacity |
| CRS | cluster randomized sampling |
| DALY | Disability-adjusted life years |
| DIC | deviance information criteria |
| EB | elementary bodies |
| EVI | enhanced vegetation index |
| FAO | Food and Agriculture Organization |
| GAT | Global Atlas of Trachoma |
| GBD | Global Burden of Disease |
| GDP | gross domestic product |
| GIS | Geographical Information Systems |
| GPS | geographical positioning systems |
| ICC | intra-cluster correlation |
| IHME | Institute of Health Metrics |

| | |
|---|---|
| ITI | International Trachoma Initiative |
| ITM | Integrated Threshold Mapping |
| JMP | Joint-Monitoring Programme |
| KNTCP | Kenyan National Trachoma Control Programme |
| LGA | local government area |
| LSHTM | London School of Hygiene & Tropical Medicine |
| LST | land-surface temperature |
| MDA | mass drug administration |
| NTD | Neglected Tropical Disease |
| OR | odds ratio |
| PBL | WHO Programme for Prevention of Blindness |
| PBPS | population-based prevalence surveys |
| PCA | principal components analysis |
| PET | potential evapo-transpiration |
| PPS | probability proportional to size |
| RAAB | rapid assessment of avoidable blindness |
| RB | reticulate body |
| ROC | receiver operator characteristics |
| SAFE | Surgery for trichiasis, Antibiotics for infection, Facial cleanliness and Environmental improvement |
| SES | socioeconomic status |
| STH | soil-transmitted helminths |
| TF | trachomatous inflammation–follicular |
| TRA | trachoma rapid assessment |

| | |
|---|---|
| TT | trachomatous trichiasis |
| UIG | ultimate intervention guidelines |
| WHO | World Health Organization |
| YLD | years of life lived with disability |
| YLL | years of life lost to premature mortality |

# Chapter 1: Introduction

## 1.1 Background and context

Trachoma, a blinding eye infection, is a significant public health risk in developing countries and has caused an estimated 1.3 million current cases of blindness globally [2]. In 1998, the World Health Assembly called for the global elimination of trachoma as a public health problem by the year 2020 (GET2020). The specific targets are to reduce the burden of trachoma to less than one prevalent case of trachomatous trichiasis (TT) unknown to the health system per 1000 total population, and prevalence <5% "trachomatous inflammation–follicular" (TF) in children aged 1-9 years, at sub-district level [3]. To meet these targets, it will be necessary to scale up control activities in all endemic areas of every country by 2016-2018 in order to allow sufficient time for programme impact.

A broad consortium of partners have successfully mobilised political will and funding in order to identify key milestones in meeting this target and address current gaps in our knowledge. One critical limitation identified early in the process was that the exact distribution and disease burden is unknown in the majority of trachoma endemic countries. While many surveys have been carried out in the last few decades, some epidemiologic data are out-of-date while other areas suspected to be endemic remain unmapped. Cost-effective implementation of trachoma control relies on an evidence based understanding of the distribution of trachoma at different scales, and effective use of this knowledge to geographically target resources and inform survey design.

Spatially referenced trachoma prevalence data increasingly form the evidence base for a wide range of programmatic and epidemiologic activities. Maps of the current geographical distribution of trachoma allow identification of gaps where data are lacking, determine the geographic overlap of different Neglected Tropical Diseases (NTDs) and

provide a common platform to mobilise partners for action. While aggregated maps of district-level prevalence estimates within countries are used to target implementation of trachoma control, information at larger scales can also refine estimates of the burden of disease attributed to trachoma and broadly inform resource allocation. More recently, as the geographic coverage of underlying trachoma surveys have increased and mapping methodologies are standardised, maps have provided an important tool in visualising spatial heterogeneity, or variation, at smaller scales.

As a first step in interpreting patterns of disease distribution, it is important to understand heterogeneities and underlying determinants of the risk of trachoma in different contexts. Heterogeneity in risk between individuals, households and communities has long been recognised as a hallmark of the disease [4]. An extensive body of research has investigated sources of this variation at smaller scales, including individual-level risk factors associated with heterogeneities in the pathogenesis of disease and within-community patterns of disease and determinants of risk. However, only a few studies have systematically investigated spatial variation in risk at larger scales and identified underlying context specific associations. As well as adding to our epidemiological understanding of disease ecology in different contexts and at varying scales, associations may be useful in targeting survey activities to areas predicted to be at higher risk and better defining the burden of disease.

There is a critical need to rapidly and accurately scale up trachoma mapping to all endemic areas if the goals of control efforts are to be achieved by 2020. Thus, survey designs used for this baseline mapping must meet required standards of reliability and validity in order to inform control decisions, as well as context-specific restrictions on costs and feasibility. These requirements may change as control efforts are scaled up, and transmission levels are reduced. Consequently, evaluation of alternative survey designs represents an important and potentially expensive process over the course of a control programme. Furthermore, spatial characteristics of diseases, both in terms of variability and spatial

clustering, will likely have important implications for survey designs. With advances in spatial statistical methods, and an increasing amount of high resolution data, there is now an opportunity to quantifiably investigate and take into account spatial aspects of trachoma to assist in the evaluation and rational design of surveys.

The aim of this thesis is to describe the geographical distribution of trachoma at global, national and local levels to address current gaps in the evidence base, inform estimates of disease burden, and assess potential applications in disease mapping and survey design. This chapter provides an introduction to a number of topics that provide contextual information to the thesis. In particular, it provides an overview of the pathology and clinical features of trachoma, before describing current diagnostic and survey methods used to estimate the burden of disease. Next, the epidemiology of trachoma is reviewed in the context of differing spatial scales, followed by an introduction to the use of statistical tools for analysing spatial data and applications in disease mapping.

## 1.2 Overview of Trachoma

### 1.2.1 Biology and developmental cycle

Trachoma is a chronic bacterial conjunctivitis caused by *Chlamydia trachomatis*. A subset of four serovars (A-C) that are selective for ocular epithelial tissue are responsible for blinding endemic trachoma, while other serovars infect genital tissues and may occasionally cause self-limiting conjunctivitis [5,6]. As an obligate intracellular bacterium, reproduction is metabolically dependent on the host cell. Extracellular, infective stages are called elementary bodies (EB), which bind to and enter epithelial cells in the conjunctiva. Upon entry, the EB expands to form a larger reticulate body (RB) and begins transcription, causing the formation of inclusions within infected cells. Newly formed EBs rupture from the cell and are found in ocular and nasal secretions [7]. Both resolution of infection, which is age-dependent and typically occurs over a period between one to14 weeks [8],

and pathogenesis of disease are driven by complex inflammatory processes governed by the host's cell-mediated immune response [9].

### 1.2.2 Clinical features and natural history of disease

Trachoma encompasses a broad spectrum of clinical manifestations that include inflammatory signs of disease broadly referred to as "active trachoma" and later cicatricial changes resulting from multiple episodes of infection (Figure 1.1). Although these distinct signs generally reflect progress of the disease, they may overlap in an individual at a given time point. Many infections are symptomless or may present as a mild conjunctivitis and a single episode is unlikely to result in important sequelae [9,10]. However, prolonged infection causes clinical signs of active trachoma, including development of characteristic lymphoid follicles and an intense, progressively chronic, inflammation of the conjunctiva [9].

Increased inflammatory and follicular responses are believed to be dependent on a number of factors, including number of episodes of infection and bacterial loads [10-13], and are likely mediated by individual immune responses [11]. Increasingly, there is evidence that some individuals may develop a hypersensitivity to infection, so that clinical signs of active disease are sustained even without recent exposure [9]. Non-chlamydial pathogens may also play a role in development of clinical signs of active trachoma in the absence of infection [14,15].

Recurrent episodes of infection and associated inflammation are responsible for chronic disease states and disabling sequelae through progressive scarring of the conjunctival stroma, causing the lashes to turn inwards and scratch the cornea (trichiasis). This mechanistic pathway to blinding complications (Figure 1.1) is supported by several longitudinal studies, although there is considerable variation in disease progression rates.

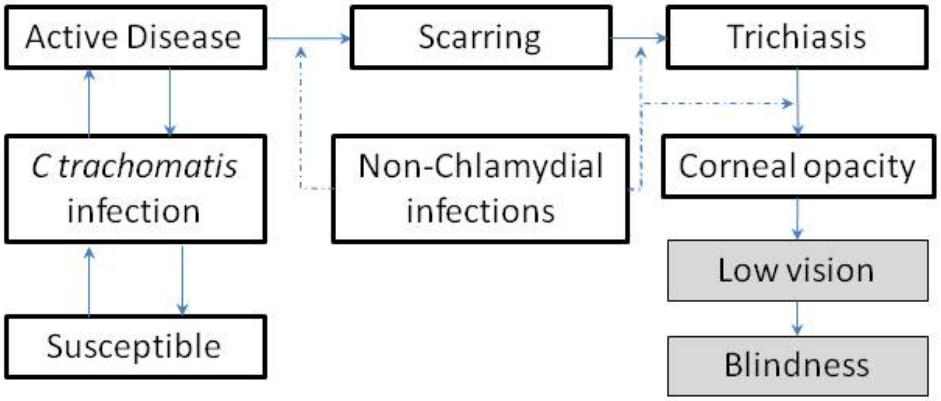



Figure 1.1 Simplified diagram showing progressive disease states and sequelae for trachoma. Developed for this chapter

Studies in Tunisia, Tanzania and the Gambia have shown that severe inflammatory disease was predictive of higher incident conjunctival scarring [10,16], which in turn is predictive of trichiasis [17] and subsequent corneal opacity [18,19]. The main factor determining variation in reported rates is likely to be the number and duration of infections experienced, with evidence supporting faster progression associated with increasingly severe inflammatory disease and persistent infections [10,20,21]. However, the role of host immune responses in disease pathogenesis remains incompletely understood and a growing body of evidence suggests that collateral damage caused by individual cell-mediated immune responses against chlamydial antigen may mediate tissue damage and fibrosis as well as age-dependent clearance of infection [8,22,23]. Other bacterial infections may contribute to the scarring process, with greater severity of trachomatous scarring associated with a higher prevalence of bacterial isolates in case-control studies in the Gambia and Tanzania [24,25]. As a consequence of the long development of chronic sequelae and its multifactorial aetiology, incident cases of trichiasis are expected to arise in the adult population for some time, even where control of infection in children is immediately effective, and contribute to subsequent vision loss.

### 1.2.3 Diagnostic methods

For programmatic purposes, cases of trachoma are diagnosed based on clinical signs of disease during an ocular examination, usually using the 1987 WHO simplified grading system presented in Figure 1.2 [26]. This system distinguishes between five distinct grades of trachoma which can be easily recorded in the field at low cost and where laboratory capacity may be limited. Typically, a minimum of two key clinical signs are collected for use in trachoma control programmes: trachomatous inflammation–follicular (TF) in children aged 1-9 years and trachomatous trichiasis (TT) in adults aged over 14 years. These two signs are directly relevant to programmatic action as they are interpreted as a proxy for active infection (TF, at least in highly endemic contexts) and the surgical burden (TT). Other signs distinguished in this grading scheme, in particular trachomatous inflammation – intense (TI) and trachomatous scarring (TS), are not collected as often, in part due to the greater inter-grader variability in grading these signs, and in part due to the lack of programmatic action mandated by high prevalence of these signs. Numerous more detailed grading systems that distinguish between major and minor forms of chronic disease stages exist [27,28], but are typically restricted to research studies tracking pathological process and increasing severity of disease. Signs of scarring could serve as a more sensitive indicator for monitoring the impact of control programmes on chronic sequelae, as a certain proportion of individuals with signs of TS would be expected to progress to TT over time.

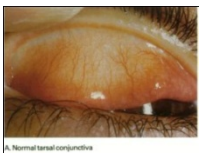
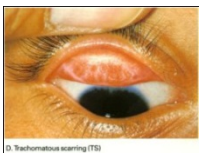
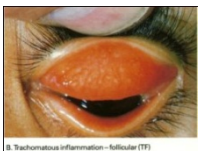
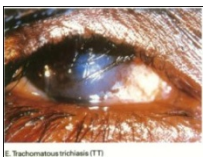
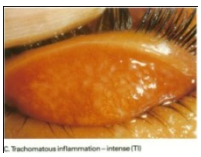
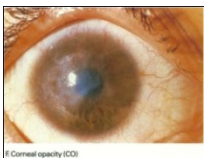
| Grade & Clinical signs | | Grade & Clinical signs | |
|---|---|---|---|
|  | **Normal** |  | **Trachomatous Scarring (TS)**<br>The presence of scarring in the tarsal conjunctiva |
|  | **Trachomatous inflammation – follicular (TF)**<br>The presence of five or more follicles in the upper tarsal conjunctiva |  | **Trachomatous Trichiasis (TT)**<br>At least one eyelash rubs on the eyeball or evidence of recent removal of inturned eyelashes. |
|  | **Trachomatous inflammation – intense (TI)**<br>Pronounced inflammatory thickening of the upper tarsal conjunctiva that obscures more than half of the normal deep tarsal vessels |  | **Corneal Opacity (CO)**<br>Easily visible corneal opacity so that at least part of the pupil margin is blurred when viewed through the opacity. |

Figure 1.2 WHO simplified system for community assessment of trachoma, depicting five clinical signs of trachoma. Adapted from Thylefors et al. (1987) [26].

The presence of TF is not considered a reliable marker of infection with *C trachomatis* [7,11,29,30] and increasing disparity is noted in areas of lower endemicity, particularly after mass drug administration (MDA) [31,32]. The specificity of clinical signs thus poises an increasingly important issue for trachoma control programmes attempting to target antibiotic distribution as the prevalence of trachoma declines. In some instances, TF may be present without detectable *C trachomatis* due to clearance of infection prior to resolution of clinical signs, or presence of non-chlamydial pathogens that may elicit a similar follicular reaction [30],[14,33]. Recent work using latent class analyses to quantify the performance of trachoma diagnostic tests have highlighted the low specificity of TF, which was reported as 36.6% in a hyperendemic, pre-treatment area in Ethiopia [34]. In contrast, the same analysis found that polymerase chain reaction (PCR)-based assays, which are commonly used as a diagnostic gold standard, provide a highly specific (100%) and reasonably sensitive (87.5%) test for *C trachomatis* infection. While PCR methods remain largely confined to research studies due to their high cost and technical requirements, recent field evaluation of various assays for ocular Chlamydia infection show promising results for monitoring treatment impact and targeting interventions in hypoendemic areas [35,36].

### 1.2.4 Control strategies

A multi-faceted approach to trachoma control is advocated, known as SAFE: Surgery to correct trichiasis, Antibiotic to clear *C trachomatis* infection, Facial cleanliness and Environmental improvement to reduce transmission. The strategy is targeted based on the prevalence of clinical signs of active disease (TF or TF/TI) in children aged 1 to 9 years and blinding sequelae (TT) in adults. While the first two components are typically implemented directly through national trachoma control programmes, behavioural and environmental improvements often require intersectoral collaborations and support from ministries of education, water and sanitation or rural development. As a consequence, in practice the SAFE strategy is not always implemented in full and the F & E components have lagged behind antibiotic distributions and surgical interventions.

Despite widespread support and a solid rationale for this strategy, the evidence base for individual SAFE components is relatively weak. Individuals with trichiasis are at immediate risk of going blind, and so an important component of SAFE is providing surgical interventions. Prevalence estimates are used to estimate the expected surgical burden for planning ophthalmic services, as programmes should aim to operate on all cases. Although surgeries have been shown to be effective at reducing discomfort associated with trichiasis, there is inconclusive evidence of their effect on preventing vision loss [37]. Surgeries have up to a 20-60% recurrence rate and uptake is poor in many communities [38,39]. The remaining components of this strategy and best methods for reducing the burden are aimed at preventing the blinding complications of trachoma as a result of repeated infection.

Two antibiotics are currently used to treat infection (oral azithromycin and topical tetracycline) and are administered through repeated mass distributions in order to reduce the community pool of infection and suppress re-emergence of infection [40]. While evidence of the impact of antibiotics on infection is relatively robust, there is less evidence of a corresponding impact on signs of clinical disease and high levels of heterogeneity in

estimates of effect for both outcomes. The recent update to the Cochrane review reported estimates of the relative risk (RR) of active trachoma from individually randomised antibiotic treatment trials to be 0.78 (95% 0.69 to 0.89) and 0.74 (95% 0.55 to 1) at 3 and 12 months following antibiotic treatment. Estimates of the effect on active trachoma from community randomised trials (using oral azithromycin) are scarce, but a large study randomising 24 communities (TANA 2009 study) reported a 95% confidence interval of the relative risk of active trachoma ranging from 0.47 to 0.72 at 12 months following treatment [41]. Estimates of the effect of antibiotic treatment on infection with *C trachomatis* from individually randomised trials are similar (3 months: RR 0.81, 95% 0.63 to 1.01; 12 months: RR 0.25, 95% 0.08 to 0.78), while community randomised trials provide stronger evidence of an effect (RR 0.35, 95% 0.21 to 0.60) despite considerable heterogeneity in relative risk estimates and quality of evidence [42].

Variation between estimates of effect between studies may arise from differences in the trial designs, populations and compliance. Although both antibiotics have been found to be effective against *C trachomatis*, single-dose oral azithromycin is suggested to have higher compliance and potentially greater impact on TI [43,44], treat extraocular reservoirs of infection [45], and reduce overall mortality in young children [46]. The possibility of local elimination of infection has been demonstrated with moderate to high coverage of repeated treatments of azithromycin, even in hyperendemic communities [47,48].

While local elimination of infection might be achieved through frequent administration of antibiotic distribution alone at sufficiently high coverage, its reintroduction poses an unknown threat to control efforts. Reintroduction of infection after MDA through travel between villages has been observed in the Gambia [49]and Ethiopia [50,51], but not in Tanzania [52]. A number of factors are likely to influence the likelihood of this being an important source of re-emergence, including initial endemicity levels, scale of treatment, coverage and contact patterns between communities. To successfully eliminate blinding

trachoma, it is probably not necessary to eliminate all infection or all active disease (as recognized by the 5% prevalence targets), but rather keep transmission below an undefined threshold. Promotion of facial cleanliness and environmental improvement through provision of latrines and increasing water availability is aimed at making more sustainable changes in transmission, although the evidence base for these components is fairly weak [53-55].

The prevalence of clinical signs of disease are used to guide the planning and implementation of control strategies at the district (second administrative) level, as outlined in Table 1.1, with subsequent focus on sub-districts as appropriate [56]. For example, district-wide mass antibiotic treatment is recommended for districts where the prevalence of active trachoma is 10% or more among children aged 1-9 years. In districts with prevalence between 5% and 10%, sub-district level assessment and treatment is advocated. Clinical indicators are also used to define GET2020 ultimate intervention guidelines (UIGs), signifying when trachoma has been eliminated, which are less than one case of TT unknown to the health system per 1000 total population and <5% TF in children aged 1-9 years , at the sub-district level [3].

Treatment thresholds are currently based on expert option, rather than empiric studies. While longitudinal trials and post-impact intervention surveys have provided support for more than 3 years of treatment in highly endemic contexts [57,58], the lower (10%) threshold is increasingly questioned due to the large discrepancy between clinical signs and infection in low endemicity settings [59]. MDA stopping points, represented by the 5% threshold, have been the subject of a recent mathematical modelling studies by Ray et al. (2009) that found a graduation strategy based on this threshold to be reasonably effective in hypo- and meso-endemic areas but more variable in hyper-endemic areas [60]. These results highlight the sensitivity of stopping points to different transmission parameters, likely influenced by behavioural (i.e. contact patterns between and within communities), as well as environmental factors.

Table 1.1 Endemicity classes for implementation of SAFE based on trachomatous inflammation–follicular (TF) and trichiasis (TT).

| TF Prevalence band | Classification | Implementation |
|---|---|---|
| <5% | Non-endemic | No need for implementation of AFE |
| ≥5% and <10% | Hypo-endemic | Mapping, F and E can be applied, focal A |
| ≥10% and <30% | Meso-endemic | AFE at district level (≥3 years then review) |
| ≥30% | Hyper-endemic | AFE at district level (≥5 years then review) |
| **TT Prevalence band** | **Classification** | |
| <0.1% | UIG achieved | |
| ≥0.1% | UIG not yet achieved | |

## 1.3 Defining the distribution of trachoma

### 1.3.1 Global distribution of trachoma

The global distribution of trachoma has evolved over time as a reflection of population movements, including through trade routes and military action, as well as economic and social development. MacCallan (1931) recounted evidence of its presence dating as far as 1800 BC in Egypt, where early epileption forceps have been found [27], and various references document its presence in antiquity across Asia, Greece and the Middle East [61,62]. Trachoma is thought to have been introduced to Europe in 13$^{th}$ century through the Crusades and became entrenched in preindustrial cities following the return of soldiers from Egypt during the Napoleonic wars [62,63]. During the 1800s and early 1900s, trachoma was a public health problem in much of Europe and parts of the United States, but the disease was eliminated from these countries as a result of general socioeconomic improvement and specific public health measures including ophthalmic hospitals, boarding schools for infected children, immigration control and treatment with sulfa antibiotics [64,65]. Thus, the wide-ranging distribution of trachoma does not lend support to the idea of climatic limits, whilst both geographical heterogeneity and disease persistence have traditionally been associated with socioeconomic development and inequalities at various scales.

Currently, the WHO classifies 53 countries as endemic for trachoma, which include areas of Africa, Latin America, South East Asia, the Eastern Mediterranean and the Western Pacific (Figure 1.3) [66]. An endemic disease is usually defined as one that is present in a population or area at all times, although in practice most countries that are classed as endemic are those where trachoma has been found to be a public health problem or are lacking data to the contrary. Many countries where early studies of trachoma were focused are now considered post-endemic, including Saudi Arabia, Lebanon, Tunisia, and Palestine. Some countries which remain classified as trachoma endemic (such as India and Namibia) have experienced significant economic growth and socioeconomic changes; however no recent data are available to support a reduction in prevalence. Trachomatous blindness has been found by national blindness surveys in some countries currently considered to be non-endemic (including Tonga and Occupied Palestinian Territories), while cases of trichiasis have been found in Rwanda as part of rapid assessment surveys.

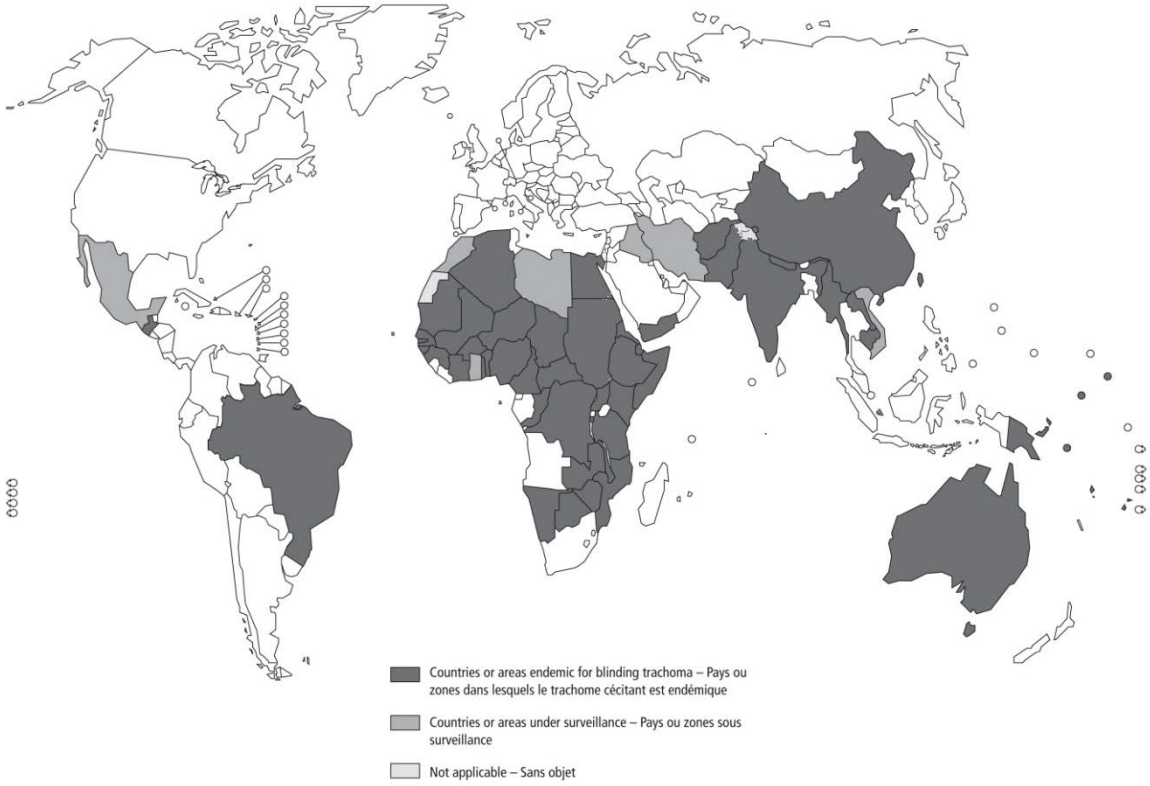



Figure 1.3 Map of countries classified as trachoma endemic by the World Health Organization. Adapted from World Health Organization (2012) [66].

### 1.3.2 Disease mapping

Several global maps of trachoma have been developed over the past 100 years and these have been important for documenting the distribution of the disease at different time points. Figure 1.4 shows the first known world map of trachoma which was developed in 1929 by the International Council of Ophthalmology [67]. This map was based on data collated from existing ophthalmic societies and shows trachoma still present in Europe and North America. Later global mapping efforts by the WHO in 1949 and 1961 showed trachoma disappearing from northern Europe, but remaining widespread in much of Africa, Asia, the Pacific and some countries of Latin America [68,69]. Common to all these early mapping efforts is that they described broad areas/countries where endemic trachoma was found, however, they did not indicate the levels of trachoma prevalence or variations in prevalence within and between countries.

Figure 1.4 World map of trachoma developed in 1929 by the International Council of Ophthalmology Block colour indicates 'The approximation is of a certain exactness'. Hashed colour indicates 'The approximation is more roughly'. Circles are 'figures of American Indians'. The percentage categories are as follows: 0-0.1% (green), 0.1-1.9% (yellow), 2-4.9% (orange), 5-10.9% (pink), 11-30.9% (purple), 31-30.9% (brown), 61-100% (grey), unknown (white)

In 2005, the London School of Hygiene & Tropical Medicine (LSHTM), in collaboration with WHO Prevention of Blindness and Deafness and with support from the International Trachoma Initiative (ITI), created the first global trachoma atlas displaying systematically collated district-level prevalence data [70]. These maps highlighted the variation in trachoma prevalence within and between countries, but also the paucity of reliable data in many trachoma endemic countries, particularly in Africa. The development of Geographical Information Systems (GIS) - computerised systems for managing, analysing and mapping disease information – together with standardisation of trachoma survey methodology has facilitated efforts to develop a framework for trachoma mapping. Increasingly, high resolution trachoma prevalence data are collected through the use of smartphones with geographical positioning systems (GPS), providing the potential for more detailed and accurate mapping of disease distribution at various scales. Increasingly, the term "disease mapping" is used interchangeably with "baseline surveys", reflecting the growing recognition of the use of maps to inform control decisions – and the ease with which modern technology allows survey data to be displayed in map form.

### 1.3.3 Global disease burden

Within the context of a control programme, the burden of trachoma is typically defined as the prevalence of active trachoma and trichiasis. However, it is well recognized that prevalence estimates do not capture the full burden experienced by the individual over their lifetime as a result of a disease. For example, a condition which is severely disabling or of long duration might be considered to confer a larger burden than one that is of shorter duration or causes no discomfort. In order to compare the burden attributed to different diseases, various summary measures are used to quantify this disability. The disability-adjusted life year (DALY) is a commonly used measure for setting health research priorities, used by the Global Burden of Disease (GBD) study, and is calculated as a weighted measure of morbidity and mortality [71]. Specifically, the DALY is the sum of

the years of life lost to premature death (YLL) and years of life lived with a disability (YLD). YLLs are computed by multiplying the number of deaths at each age by the corresponding standard life expectancy, while YLDs are calculated as the prevalence of a given disease and injury sequelae multiplied by a disability weight. Weights represent a measure of the severity of health lost and range from 0 to 1, where 0 is commensurate with perfect health and 1 is commensurate with death [72]. The current estimation of DALYs differs from the original study in 1990, primarily due to how YLDs are calculated. Previously, estimates of YLDs were based on incidence, rather than prevalence, and incorporated age-weighting and discounting to reflect respective social valuation of years of life at different ages and the present time over the future [73]. More recent YLD estimates are based on prevalence data and do not use discounting.

DALYs attributed to trachoma are based solely on the morbidity associated with resulting visual impairment. Cause-of-blindness surveys are used to generate trachomatous blindness prevalence estimates, which are the epidemiological data used to estimate disability. These are population-based prevalence surveys that typically use probability proportional to size (PPS) to ensure population representativeness (see section 1.4). Rapid Assessment of Avoidable Blindness surveys (RAAB) are increasingly common, which are usually conducted in representative populations aged 50 and above years, using a streamlined examination technique [74]. In both of these protocols, visual acuity is typically measured using Snellen E optotypes or Landolt C charts at specified distances, where low vision is typically defined as best corrected vision being < 6/18 and blindness as 3/60 [75]. While this definition of vision loss is typically used in trachoma endemic counties, with the exception of India, many more developed countries use 6/60 [76]. Where vision is less than 6/18 in either eye, the cause of vision loss is diagnosed based on a detailed ophthalmic examination using a slitlamp or reflected light [77]. Trachoma is most commonly reported as a cause of blindness where there is central scarring in the presence of trichiasis/entropion, conjunctival scarring, pannus or Herbert's pits [77]. Often more than one contributing cause can be diagnosed, but only one principal cause is

reported based on which is readily curable or most easily preventable. As cataract is widely reported to the most common cause of blindness in Africa and Asia, and is prioritised above trachoma based on these criteria, the burden of blindness attributed to trachoma based on causes-of-blindness surveys is likely to be an underestimate [78].

A first seminal effort to define the global prevalence of trachomatous blindness was provided by Ranson and Evans (1995) for the 1990 GBD study [79]. This study has informed subsequent estimation strategies [80,81], with the most recent GBD 2000 estimates provided by WHO [82]. Unfortunately, epidemiological data on the prevalence of trachomatous blindness are limited and national surveys on causes of blindness are even scarcer. This lack of data has been overcome by either (i) extrapolating the results of national surveys to all endemic countries within a world region [79,80] or (ii) modelling the data using national gross domestic product (GDP) estimates to provide estimates for countries that are lacking blindness data and those without data at multiple time points [82]. Regional extrapolation makes use of the more similar socioeconomic and climatic conditions in countries within a region, but ignores the sizeable variation that often exists between neighbouring countries.

There are a number of clear limitations of current estimates of the global burden of trachoma. First, chronic stages of trachoma are not included in current global blindness estimates or modelling strategies despite some evidence that trichiasis imposes functional limitations on activities [83]. Longitudinal studies measuring subjective assessments of pain and photophobia have shown an improvement following TT surgery, and support an added impact on quality of life [84]. These limitations may confer an economic burden that is not captured in current measures. Second, the scarcity of data in space and time limit current methods to estimate the burden of trachomatous blindness. Finally, existing causal blindness surveys are powered for all cause blindness, rather than trachomatous blindness, and limit the precision of cause-specific estimates.

## 1.4. Survey design

### 1.4.1 Current trachoma survey methodologies

In order to target interventions appropriately, surveys must use reliable indicators of disease and be designed so that the sample is both representative of the underlying population and able to correctly identify districts requiring treatment. Surveys assessing the burden of trachoma use one of four methodologies: population based prevalence surveys (PBPS); trachoma rapid assessment (TRA); acceptance sampling trachoma rapid assessment (ASTRA); or "Integrated Threshold Mapping" (ITM) [85]. PBPS are the recommended method since they provide a representative measure of the prevalence of trachoma within a population, and are currently the basis for targeting SAFE interventions according to treatment thresholds.

The most common PBPS strategy is cluster randomized sampling (CRS) which uses a representative, two-stage sampling methodology to provide a "gold" standard prevalence estimate for each district. This methodology often selects clusters using probability proportional to size (PPS) in order to provide a representative population estimate at this level. General guidelines recommend sampling 20 clusters per district, although this varies in practice. Sample size calculations are based on an expected prevalence of active trachoma of 10% in children aged 1-9 years, and use a design effect of four to account for expected clustering within the survey population [56].

More recent PBPS in Kenya have defined geographically smaller evaluation units and reduced the design effect to two, based on the assumption that risk is more homogenous within these areas [86]. Estimates of TT tend to be less precise due to their lower prevalence, although the precision of district-level prevalence estimates is rarely reported. While the majority of surveys continue to sample adults over the age of 15 years, some protocols are being adapted to restrict sampling to higher age groups [87]. Although this

will provide a more precise estimate for the same sample size, it depends upon an accurate correction factor to calculate the number of all-age cases.

TRA was developed as a rapid and inexpensive method using convenience sampling to rank communities in terms of priority for control programmes [88]. TRAs are intended to be "optimally biased" to find trachoma where it is endemic and do not provide a reliable estimate of trachoma prevalence in the overall population. Although it is generally assumed that a negative TRA reliably identifies the absence of trachoma, sampling is targeted to populations assumed to be at the highest risk and thus estimates of risk and subsequent rankings are reliant on the accuracy of the informant and noted to be highly subjective and variable [89,90]. Increased discrepancy between PBPS and TRA rankings has been reported where prevalence is low [91], and a validation study in China by Liu et al. found that the presence of TF was not detected in several sites where it was found by PBPS [92]. ASTRA is a form of lot quality assurance sampling and can reliably classify communities in relation to a threshold value [93], but has in practice rarely been used as it requires modification to derive overall population estimates of trachoma prevalence.

More recently, ITM has been developed as a rapid and cost-effective alternative design for undertaking baseline assessment of populations for trachoma and other NTDs simultaneously. This methodology employs convenience sampling of school children, pre-school children and women of child-bearing age [94]. A minimum of 20 sites are randomly selected from the district (with a minimum of two per subdistrict), and an equal number of children aged 1-5 years and 6-9 years are sampled from each site to make a total of 50. This methodology was piloted in Mali and Senegal [95], and it has since been used in a nationwide NTD mapping in Togo [96]. The key concern cited regarding the use of ITM for trachoma surveys is the use of a school based platform. A study by Courtright et al. (1991) highlighted the limitations of a school-based approach in Egypt, and found that only 50% of pre-school age children were in a household with a trachomatous school-age sibling and not all children attended school in this context [97]. School attendance is likely to be low in

many endemic settings in Africa and, furthermore, the sample pool of young children and adults relies on active engagement from the community.

### 1.4.2 Evaluation of survey designs

Surveys to estimate trachoma prevalence are important in determining appropriate control strategies, yet large-scale evaluation of alternative survey designs may be prohibitively expensive. Typically, an internal validation is carried out in multiple settings, in which the new methodology is compared to an existing "gold standard". Ideally, this gold standard would be the true prevalence, and the performance of surveys could be assessed in terms of reliability of district classification, which is directly relevant to control thresholds. In practice, it is unfeasible to generate gold standard data of that type, and so validation studies use the recommended survey strategy (CRS) for comparison and are restricted in geographical scope. Validation studies for TRA and ITM have all been based on relatively few clusters within a subdistrict [90], district [91,95] or province [92], although one country-wide evaluation of ITM was carried out in Togo [96]. Clearly, these studies are limited by choice of a gold standard, which is subject to sampling error, and cannot allow full exploration of the performance of sampling methodologies in different contexts.

A relatively unexplored method for evaluation of NTD survey designs, and one that can help to address some of the limitations of trachoma survey evaluations, are computerised sampling simulations. A simulation approach can been used to solve a wide variety of computational problems and provide a convenient platform for evaluating more complex survey designs, particularly where there is an interest in understanding what factors impact on survey performance. In this thesis, simulation methods are used in two different ways: 1) as a computational tool that serves as an alternative strategy to frequentist statistics in solving complex models, where a direct analytic solution may not be possible (as used for risk analyses) and 2) in order to produce realistically varying data and

mechanically test a specific sampling procedure under different conditions. Geostatistical and statistical algorithms can be used to simulate a realistic gold standard disease data that reproduce spatial variation in risk at different scales and aspatial variation within various demographics or subpopulations. Sampling can then be repeatedly carried out, using Monte Carlo simulations, according to defined protocols. There are many advantages to this approach. First, subsequent comparisons between different survey designs can fully incorporate sampling error. Second, key parameters can be varied to assess their impact on survey performance. Third, the use of Monte Carlo methods allows calculation of probability. Finally, data and sampling simulations can be conducted at larger scales, in order to quantify the performance of a survey methodology in correctly identifying areas requiring treatment. Sampling simulations offer a cost-effective means of comparing survey methodologies in different endemic settings, and have been utilized in animal populations [98] and more recently in human populations [99-101]. With an increasing amount of geographically referenced data on the distribution of trachoma in populations at multiple scales, this approach has enormous potential to explore the performance of alternative survey designs in different contexts.

## 1.5 Epidemiology of trachoma

The following sections provide an overview of the principles behind the ecology of disease and their application to the epidemiology of trachoma. This is followed by a review of the evidence base for specific risk factors for trachoma and pathways through which they may influence the distribution of disease at different scales.

### 1.5.1 Ecology of disease and spatial scales

Spatial heterogeneity, or variation in an outcome or process through space, is found throughout nature; indeed true uniformity or randomness is rarely observed. From an

ecological perspective, the geographical distribution of disease is determined by variation in underlying factors which create suitable environmental niches for transmission. Some of the observed variation in disease prevalence is structured by spatial location, due to factors that are themselves clustered at different spatial scales and affect transmission. The remaining variation is aspatial and should, by definition, be distributed randomly between sites without any quantifiable spatial structure (e.g. personal hygiene).

The issue of scale is an important one, as relevant processes influencing the distribution of disease will depend on the level of the analysis and the influence of unpredictable, stochastic events will be less at larger scales [102]. Spatial patterns are described by variation that may be classified as macro-scale (large-scale and usually caused by risk factors that form a gradient across wide areas), meso-scale (local dependence structure caused by risk factors that cluster within a country) or micro-scale (occurring within meso-clusters and caused by differences in risk factors at a very small scale, such as at the household level)[103].

### 1.5.2 Overview of risk factors for trachoma

The spatial distribution of trachoma is likely to arise from complex relationships between risk factors for trachoma at different scales, as illustrated in Figure 1.5. Micro-scale risk factors are those which are associated with transmission at the individual or household levels, including genetic factors that are likely to influence immune responses, hygienic behaviours and differences in contact patterns or exposures correlated with sociodemographic factors. Meso-scale risk factors vary at a larger scale, between villages or subnationally, and may include water availability, latrine coverage, livestock and access to treatment facilities. Macro-scale risk factors operate over a larger scale and include climatic factors such as rainfall, aridity and temperature that may influence risk factors at other levels.

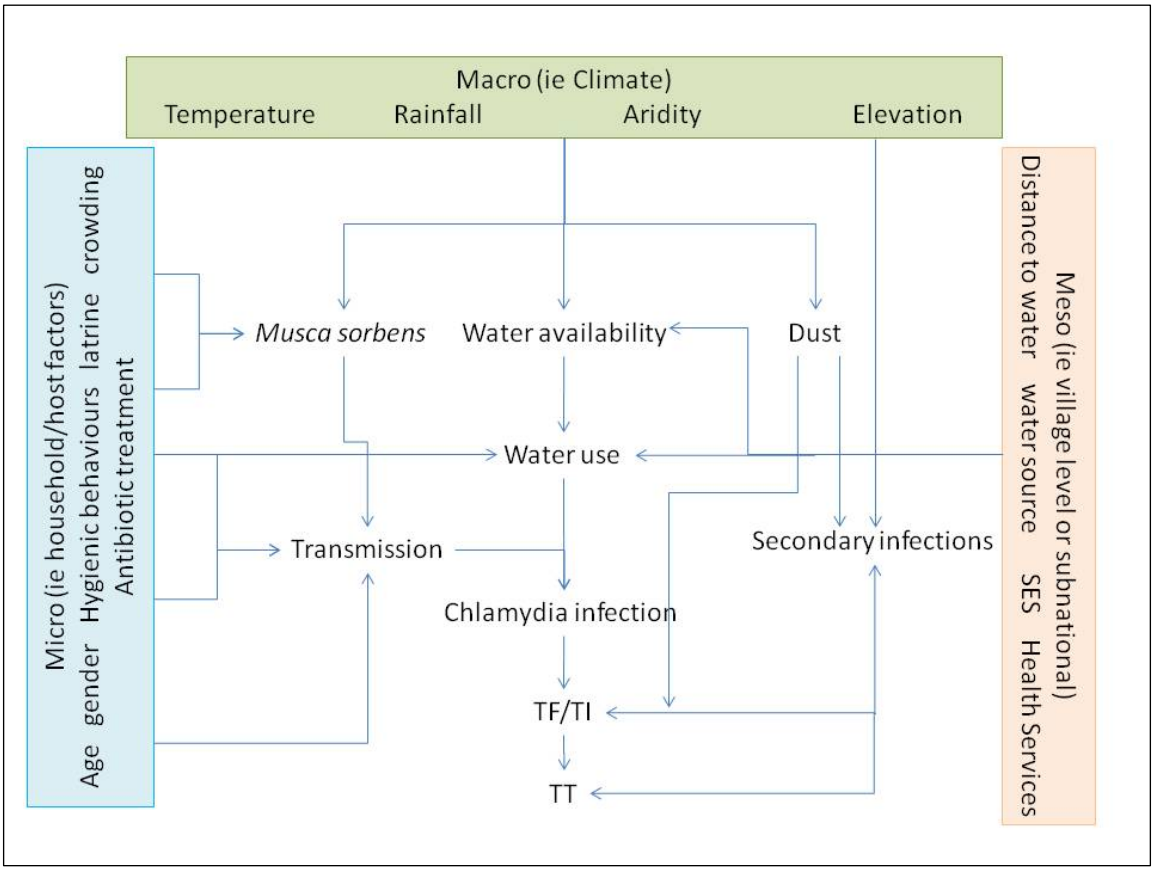



Figure 1.5 Diagram of factors at different scales associated with risk of trachoma. Genetic risk factors and TS are not shown. TF/TI: trachomatous inflammation-follicular/ trachomatous inflammation-intense; TS: trachomatous scarring; TT: trachomatous trichiasis; SES: socioeconomic status. Figure developed for this chapter.

Small scale patterns of the risk of trachoma develop as a result of exposures to determinants of infection within households, as well as stochastic contact patterns and individual immunological differences. Over larger scales, spatial dependency is more likely to reflect shared exposure to underlying environmental determinants, and may mask stochastic variation occurring at smaller scales. However, it is noted that transmission dynamics (and underlying contact patterns) between communities undoubtedly plays a role in reinfection after MDA in some contexts [49] and will influence subsequent spatial patterns at varying scales [104]. Risk factors at multiple scales will influence the same transmission pathways; with macro-scale risk factors such as climatic conditions mediated by risk factors operating at smaller scales (including access to treatment and behavioural differences) potentially breaking up large scale patterns of risk.

### 1.5.3 Transmission pathways and small-scale dynamics

As described in section 1.2.1, infectious stages of *C trachomatis* are found in ocular and nasal secretions. Mechanisms of transmission include direct contact or indirect transmission through fomites such as towels and other shared objects or transmitted via eye-seeking flies [27]. Although the relative importance of these pathways is likely to vary in different contexts, transmission requires relatively close contact which defines observed epidemiological clustering of disease at small scales [105]. Active disease has been observed to cluster within households and bedrooms [11,106-108], which is consistent with findings of higher risk in individuals cohabiting with an infected individual at baseline and after MDA [49,97]. Mathematical modeling studies have supported these findings, and suggest that an average of 71% of incident infections result from transmission within the household [109].

### 1.5.4 Individual risk factors

The prevalence of active disease (trachomatous inflammation) is consistently found to be highest in young children, who also have a greater bacterial load, particularly in highly endemic contexts, and are likely to be an important source for transmission of infection [11,13]. Increased prevalence in this age group is likely to reflect closer contact patterns and poorer hygiene of small children, as well as higher rates of clearance in older individuals resulting from accumulated exposure [8,110]. Progressive scarring results from cumulative infections, consequently entropion and trichiasis are more prevalent in older age groups [18,19,111].

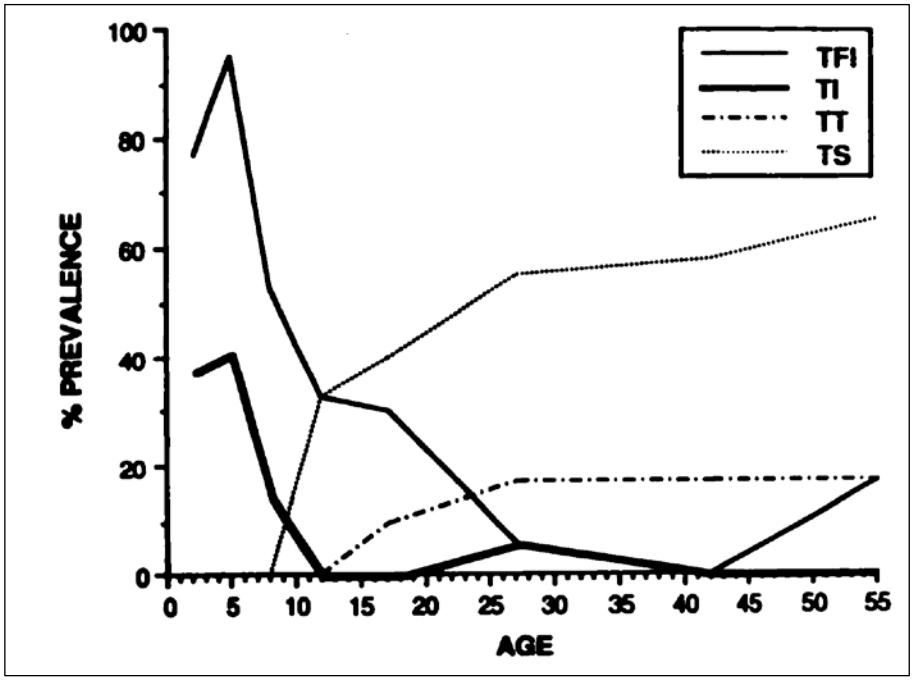



Figure 1.6 Age-prevalence curves of clinical signs of trachoma. Taken from Taylor et al. (1992) [112]

Females are often observed to have higher risk of trichiasis compared to males across different contexts, with a recent meta-analysis by Cromwell et al. (2009) [113] estimating an increased odds of 1.8 for trichiasis (95% CI 0.40 – 1.73). The variability in risk associated with gender is hypothesized to relate to division of labor and subsequent differences in contact with young children [107,114].

Finally, an increasing body of research supports a genetic basis for observed individual differences in persistence of inflammatory disease and intrinsic variation in susceptibility to infection, clearance and progression to chronic disease states [9,23].

Individual and household-level hygienic behaviors are expected to be proximal factors that influence transmission of *C trachomatis*, although their measurement presents a challenge for epidemiological studies. Routine face, hand and laundry washing should decrease discharge present on the face, fingers, and the personal environment; thus reducing spread of infection through fomites and eye-seeking flies and the likelihood of auto-reinfection [54,115]. Self-reported measures may be biased, due to the perceived

desirability of the hygienic behavior, but more objective measures of facial cleanliness have also been frequently associated with individual risk of active trachoma in observational studies in different settings [59,115-120].

Clearly, risk factors discussed above influence transmission between individuals and correlate with variation in risk of trachoma at micro scales (i.e. individual or household level). Access and use of antibiotic treatment will also influence risk of transmission and disease at this level. Although active trachoma typically causes minimal discomfort (compared to other ocular bacterial infections) and is unlikely to prompt health-seeking behaviour, treatment for other bacterial infections (including pneumonia, typhoid and gastroenteritis) with azithromycin and other antibiotics will also be effective against *C trachomatis*. Individual behavioral factors also provide a potentially important link to environmental risk factors at larger scales, as discussed in subsequent sections.

### 1.5.5 Environmental risk factors

*Water*

Particularly in underdeveloped or resource constrained settings, access to water for hygienic purposes represents an initial constraint on its use. Following basic availability, important determinants of water allocation within a household may include ease of access, the quantity and quality of water available, and priorities around water use. The complexity of relationships between water and its use for hygienic purposes is reflected in inconsistent associations between trachoma and different measures of water availability. Type of water source has found to be a risk factor in some [116,121,122], but not all [123,124], studies. A more proximal risk factor is likely to be the allocation of water within households, which may be better predicted by collection time, distance to or quantity of water available depending on the context [121,123,125,126]. However, there is substantial variation in how studies have defined water sources for analyses, with some

comparing piped water to all other sources [115,122] and others comparing protected and unprotected water sources [127,128].

Where a reliable water source has been put in place, subsequent behavioral changes around water use are likely to be mediated by the value that is placed on the water that is collected and perception of how much water is needed for washing in different contexts [123,129-132]. Although intensive participatory strategies to increase face washing has shown an improvement on facial cleanliness[133], subsequent randomized community based trials of a behavior modification campaign in Tanzania and Ethiopia provided limited evidence of impact on the prevalence of TF [53,134]. Clearly, modifying behaviors around water use in contexts where access is not a constraint presents a key challenge in achieving a sustainable impact on trachoma.

Water allocation for hygienic purposes within a household will be an important risk factor determining variation at micro scales, and potentially mediate the observed effect of differences in access to water between households and communities at micro and meso scales. In the absence of water and sanitation interventions, it is likely that natural resources will determine common exposures around water availability at larger scales and in some contexts may also influence hygienic behaviors. It is hypothesized that climatic factors may mediate the observed impact of interventions to improve domestic water access, which would be expected to vary based not only on the type of intervention but also existing physical and social environmental conditions [129]. For example, one would expect a greater impact on water availability and disease prevalence in semi-arid areas, where water is scarce, compared to the same intervention in a water-rich environment.

*Flies*

The Muscidae fly *Musca sorbens*, which is strongly attracted to and feeds on human secretions, acts as a vector in the transmission of ocular *C trachomatis* in some contexts.

Flies have long been associated with transmission of bacterial conjunctivitis, although much of the seasonal ophthalmic diseases observed were attributed to gonococcal infections [27]. *M sorbens* is widespread throughout Africa, Asia, and the Western Pacific and observations on its ecology, as well as experimental work, have associated its survival and relative abundance with variation in climatic factors [135]. Increased fly densities have been associated with outbreaks or an increased risk of trachoma in Ethiopia [136], Morocco [137], Tanzania [138] and Egypt [139]. Fly density has in turn been associated with increased number of flies on the face [140], which is a common risk factor for active trachoma [141]. Work by Emerson et al. (2000) definitively demonstrated the ability of *M sorbens* to carry *C trachomatis* in The Gambia, by identifying the bacteria on flies caught from the eyes of infected children[142].

Human faecal material in and around the household introduced by open defecation provides the favored breeding material for *M sorbens* [143], and its presence has been associated with an increased prevalence of active trachoma and infection [128]. Use of a pit latrine should reduce fly density around households, as larval stages have not been found in latrines nor have adult flies been caught emerging from them [142,144,145]. Access to a latrine has been associated with a lower prevalence of active TF in Egypt, Malawi, and Ethiopia [11,49,97,115`], although not in Tanzania [119]. Similarly, living in close proximity to livestock has also been associated with an increased risk of trachoma in Tanzania and South Sudan, and may be related either to increased fly density or act as a marker for other socioeconomic risk factors [118,126,138].

Despite the above evidence linking fly density with trachoma and latrines, the impact of fly control through spraying or latrine provision on trachoma has not been shown conclusively. Although a pilot cluster-randomised controlled trial of insecticide spraying showed a 61% reduction in trachoma prevalence, it was limited by a small sample size and included only 7 clusters in each arm [146]. A randomized controlled trial of fly control and latrine provision demonstrated that fly-eye contacts were reduced both by spraying

(88%) and latrine provision (30%), but a reduction in new prevalent cases of TF was statistically significant for the insecticide intervention group only [147]. A cluster-randomized trial of latrine promotion in 24 communities in Ethiopia showed no impact on the prevalence of ocular chlamydial infection after two years [148], although a subsequent secondary analysis associated uptake of latrine use with a greater reduction in trachoma that was significant for infection but not TF/TI [149]. Inconsistent associations between observational studies and variation in impact within intervention sites may be related to variable use of latrines or insufficient latrine coverage within communities to have an effect on fly density [150,151]. Furthermore, the relative importance of routes of transmission is likely to vary by settings and may reduce any observed impact of fly control where other pathways are more important.

Where flies provide an important pathway for transmission of *C trachomatis*, the above risk factors may be associated with variation in trachoma at multiple spatial scales. Fly eye contact at the individual level will be influenced by the presence of secretions on the face, which in turn are determined by individual hygienic behaviors. Behavioral factors such as open defecation may promote higher fly densities and clustering of disease risk within households and communities, while community-level latrine coverage may offer protection at meso scales. At macro scales, climatic factors may affect transmission through influence of fly density [135,142]. Currently, no evidence exists for a climatic influence on the survival of infectious elementary bodies on flies or surfaces.

*Climate*

Based on the historical distribution of trachoma and its current presence in climates ranging from humid and wet (Amazonas, Brazil) to hot and dry (Egypt), it is implausible that this represents a constraint on risk of infection. However, the links between climate and more proximal risk factors for transmission provides a basis for these factors to

influence large-scale trends in the distribution of trachoma. There may also be a potential role for climatic factors to directly affect disease progression in trachoma endemic areas, through ocular dryness or irritation [18,19].

Variation in active trachoma within continents and countries has been previously associated with broad scale geographic and climatic characteristics, including temperature [152,153], aridity and rainfall [154-158], and altitude [136,159]. On this macro scale, factors are likely to influence transmission dynamics through water availability and as determinants of fly density [160]. As noted in the recent review by Ramesh et al. (2013), studies that have investigated large scale associations with climatic factors are often limited by low power [136], statistical methods used [157,159], or have inadequately controlled for socioeconomic and environmental confounders [161]. The most robust study was conducted by Hagi et al. (2010) [152], which investigated clustering of active trachoma among 210 villages in Mali using a Bayesian hierarchical model, finding nearly 40% of observed variation was attributed to the village level and identifying four explanatory climatic risk factors. This study established strong associations between macro-scale variation in risk factors and disease risk, after controlling for numerous risk factors at the individual, caretaker, household and village level.

Although differences in water availability and subsequent allocation of water for hygiene between seasons is possible, there is limited evidence to support seasonality of transmission of *C trachomatis* or active trachoma. A longitudinal study by Holm et al. (2001) noted seasonal fluctuation in active trachoma prevalence corroborating anecdotal reports of seasonality in Nepal [162], and observed in Morocco and Australia [137,163,164]. While it is likely that fly density undergoes seasonal variation, this was not linked to trachoma prevalence in Ethiopia [136].

### 1.5.6 Socioeconomic risk factors

Disparities in socioeconomic status (SES) are consistently related with health disparities across a broad range of contexts. Although SES is traditionally associated with trachoma risk [4,27,165], associations are found to be inconsistent between different settings and often absent. An individual level analysis by Jansen et al. (2007) in Tanzania and Vietnam relating trachoma prevalence to living standards lends support to the notion that these associations are likely to vary by context [166].

Socioeconomic indicators traditionally include measures of occupation, education and income, which are distinct but may be found to correlate in some contexts. As opposed to directly causing disease, socioeconomic factors determine access to those resources and physical environments which directly modify exposure to *C trachomatis*, or influence mediating behaviors. As such, mechanisms of association may act through the proximal risk factors discussed above (access to water, latrines and presence of cattle), or influence the risk of trachoma through other unmeasured behavioural or environmental exposures. Therefore, given the complexity of this framework, associations may vary and reflect local importance of different pathways to disease. Briefly, this section provides an overview of the mechanisms through which each of the commonly collected measures of SES may influence risk of trachoma, the evidence base and the perceived limitations.

Income-based indicators of SES are typically measured at the household level, and have been associated with a reduced prevalence of active trachoma in Ethiopia [121] and Mali [116]. While a higher income provides the means for better housing, water and sanitation, translating financial resources into disease prevention requires individuals to have invested in these improvements. In addition, income is a highly unstable measure (whereas trachoma is a chronic disease), and does not capture all assets that may influence behavior (i.e. contextual services and resources). Furthermore, an income-based measure may not translate well to protective behaviors in contexts where women have little control over how income is spent in the household. Income, however, is easily linked

to housing conditions and crowding, providing a measure which is easily collected and directly relevant to transmission pathways [116,165]. Occupation is regarded as a more stable measure than income, and provides a clear structural link between education and income. Head of household occupation has been associated with active trachoma in The Gambia [59] and Malawi [167]. However, measures of occupation do suffer from a lack of precision introduced by rather broad categories and a gender bias that may not distinguish "homemakers" from unemployed persons.

Finally, the socioeconomic measure that is arguably the most robust for trachoma is education. Although measured in different ways in different studies, education is easy to quantify and fairly stable beyond early adulthood. This measure is particularly likely to capture aspects of lifestyle and behaviors, although its importance may vary according to social values [168]. Education is typically measured as the years of education of the head of household or literacy, and in most [115,121,167], but not all [124], studies has been negatively associated with active trachoma. Child education may independently influence the risk of trachoma both through behavioural education taught through schools and reduced exposure to small (infectious) children and unsanitary household conditions.

Larger-scale environmental and climatic factors are likely to be closely linked with socioeconomic status. To start with, poor people are usually pushed to more fragile, marginal lands where water access is limited, vegetation scarce and opportunities for cattle grazing or income generation are few [169]. Thus, elevation, landcover and rainfall may act as key constraints to livelihood opportunities, except where irrigation has increased potential agricultural productivity. Fewer opportunities leads to a cycle of poverty, which may strengthen rather than break associations with environmental determinants of poverty and water availability in the absence of interventions targeted to these high-risk populations.

## 1.6 Spatial analyses and disease mapping

Standardised collection of trachoma survey data and increasing use of GPS has generated a wealth of more accurate and comparable data associated with spatial locations. These data open up a range of statistical approaches to improve our understanding of spatial patterns and processes of active disease and infection. Increasingly, spatial analyses of existing prevalence data have informed control strategies for other diseases, through an improved understanding of the scales over which a disease clusters, in order to i) inform survey design and ii) use information on underlying determinants to target surveys and interventions. This section will provide an introduction to approaches to spatial analysis and modelling, an overview of their use in trachoma research, and application to disease control.

### 1.6.1 Approaches in spatial analysis

There are three main branches of spatial statistics: continuous spatial variation, discrete spatial variation and spatial point processes. The first two both rely on sampled measurements of an outcome and model the overall degree of spatial dependency; the former assuming the outcome to be continuous in space and the latter using aggregate data. In contrast, spatial point processes model events that have arisen from a population, rather than sampling from an underlying process, and so compare the physical location of events within a study region and their propensity to cluster to a random or uniform null distribution [170]. Point processes have an important place in spatial statistics, and are the subject of many detailed reviews [171,172], but as this thesis uses data that has been purposefully sampled they are not considered further.

Four key concepts are essential to understanding spatial structure that may arise in data and statistics used to characterise this structure. First is the concept of spatial dependency, or autocorrelation, which refers to the tendency for measurements taken

from sites in close proximity to be more similar than those further apart [173]. This is a well recognized phenomenon, first defined by Tobler in his first law of geography, and results from underlying formative processes (meteorological or behavioural). Spatial dependency over a study area violates basic statistical assumptions around the independence of observations and, as each observation contributes only a fraction of the information, may result in spuriously small standard errors and false inferences regarding statistical significance. Spatial heterogeneity, defined as the variation over space of the observed values from a spatially continuous process, is a distinct but related concept. Values may be heterogeneous and unstructured in space or heterogeneous in the presence of spatial dependency, which depending on the scale over which this dependency is present will result in variable levels of overall heterogeneity.

Clearly then, the second key concept is the importance of scale in determining spatial structure and its influence on the analysis. Different processes act over different ranges and therefore, spatial patterns are likely to manifest at multiple scales. For example, close contact and mixing patterns between household members may produce small scale clustering of infection with added stochastic variability, while shared risk factors and similar exposures may induce spatial dependency over larger scales [102].

These differences in scale give rise to the third key concept, which is the distinction between first order (large scale, deterministic spatial trends) and second order (small scale and stochastic) spatial effects [103]. First order trends, for example a north-south gradient in the prevalence of infection, may arise from climatic influences and can easily be modelled and accounted for by standard regression techniques. Second order effects arise from autocorrelation and represent the tendency for neighbouring values to be more similar in their deviation from the global mean. The presence of second order effects violate assumptions of independence between observations and are the main focus of any spatial analysis [174]. The categorization of first order and second order effects will change according to the scale of the analysis – for example, variation that appears as a

trend at small spatial scales may be seen as second order variation on a larger scale [102,173]. Essentially, if the range of spatial variation is outside the range of the study area, then it appears as a large scale trend. Most spatial analyses should focus on first quantifying any trends in the global mean and then investigate autocorrelation in the residuals [103].

Finally, the concepts of stationarity and isotropy are conditions in which spatial patterns of disease are i) constant in space and ii) non-directional. Many measures and tests of spatial autocorrelation are "global", in the sense that they assume a single dominant spatial structure that exists over the whole area (i.e. the mean and covariance function do not depend on location). Where this assumption does not hold true, it implies that the mean or covariance of two variables does not only depend on distance but also on their location in space (a non-stationary process) and relative orientation (an anisotropic process). While a non-stationary field (or spatial trend) can be included through quantification of first-order trends through regression analyses, non-stationary covariance may be incorporated by local estimation of the spatial structure or various smoothing approaches [175,176].

### 1.6.2 Spatial analyses in trachoma

The use of spatially explicit methods in trachoma research has mainly focused on spatial point patterns of infection potentially reflecting transmission between households [106,177] and communities [104], or micro-epidemiological studies mapping infection or disease in a community in relation to water sources [178,179]. A small study by Bailey et al. (1989) found that amongst 15 sleeping rooms in 6 compounds, cases of active trachoma showed no evidence of spatial aggregation within households, when comparing the interpoint distance between rooms with a higher modified chi-square statistic and using Monte Carlo simulations to randomly reallocate cases and establish a null

distribution [106]. A more recent study by Broman et al. (2006) found evidence of clustering of infection but not active trachoma up to 1.3 km in Maindi in Kongwa, Tanzania, using the k-function, which is based on the magnitude of difference between observed and expected cases over a given distance [177]. In contrast to these two studies, Polack et al. (2005) used spatial scan statistics to identify three clusters of elevated risk of active trachoma within a different community in Tanzania, the largest of which had a radius of 283 meters. These studies use point-process models to quantify patterns of clustering and on such a small scale, the potential effects of stochastic variation are high.

There are limited examples of larger scale analyses that have used continuous spatial statistics to quantify associations with trachoma. Recent work by Clements et al. (2010) [158] predicted the prevalence of active trachoma in Southern Sudan using a Bayesian model that included a geostatistical random effect to account for small scale residual spatial clustering (~8 km) after accounting for the effects of rainfall and land cover. Another study which controlled for many factors at the individual, household, caretaker and village levels found no evidence of residual autocorrelation using a conditional autoregressive (CAR) spatial structure defining neighbours within 128 km [152].

### 1.6.3 Bayesian hierarchical models

In a standard regression model, a common aim is to model the mean outcome as a function of predictor variables. We recognize that not all results will fit on regression line, thus an error term is added to capture this variation. Consideration of this error term and its covariance matrix leads to formulation of more complex models, which are better suited to data that are naturally grouped at multiple levels (ie individuals into households, households in clusters, clusters in districts, etc.) or are more similar in space. Survey data are often structured this way, and it is well recognised that units of measurement that come from a group may be more similar than those from different groups, leading to correlation in their error term (or residual variance). Adjusting for this clustering is

important, as it impacts on the standard error estimated from regression. Although adjustment for clustering at the household and/or community level is recommended [180], and commonly done in analyses of trachoma risk using robust standard errors [181], General Estimating Equations [49] or mixed models incorporating random effects [127], some residual error may be spatially correlated at larger scales. For example, if a relevant risk factor is not included in the analysis of survey data, and that risk factor is spatially structured over the survey area, the assumption of independence will also be violated and the standard errors affected. Thus, the presence of spatial autocorrelation in the residuals should always be assessed and, where necessary, modelled in order to account for residual spatially-structured variation.

Bayesian hierarchical models provide a convenient framework for development of spatially structured multilevel models, as they allow incorporation of covariates and assessment of variance at different levels. Bayesian estimation offers a number of practical and computational advantages to traditional, frequentist statistics [182]. First, there are not closed form solutions for generalised linear mixed models, and frequentist estimation involves various numerical approximation methods. In contrast, Bayesian methods assume parameters are variable, and not fixed, and use Markov chain Monte Carlo (MCMC) algorithms to repeatedly generate random samples from a target distribution, conditional on the data. When these simulation chains have converged to the target distribution, they no longer depend on the starting point, and are randomly sampling within a defined distribution that can then be used for parameter estimation [183]. Second, in multilevel models where there are fewer observations at higher levels, the standard errors may not be very accurate, while Bayesian estimates obtained from MCMC procedures are more robust. Finally, Bayesian approaches have the advantage of allowing access to the posterior distribution for useful summary statistics and can incorporate prior information with the empirical data for a flexible decision framework. This is particularly useful for disease mapping, where we may be more interested in the model's ability to predict

whether the prevalence exceeds a particular threshold relevant to mapping or control [184].

### 1.6.4 Applications in disease control

Bayesian model-based geostatistics have been applied to predict the spatiotemporal risk of disease by incorporating traditional modelling techniques and spatial analysis within a Bayesian framework, generating predictions at unsurveyed locations which allow for underlying uncertainties associated with each stage to be propagated in the final estimates. This approach has been used to model spatial patterns of disease risk at national and regional scales for a number of diseases, including soil-transmitted helminths (STH), malaria and schistosomiasis [185-187]. These predictions use associations with risk factors at multiple levels, including climatic covariates which may have a role in determining vector density, water availability or behavioural patterns, as well as variation in village, household and individual-level risk factors[188-191]. Although a Bayesian geostatistical approach is increasingly used to inform planning of large-scale helminth control programmes [185,187], the potential applications for trachoma have not yet been fully evaluated and only explored in two studies [152,191].

In addition, spatially continuous estimates of risk generated through disease mapping can then be used as a platform upon which to evaluate alternative sampling designs through computer simulation, as demonstrated by Sturrock et al. (2010) [100] who used data on the prevalence of STH from East Africa to compare alternative sampling strategies in Kenya. The potential of this simulation approach in identifying alternative methods for trachoma surveys warrants further investigation.

## 1.7 Aims, objectives and thesis outline

### 1.7.1 Thesis aims and objectives

The two aims of my thesis are to i) use existing data to define and quantify the spatial distribution and disease burden of trachoma and ii) to investigate the spatial heterogeneity of trachoma and underlying risk factors at different spatial scales. Addressing these aims would have the practical consequence of providing empirical and predictive maps of trachoma to inform estimates of disease burden and evidence-based choice of survey designs. This aim will be achieved through the following specific objectives:

1. To define the spatial distribution of active trachoma and trichiasis using existing survey data and identify current mapping gaps

2. To quantify the global disease burden of trachoma for 1990 and 2010, using data on the spatial distribution of trachoma and trachomatous blindness

3. To identify individual and cluster-level factors underlying spatial patterns of trichiasis risk in Nigeria and explore the relative importance of variation at different scales

4. To identify cluster-level factors underlying spatial patterns of active trachoma in Kenya, and evaluate the potential use of geostatistical risk mapping to predict the distribution of trachoma.

5. To use computerised simulations to evaluate Integrated Threshold Mapping as an alternative trachoma survey design to cluster randomised surveys

### 1.7.2 Thesis outline

Chapter 2 describes the assembly of a global database of trachoma prevalence surveys and the current distribution of disease. These data are subsequently used in Chapters 3 through 5 to define disease burden and for more detailed epidemiological analyses, investigating potential spatial correlates and the utility of predictive risk mapping. Chapter 3 uses available data to define the burden of trachoma in Africa, in terms of trichiasis and trachomatous blindness. Chapter 4 uses a multilevel modelling approach to explore the respective roles of individual, socio-demographic, and environmental factors on trichiasis prevalence and the relative importance of variation in risk at different levels. Chapter 5 uses a similar approach to identify spatial correlates of the risk of active trachoma at the cluster level and validate a predictive model. To evaluate the use of a novel survey methodology, Chapter 6 compares the performance of Integrated Threshold Mapping to cluster randomised sampling according to standard treatment thresholds. Finally, Chapter 7 discusses the main findings and highlights the important issues that have arisen from this work.

# Chapter 2: The global distribution of trachoma

## 2.1 Overview

In order to effectively target trachoma control efforts and define the burden of disease, the geographic distribution of trachoma within countries must be known. As a first step in defining this distribution, existing surveys may be collated and mapped to describe global geographical trends in endemicity and identify gaps where data are lacking.

This section describes the assembly of a database and maps for a Global Atlas of Trachoma (GAT), which have been made freely available online at www.trachomaatlas.org since 2011. The first global trachoma atlas was developed in 2005 by the International Centre for Eye Health at the London School of Hygiene and Tropical Medicine (LSHTM) and the Programme for the Prevention of Blindness and Deafness at the WHO [70]. As part of a recent collaboration between LSHTM and the International Trachoma Initiative (ITI), this atlas was updated to reflect the availability of new survey data and changing burden of trachoma. Iterations of these data have been used at the global level by the International Task Force for Disease Eradication [192] and the International Coalition for Trachoma Control [193]. As the coordinator of the GAT during its development, I conducted the literature review, developed the initial atlas in 2011 and was responsible for the analysis presented in this chapter and published paper [194]. Data collected in the GAT directly informs the analysis of the burden of disease presented in Chapter 3.

## 2.2 Introduction

While the global distribution of trachoma has evolved over time in response to population movements, socioeconomic development and control measures, it continues to be an important cause of blindness in many countries in Africa, Asia and the Americas. Recent

estimates have suggested trachoma accounts for 3.6% of the global burden of blindness [2]. The most important limitation of our understanding of the distribution of trachoma, identified in the last review of the global distribution of trachoma in 2005, is the paucity of data in many endemic countries [70]. Since that time, international political support for the global elimination of trachoma as a public health problem has increased, with commensurate scale up of mapping activities and funding for national trachoma control programmes. Targeting available resources cost-effectively requires an understanding of the known geographic distribution of trachoma at sub-national levels and identification of gaps in survey data where further mapping is required.

The aim of this chapter is to describe the global geographical distribution and population at risk of trachoma using existing data from the GAT. Specifically, the methods of data assembly and mapping will be described and data then used to define the current geographical distribution, calculate the population at risk of TF and TT, and estimate numbers requiring treatment. Ongoing trachoma mapping efforts and remaining areas requiring further surveys will also be discussed.

## 2.3 Methods for assembly of a Global Atlas of Trachoma

### 2.3.1 Overview of the Global Atlas

The GAT adopted an identification and data assembly strategy similar to other mapping initiatives, including those for malaria [195], helminth infections [196-198] and human African trypanosomiasis [199]. In brief, epidemiological data on the burden of trachoma were identified through structured searches of published and unpublished literature, with a number of inclusion rules applied to identified information. Data were then abstracted into a standardised database and mapped using geographical information systems (GIS) software.

### 2.3.2 Identification of survey data

To assemble a global database of trachoma risk, survey data were identified through a combination of: (i) searches of electronic bibliographic databases; (ii) review of programmatic data submitted to ITI; (iii) manual searches of local archives and WHO GET2020 documents; and (iv) direct contact with programme managers and researchers. These searches, conducted in 2010 and annually thereafter, build on an earlier effort in 2005 as part of a collaboration between the International Centre for Eye Health at the LSHTM and the Programme for the Prevention of Blindness and Deafness at the WHO, to develop a first global atlas of trachoma [70]. The online bibliographic databases PubMed [200] and Embase [201] were searched to identify relevant studies, using the Medical Subject Headings "trachoma", "trichiasis", and "*Chlamydia trachomatis*". These searches were restricted to surveys conducted after 1980 for trichiasis and 1988 for active trachoma. The latter restriction was applied because 1987 is when the new simplified grading system for trachoma was introduced [26]. Authors were contacted if additional information was required on survey design or indicators collected. Countries for which no up-to-date information was available from the literature, GET 2020 country data forms, or submitted to ITI, were contacted on an individual basis for local knowledge and clarification. As a whole, these data are unpublished and use the standardised survey methodologies recommended by WHO. Work initially focused on the 55 countries classed in 2004 as trachoma endemic by WHO: 36 of which are in Africa, six in the Middle East, 10 in Asia and the Western Pacific and three in Latin America [202]. There are currently no reliable data indicating the status of trachoma in Iraq, Libya, Namibia or Zimbabwe, and these countries were also excluded from the most recent WHO update on GET2020 [203]. These countries were therefore not included in the analysis. The aim was also to collect the most contemporary data possible in order to inform current control efforts. Literature searches are conducted annually (most recently in September, 2012), and additional data submitted directly to ITI by national trachoma program managers are routinely used to

update the database and resulting prevalence maps. Data available as of September, 2012, were used in the preparation of this chapter.

### 2.3.3 Data selection and entry

The title and abstract of each source of information were reviewed and evaluated against a number of pre-defined inclusion and exclusion criteria. Only cross-sectional population based prevalence surveys were included as measures of trachoma prevalence, while TRAs were used to indicate the presence or absence of trachoma where no prevalence data were available. Data were excluded if based on hospital or clinic surveys, or surveys among sub-populations such as among refugee populations. Where multiple surveys were available from the same district but surveyed at different times, they were included as separate entries and coded as "current" or "historical" in order to ensure that only the most recent data are used to estimate the current burden of disease. Estimates of disease prevalence were typically available at the district level as this is the administrative unit at which control is implemented. Where estimates were representative of point locations or the result of a non-random selection of communities within a district, data were only used to provide information on the presence of trachoma. Abstracted data included details on the source of the data, location of survey (including geographical co-ordinates for cluster data when available), survey year, characteristics of the surveyed population, survey methodology, the numbers of children aged 1-9 years and adults aged over 14 years examined, the number of children graded positive for TF and the number of adults graded positive for TT. Any variation in clinical indicator or age group was also recorded in the database. A unique identifier linked each record in a bibliographic database to the survey data and to an electronic copy of the source when this could be obtained.

### 2.3.4 Mapping

All data were entered into a standardized Microsoft Access 2007 geodatabase (Microsoft Corporation, Redmond, WA, USA), which is linked to a geographic information system (GIS). Data can be queried to produce custom tables, thus allowing simple and rapid generation of country and regional maps using Arc GIS 10.1 (ESRI, Redlands, CA, USA). Data were assigned wherever possible to the second administrative level (e.g., district level), which has direct relevance to implementation of trachoma control. However, data from some older surveys and hyperendemic areas are available at the first administrative level (e.g., province, region), and data were assigned accordingly. The most recent data are displayed on the main maps. Where historical data are also available they are displayed on separate maps online.

Prevalence data were banded into categories corresponding to current intervention guidelines for TF and TT (Table 2.1). TRA data were categorized into three bands for active trachoma (No active trachoma found, <10% and ≥10% of children aged 1-9 years examined found positive) and two bands for trichiasis corresponding to UIG targets (prevalence of TT unknown to the health system of <0.1% or ≥0.1% of the total population). Geographical boundaries used for mapping were derived from: (i) the United Nations Second Administrative Level Boundaries data set project (http://www.unsalb.org/); (ii) Global Administrative Areas (http://www.gadm.org/); and (iii) shapefiles created specifically for this project from maps provided by programme managers. Updated district-level maps were launched in 2011 on an open-access website (www.trachomaatlas.org).

Table 2.1 Endemicity classes for implementation of SAFE based on trachomatous inflammation–follicular (TF) and trichiasis (TT)

| **TF Prevalence** | **Classification** | **Implementation** |
|---|---|---|
| <5% | Non-endemic | No need for implementation of AFE |
| ≥5% and <10% | Hypo-endemic | Mapping, F and E can be applied, focal A |
| ≥10% and <30% | Meso-endemic | AFE at district level (≥3 years then review) |
| ≥30% | Hyper-endemic | AFE at district level (≥5 years then review) |
| **TT Prevalence** | **Classification** | |
| <0.1% | UIG achieved | |
| ≥0.1% | UIG not yet achieved | |

### 2.3.5 Analysis

The characteristics of all surveys that met the inclusion criteria were summarized by country according to data source, time period, and survey methodology. Districts and regions were categorized as suspected endemic or assigned to a prevalence category using the most current PBPS data representative at this level. Districts were classified as 'suspected endemic' or 'suspected non-endemic' based on information from TRA surveys, point locations, reported cases or anecdotal information from national programs. Surveys which only collected data on one clinical sign were also used to inform this classification (i.e. a district known to be endemic for TF was classified as 'suspected endemic' for TT where no other data were available). Note that, in some cases, identified districts may not include all endemic or non-endemic areas within a country, but their classification does reflect available evidence supporting the presence of trachoma.

A total of 24 district-level surveys of TT were conducted in populations aged 40 or 50 years and over. Based on a review of age-stratified TT prevalence ratios from published and unpublished data, a conversion factor of 0.54 was applied to estimate the corresponding prevalence in adults aged ≥ 15 years.

Survey data at the region and district level were presented separately in this analysis, with district defined here as the unit of implementation typically used for SAFE control activities. While this is usually the second administrative unit within a country, in some

cases these are distinct health districts (Burkina Faso and Cameroon), third administrative areas (Ethiopia) or first administrative areas (Chad, CAR, Guinea-Bissau, Iran, Oman and the Pacific Islands). This was based on the aim to present data most relevant to current guidelines relating to the implementation of SAFE control strategies. In Australia, the most reliable prevalence data come from the National Indigenous Eye Health Survey, as opposed to routinely collected surveillance data, which provides prevalence estimates in populations classified by the 2006 Remoteness Structure [204,205].

Estimates of the current population at-risk were calculated for each country using district-level population estimates and summarised by endemicity class. In this chapter, "population at-risk" is defined as the total population living in districts that fall within a given endemicity category. Population figures for 2012 in Africa were derived from the Afripop project, which provided a continental 1 km gridded population map produced using projected population census data for 2010 and settlement extents (www.afripop.org). Population estimates for the remaining endemic countries were derived from the Gridded Population of the World, Version 3 data set (GPWv3) at the same resolution. All estimates were projected to 2012 using country-specific growth rates from the United Nations World Population Prospects [206,207]. This map was overlaid with district classification to allow summation and mapping of the population in each category of risk. Districts which had TRA data were classified as known non-endemic (if the survey found an absence of TF/TT) or suspected endemic (if clinical signs were found). Where PBPS data were available, prevalence estimates were used to classify districts according to endemicity classes (Table 2.1). Estimates of total population were used for all countries with the exception of Australia and Brazil. In these countries, indigenous population census figures were used to correspond with available prevalence estimates and present more accurate estimates of the population at risk in these countries.

## 2.5 Results

### 2.5.1 Survey database and geographical coverage

A total of 266 unique surveys with data on either active trachoma or trichiasis met GAT inclusion criteria. These included data from CRS (193), ASTRA (1), TRA (30), ITM (2), screening (3) and surveys at single sites (35) from 50 countries globally classified as endemic. Prevalence data on active trachoma were available in 40 countries and data on TT in 39. In total, there are 2131 records included in the database representing surveys conducted between 1985 and 2012, 1631 of which provide implementation unit-level estimates of prevalence (usually district-level) and an additional 80 records that provide region–level estimates. The remaining 420 records were TRAs, site-specific surveys or those of unclear methodology, which were used only to provide information on the presence or absence of trachoma at the district level. The primary source of included survey data was direct contact with national control programmes and academic researchers (70%), followed by peer-reviewed publications (15%) and unpublished reports or theses (15%). These sources of data were found to vary considerably by country, with a good deal of overlap between sources in countries with established control programmes (Table 2.3).

The majority of available prevalence data (70%) are from African countries, with a lesser proportion coming from countries in the Middle East (2%), Latin America (18%) and the Asian and Western Pacific (10%). Geographical coverage of survey data, however, is highly variable within regions and endemic countries outside of Africa tend to have a patchier, more focal distribution both of trachoma and supporting data. The number of surveys available has consistently increased over the last two decades, as highlighted in Figure 2.1 which presents the total number of PBPS surveys conducted by year for each geographic region. This figure reflects international support and progress of established national control programmes in sub-Saharan Africa in recent years, and also the historical

focus on trachoma in a number of Asian, Latin American and Middle Eastern countries such as Mexico, Myanmar, Oman and Vietnam.

Table 2.2 Countries for which no trachoma data were identified for the current atlas

| **Africa** | **Middle East** | **Asia & Western Pacific** | **Latin America** |
|---|---|---|---|
| Benin | | Papua New Guinea | Guatemala |
| Botswana | | Lao People's Democratic Republic | |
| Somalia | | | |
| Zimbabwe | | | |

Table 2.3 Total number of district-level population based prevalence surveys (PBPS), trachoma rapid assessments (TRA) and site-specific surveys in the database, summarised by source of data

| | Total Number | | Number | | Primary Source n (%) | | |
|---|---|---|---|---|---|---|---|
| Country | Surveys | Number TRA | PBPS | Number Other[a] | Direct contact[b] | Published Papers | Reports |
| **Africa** | | | | | | | |
| Algeria | 1 | 0 | 1 | 0 | 0 | 1 (100) | |
| Benin | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Botswana | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Burkina Faso | 108 | 0 | 108 | 0 | 101 (94) | 0 | 7 (6) |
| Burundi | 23 | 0 | 23 | 0 | 10 (43) | 13 (57) | 0 |
| Cameroon | 41 | 0 | 41 | 0 | 40 (98) | 1 (2) | 0 |
| CAR | 10 | 1 | 9 | 0 | 2 (20) | 1 (10) | 7 (70) |
| Chad | 8 | 0 | 8 | 0 | 0 | 5 (63) | 3 (38) |
| Cote d' Ivoire | 6 | 0 | 6 | 0 | 0 | 0 | 6 (100) |
| Djibouti | 4 | 0 | 0 | 4 | 0 | 4 (100) | 0 |
| Egypt | 5 | 1 | 2 | 2 | 0 | 5 (100) | 0 |
| Eritrea | 36 | 0 | 36 | 0 | 36 (100) | 0 | 0 |
| Ethiopia | 138 | 0 | 138 | 0 | 90 (65) | 21 (15) | 27 (20) |
| Ghana[d] | 62 | 1 | 61 | 0 | 0 | 61 (98) | 1 (2) |
| Guinea | 20 | 5 | 15 | 0 | 0 | 0 | 20 (100) |
| Guinea Bissau | 9 | 0 | 9 | 0 | 9 (100) | 0 | 0 |
| Kenya | 32 | 2 | 30 | 0 | 29 (91) | 0 | 3 (9) |
| Malawi | 5 | 0 | 5 | 0 | 1 (20) | 4 (80) | 0 |
| Mali | 62 | 1 | 61 | 0 | 29 (47) | 24 (39) | 9 (15) |
| Mauritania | 64 | 2 | 62 | 0 | 64 (100) | 0 | 0 |
| Morocco[d] | 13 | 0 | 13 | 0 | 13 (100) | 0 | 0 |
| Mozambique | 6 | 0 | 6 | 0 | 3 (50) | 0 | 3 (50) |
| Niger | 53 | 6 | 47 | 0 | 53 (100) | 0 | 0 |
| Nigeria | 282 | 87 | 195 | 0 | 157 (56) | 19 (7) | 106 (38) |
| Senegal | 11 | 0 | 10 | 1 | 11 (100) | 0 | 0 |
| Somalia | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| South Sudan | 43 | 13 | 30 | 0 | 24 (56) | 17 (40) | 2 (5) |
| Sudan | 93 | 0 | 92 | 1 | 92 (99) | 1 (1) | 0 |
| Tanzania | 66 | 0 | 66 | 0 | 58 (88) | 8 (12) | 0 |
| The Gambia | 46 | 0 | 46 | 0 | 30 (65) | 16 (35) | 0 |
| Togo | 31 | 0 | 28 | 3 | 3 (10) | 28 (90) | 0 |
| Uganda | 38 | 0 | 38 | 0 | 38 (100) | 0 | 0 |
| Zambia | 26 | 0 | 26 | 0 | 8 (31) | 0 | 18 (69) |
| Total | **1342** | **119** | **1212** | **11** | **901 (67.1)** | **229 (17.1)** | **212 (15.8)** |
| **Middle East** | | | | | | | |
| Iran[d] | 4 | 4 | 0 | 0 | 4 (100) | 0 | 0 |
| Oman[d] | 29 | 0 | 28 | 1 | 10 | 19 | 0 |
| Yemen | 14 | 10 | 4 | 0 | 0 | 10 (71) | 4 (29) |
| Total | **49** | **14** | **32** | **1** | **14 (28.6)** | **29 (59.2)** | **4 (8.2)** |
| **Asia and Western Pacific** | | | | | | | |
| Afghanistan | 8 | 5 | 0 | 3 | 3 (38) | 0 | 5 (63) |
| Australia[c] | 36 | 0 | 0 | 36 | 0 | 22 (61) | 14 (39) |
| Cambodia | 22 | 22 | 0 | 0 | 19 (86) | 0 | 3 (14) |
| China | 3 | 0 | 1 | 2 | 0 | 3 (100) | 0 |
| Fiji | 1 | 1 | 0 | 0 | 0 | 1 (100) | 0 |
| India | 21 | 14 | 1 | 6 | 0 | 5 (24) | 16 (76) |
| Kiribati | 2 | 2 | 0 | 0 | 0 | 2 (100) | 0 |
| Laos | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Myanmar[d] | 29 | 6 | 23 | 0 | 17 (59) | 1 (3) | 11 (38) |
| Nepal | 77 | 39 | 38 | 0 | 74 (96) | 0 | 3 (4) |
| Pakistan | 70 | 36 | 33 | 1 | 35 (50) | 1 (1) | 34 (49) |
| Papua New Guinea | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Solomon Islands | 8 | 3 | 5 | 0 | 5 (63) | 3 (4) | 0 |
| Vanuatu | 2 | 2 | 0 | 0 | 0 | 2 (100) | 0 |
| Viet Nam | 127 | 0 | 64 | 63 | 127 (100) | 0 | 0 |
| Total | **406** | **130** | **165** | **111** | **280 (69)** | **40 (9.9)** | **86 (21.1)** |

Table 2.3 continued

| Country | Total Number Surveys | Number TRA | Number PBPS | Number Other[a] | Primary Source n (%) Direct contact[b] | Published Papers | Reports |
|---|---|---|---|---|---|---|---|
| **Americas** | | | | | | | |
| Brazil | 334 | 0 | 300 | 34 | 300 (89) | 29 (7) | 5 (1) |
| Guatemala | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Mexico | 2 | 0 | 2 | 0 | 0 | 0 | 2 (100) |
| Total | **336** | **0** | **302** | **34** | **300 (89.3)** | **29 (8.6)** | **7 (2.1)** |
| **Total** | **2131** | **263** | **1711** | **157** | **1495 (70.2)** | **327 (15.3)** | **309 (14.5)** |

[a] Site specific surveys or those in which the sampling methodology was unclear and have been used to provide evidence of suspected endemicity where no district level PBPS or TRA were available; [b] Direct contact includes contact with National Control Programmes, NGOs and academic researchers; [c] Annual Surveillance Reports are published by the National Trachoma Surveillance and Reporting Unit in Australia, only screening data from 2012 are included here and data from the National Indigenous Eye Health Survey; [d] Reported as achieved elimination targets [203]

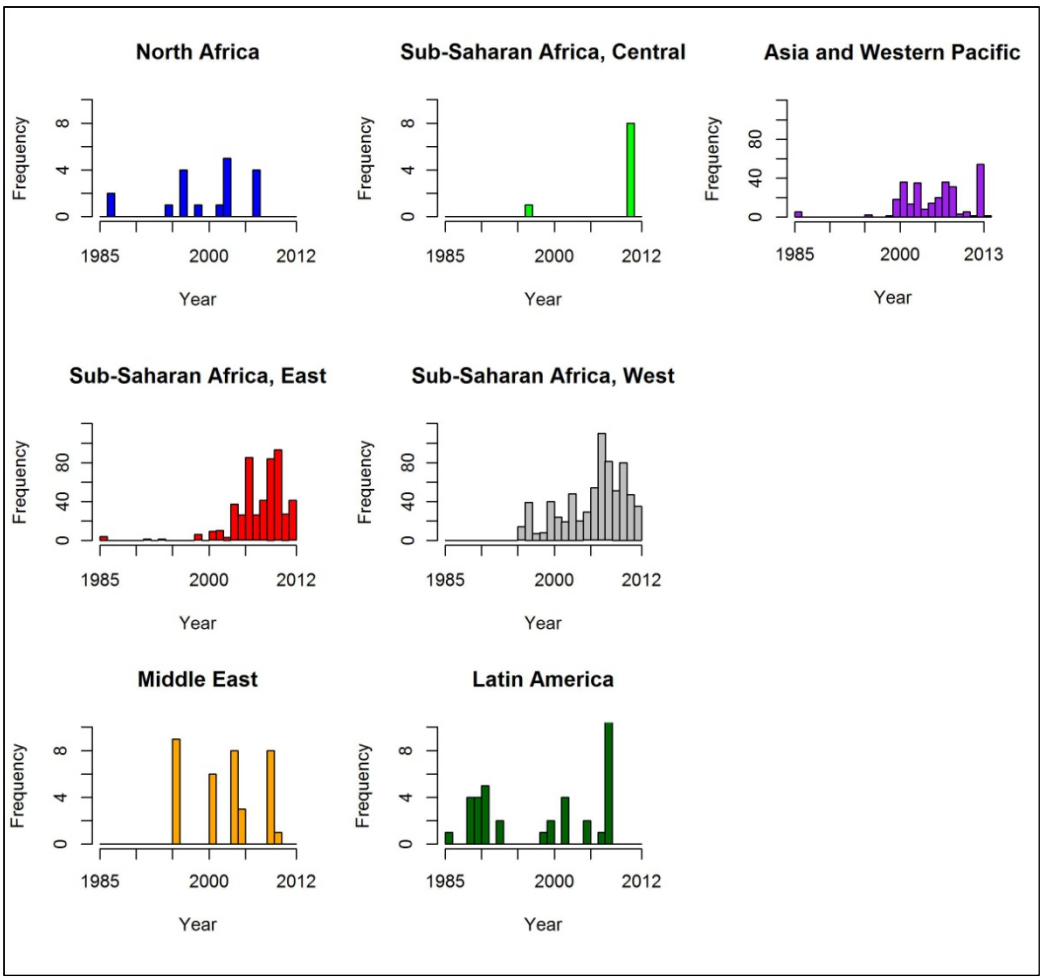



Figure 2.1 The number of population-based prevalence surveys identified by year and region globally, 1985-2013. In Africa, the graphs show a shift in survey activities from North Africa to other endemic areas, with a recent increase in the number of surveys conducted since 2005 in sub-Saharan Africa.

*Africa*

A total of 167 unique surveys with data on either active trachoma or trichiasis met GAT inclusion criteria. These included data from CRS (152), ASTRA (1), TRA (9), ITM (2) and surveys at single sites (3) from 31 of the 33 countries in Africa classified as endemic. Prevalence data on active trachoma were available in 29 countries and data on TT in 25. In total, there are 1342 records included in the database representing surveys conducted between 1985 and 2012, 1212 of which provide implementation unit-level estimates of prevalence (usually district-level) and an additional 79 records that provide region–level estimates. The remaining 10 records were site-specific surveys or those of unclear methodology, which were used to provide information on the presence or absence of trachoma at the district level.

The primary source of included survey data was direct contact with national control programmes and academic researchers (67%), followed by peer-reviewed publications (17%) and unpublished reports or theses (16%). These sources of data were found to vary considerably by country with a good deal of overlap between sources in countries with established control programmes (Table 2.3). The number of surveys available has consistently increased over the last two decades, as highlighted in Figure 2.1 which presents the total number of PBPS surveys conducted by year for each region of Africa. Surveys in north Africa and the Middle East were conducted earlier than other regions, mainly reflecting active control programmes in Morocco and some earlier surveys in Egypt. While west Africa has some historical surveys, recent survey activities are increasingly focused in this region and in east Africa.

The 33 African countries endemic for trachoma consist of 5308 districts. Of these, 1095 (20.6%) districts had representative TF data collected through PBPSs, 1024 (19.3%) with PBPS prevalence estimates for TT (Tables 2.4 & 2.7), and data from TRA surveys for an additional 101 districts. While the majority of data collected at the first administrative level are outdated and have been replaced by more recent second administrative level

surveys, Tables 2.5 and 2.8 present current data on TF and TT available at this level. Only 5 first administrative level units have trachoma prevalence data that are being used programmatically. At the time of writing, 38% of the trachoma endemic countries in Africa have more than 50% of their districts mapped by PBPS and this number is even higher when excluding districts presumed to be non-endemic from the denominator, as illustrated in Figure 2.2. These data reflect a rise in the number of large-scale national or regional surveys taking place in recent years (e.g. in Republic of Sudan and South Sudan) as well as conduct of pre-and post-implementation surveys in the context of large-scale control programmes in several countries. Since 2007, surveys have been conducted in a number of countries that previously had no data, including Burundi, Cameroon, Central African Republic, Cote d'Ivoire, Eritrea, Rwanda, Uganda and Zambia. While a number of other countries have seen a rise in survey activities during this period (e.g. Ethiopia, Guinea Bissau, Nigeria, Republic of Sudan, South Sudan, Tanzania, Togo and Zambia), prevalence estimates are still lacking in Algeria, Chad and Djibouti and no data are currently available for Benin, Botswana, or Somalia (Tables 2.3).

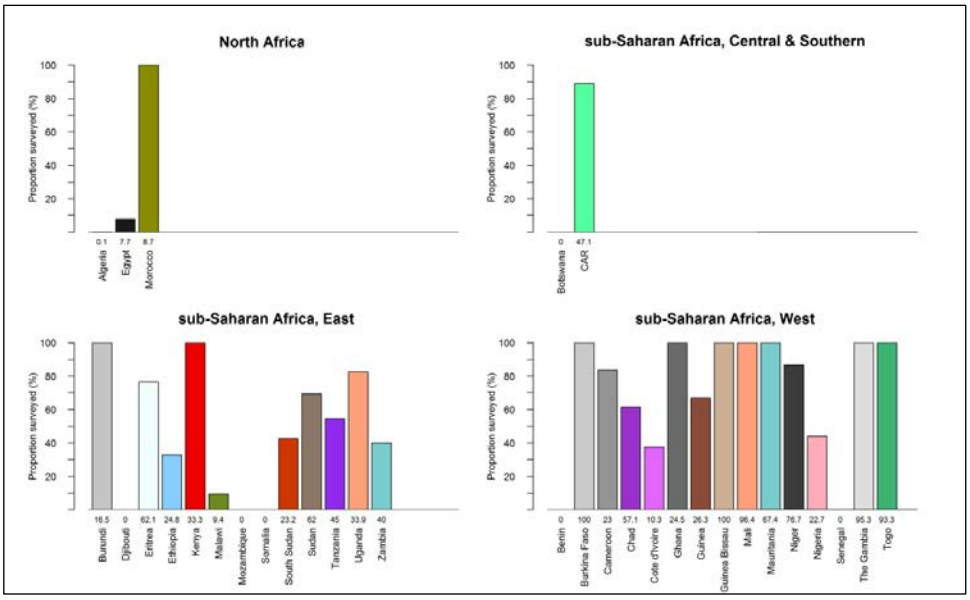



Figure 2.2 Proportion of districts surveyed by population based prevalence surveys between 1985 and 2012 in Africa. Bar plots exclude non-endemic areas from the denominator where information on suspected endemicity is available for the entire country, while numbers indicate the proportion of all districts surveyed. The graph highlights progress in mapping many endemic countries in east and west sub-Saharan Africa and the need for additional surveys in many countries in Africa.

*Middle East*

A total of 8 unique surveys with data on either active trachoma or trichiasis met GAT inclusion criteria. These included data from CRS (5), TRA (2), and screening of children (1) from the 3 countries in the Middle East classified as endemic. Prevalence data on active trachoma and trichiasis were available only in Oman and Yemen. In total, there are 49 records included in the database representing surveys conducted between 1985 and 2012, 34 of which provide estimates of prevalence at the implementation unit. An additional 15 records were site-specific surveys, TRAs or those of unclear methodology, which were used to provide information on the presence or absence of trachoma at the district level. The majority (59%) of these data are from published sources, which is much higher than other regions and mainly reflects research linked with the national trachoma control programme in Oman.

The 3 Middle Eastern countries currently classified as endemic for trachoma consist of 55 health districts. Of these, 16 (29.0%) districts had representative TF data collected through PBPSs, 12 (21.8%) with PBPS prevalence estimates for TT (Tables 2.5 & 2.8), and data from TRA surveys for an additional 12 districts. Oman has complete geographical coverage of survey and/or screening data, reported achievement of the intervention targets for the elimination of blinding trachoma in 2012 and was the first country in which verification was carried out [203]. TRA and limited prevalence surveys have indicated that trachoma may be highly endemic in Yemen while TRA data from 2004 suggest it may not be a public health problem in Iran.

*Asia and Western Pacific*

A total of 67 unique surveys from Asia or the Western Pacific regions had data on either active trachoma or trichiasis and met GAT inclusion criteria. These included data from CRS (30), TRA (19), and screening in indigenous populations or specific sites (16) from 14

countries in Asia and the Western Pacific. Prevalence data on active trachoma were available in 7 countries and data on TT in 6. In total, there are 406 records included in the database representing surveys conducted between 1985 and 2012, 164 of which provide implementation unit-level estimates of prevalence (usually district-level) and an additional record that provides region–level estimates. A further 241 records were site-specific surveys, TRAs or those of unclear methodology, which were used to provide information on the presence or absence of trachoma at the district level.

There are 15 countries in Asia or the Western Pacific classified as endemic for trachoma, which consist of 2,448 health districts. Of these, 118 (4.8%) districts had representative TF data collected through PBPSs, 95 (3.9%) with PBPS prevalence estimates for TT (Tables 2.5 & 2.8), and data from TRA surveys for an additional 131 districts. The greatest coverage of data are in Nepal and Pakistan, which have completed extensive TRAs and PBPS. The large number of districts in China, India and Vietnam partly account for the low geographical coverage observed in this region.

*Latin America*

A total of 24 unique surveys with data on either active trachoma or trichiasis met GAT inclusion criteria. These included data from CRS (6), screening (1) and surveys at single sites (17) from 2 of the 3 countries in Latin America currently classified as endemic. Prevalence data or data from regular screening for active trachoma and trichiasis were available in Brazil and Mexico, but data are lacking for Guatemala. In total, there are 336 records included in the database representing surveys conducted between 1985 and 2012, 302 of which provide implementation unit-level estimates of prevalence or screening (usually district-level). The remaining 34 records were site-specific surveys or those of unclear methodology.

Of the three countries currently classed as endemic in Latin America, a minority (5%) of the 5,616 districts have TF survey data and only three districts have TT data (Tables 2.5 &

2.8). However, only one (surveyed) state in Mexico is classified as endemic and the final assessment for achieved elimination is pending [203,208]. Brazil has a long history of trachoma and survey data are patchy both geographically and temporally, with school surveys used for assessment of TF and TT data coming from limited surveys in indigenous populations

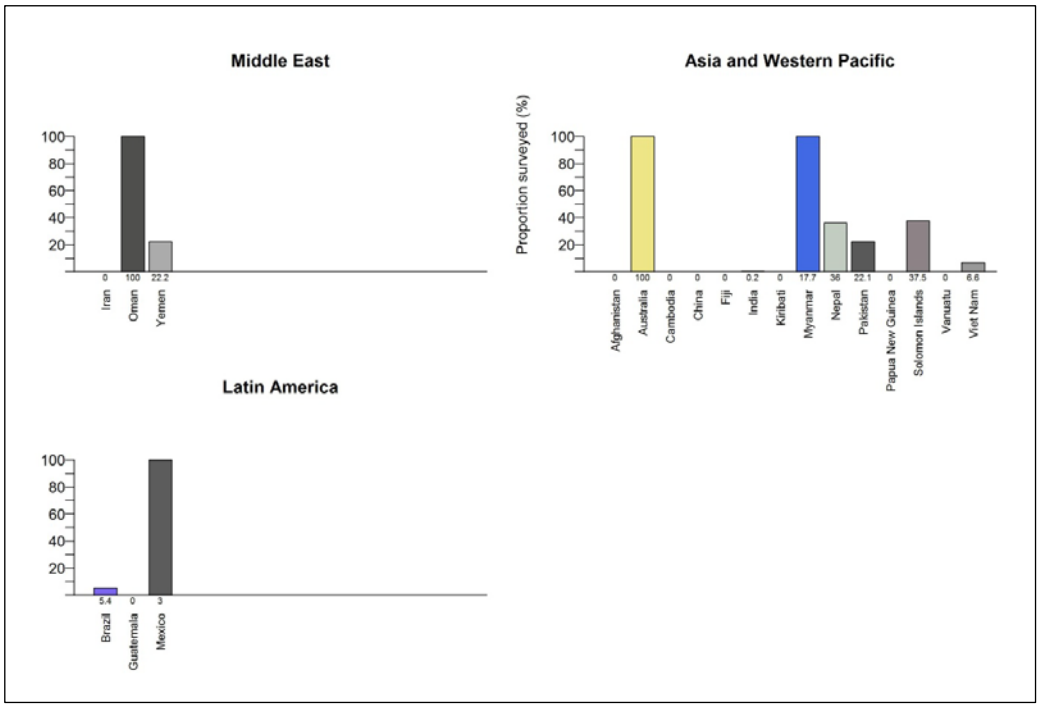



Figure 2.3 Proportion of districts surveyed by population based prevalence surveys between 1985 and 2012 in non-African regions. Bar plots exclude non-endemic areas from the denominator where information on suspected endemicity is available for the entire country, while numbers indicate the proportion of all districts surveyed. The graph reflects post-control surveys and surveillance activities in Oman, Australia, Myanmar and Mexico but also highlights the dearth of information in many of these countries.

.

### 2.5.2 Distribution of trachoma and population at risk

Globally, the distribution of trachoma varies markedly between regions, which has important implications for estimates of the population at risk of disease. Active and chronic disease remain highly endemic in much of sub-Saharan Africa, while their distribution is much patchier outside of Africa. While in some countries, such as Myanmar, Oman and Vietnam, effective control programmes have contributed to a general reduction in the prevalence of disease, there continues to be a backlog of trichiasis in areas of high historical endemicity. In other countries throughout Asia, the Western Pacific and Latin America, lower levels of endemicity are often correlated with socioeconomic development. However, as seen in Australia, Brazil, China and India, pockets of active disease may persist amongst the poorest and most marginalised communities and mirror inequities in access to water and sanitation. Empirically defining the distribution of trachoma in these countries presents a substantial challenge in generating reliable estimates of the population at risk.

In the remainder of this chapter, I will present data to show that globally, 141 million individuals live in areas confirmed empirically to be trachoma endemic (based on district-level prevalence of TF in 1-9 year-olds greater than 5%). African countries account for the vast majority (92%) of the population known to live in these areas and over 93% of the population living in meso- and hyper-endemic areas. While available data in Asia supports lower levels of endemicity, the large population and gaps in PBPS data continues to limit the reliability of estimates of the population at risk. This is reflected in the staggering figure of nearly 1.3 billion individuals living in areas suspected to be endemic. The number of individuals affected by trachoma is likely to be far lower, as it is expected to have a focal distribution throughout China and India. However, the lack of data in these countries precludes any more refined estimates of risk.

The global distribution of trichiasis (Tables 2.7 & 2.9) reflects both the known distribution of trachoma as well as areas where trachoma was historically a public health problem and a backlog of cases remain.

*Africa*

The geographical distribution of trachoma in Africa varies between regions. Trachoma is believed to be endemic in 33 of the 56 countries in Africa, which are mainly located in east and west sub-Saharan Africa, north Africa and a few endemic coastal countries in central Africa (Figure 2.2).

Based on available data, the highest prevalence of active trachoma and trichiasis remains in the Sahel area of west Africa and Savannah areas of east and central Africa (Tables 2.4, 2.6, 2.7 & 2.9). A high proportion of surveyed districts are hyper-endemic (defined as TF prevalence in 1-9 year-olds of ≥30%) in South Sudan (83%), Ethiopia (64%), Guinea (50%), Uganda (37%), Chad (38%), CAR (38%) and Tanzania (32%), but large areas suspected to be endemic remain unmapped in each of these countries. West African countries have been the focus of a number of national surveys in the last decade (Figure 2.1) providing both pre- and post-intervention data for a high proportion of districts in Burkina Faso, The Gambia, Ghana, Mali and Mauritania. Many countries in Central Africa continue to lack data, making estimation of the burden in this region difficult. Based on survey data currently included in the atlas and population estimates, an estimated 129.4 million people live in areas that are confirmed empirically to be trachoma endemic (based on district-level prevalence of TF in 1-9 year-olds greater than 5%) and a further 155 million in areas suspected to be endemic (Table 2.4). The latter is likely to be a conservative estimate, as it only includes areas classed as suspected endemic based on available TRA or anecdotal information about cases presenting to the health care system.

A substantial burden of disease is likely in Ethiopia and Nigeria due to their large populations in areas of high endemicity (Table 2.4).

As a direct consequence of repeated infections, the burden of TT follows similar geographical trends to TF within Africa. However, there is a significant backlog of TT surgeries remaining in countries with historically high endemicity levels (Tables 2.6 & 2.9). These countries include Ghana and Morocco which, despite success in reducing the burden of active disease, continue to have a high burden of TT arising from both prevalent and incident cases.

*Middle East*

Although trachoma was historically present in Saudi Arabia and Tunisia, and represented a significant public health problem in much of the Middle East and North Africa, there is evidence that it is no longer widely endemic in this region. Both Oman and Iran have already been certified as non-endemic by the WHO [203], based on PBPS in the case of Oman and information from cause of blindness surveys and primary health services in Iran [209]. The backlog of trichiasis reflects high levels of historical endemicity, although there is a low proportion of unmanaged cases in these countries [210]. There remains a lack of information on the distribution of trachoma in Yemen (Figure 2.4A), as available prevalence estimates are outdated (2001).

*Asia and Western Pacific*

Within Asia, trachoma is reported to occur in Afghanistan, Pakistan, India, Myanmar and Nepal [202,208,209,211] (Figure 4c). However, the distribution of trachoma and availability of data is patchy in this region and epidemiological data are limited, particularly for Afghanistan, India and Myanmar. In India, trachoma has been found by

TRA in some areas in the north and east, but recent PBPS data exist only in Bulandshahar and indicates a low prevalence of disease. Historically, trachoma has been a public health problem in rural central Myanmar but recently WHO has certified that the country has reached the elimination targets [203]. In contrast, Nepal has conducted extensive surveys of suspected endemic areas in the last ten years, and made encouraging progress in control. In 2005, Nepal had information for 9 (12%) of its 75 districts (all from PBPS). This has now increased to 53 districts: 27 (36%) of which are prevalence estimates from PBPS and 38 additional districts with TRA data. Pakistan is currently conducting extensive PBPS in all districts, and has completed two of the three phases of the survey to date, showing the majority of districts to be non-endemic.

Within the Western Pacific Region, the geographical distribution of active trachoma and trichiasis is well known only in Vietnam, and in Australia, where the disease is limited to Aboriginal communities (Figure 2.4C). There is some evidence that trachoma may be a public health problem in some communities in the Pacific Islands of Vanuatu, Fiji, Kiribati, and Solomon Islands, and TRAs have also found trachoma in Cambodia [212]. There is a lack of recent survey data for China although active trachoma has been reported to be present in several small, surveys in Beijing and Guangxi, and trichiasis has been reported in Hainan Province and Beijing [92,213-215]. No data could be identified for Lao People's Democratic Republic and Papua New Guinea.

*Latin America*

Only three countries in the Americas are thought to be trachoma endemic: Brazil, Mexico and Guatemala. National school surveys and PBPS among indigenous communities in Brazil suggest that a high prevalence of trachoma remains in northern areas of Brazil and focal populations (Figure 4B). Active trachoma is widely reported as eliminated from Chiapas, the last endemic region in Mexico, although it has yet to be certified as meeting its

elimination targets [203,208]. Some evidence exists to support ongoing transmission of *C trachomatis* in rural areas of Chiapas and other states in Mexico as recently as 2004 [216]. No data are currently available for Guatemala, although surveys were reported in 2012 and two districts found to be endemic [208,217]. A number of countries border highly endemic areas in Brazil, including Bolivia, Peru, Colombia and Venezuala, and imply that trachoma may be endemic in some areas. Recent surveys have found evidence of *C trachomatis* infection and later stages of disease sequelae in indigenous communities in Colombia, suggesting sustained transmission over a long period of time [218].

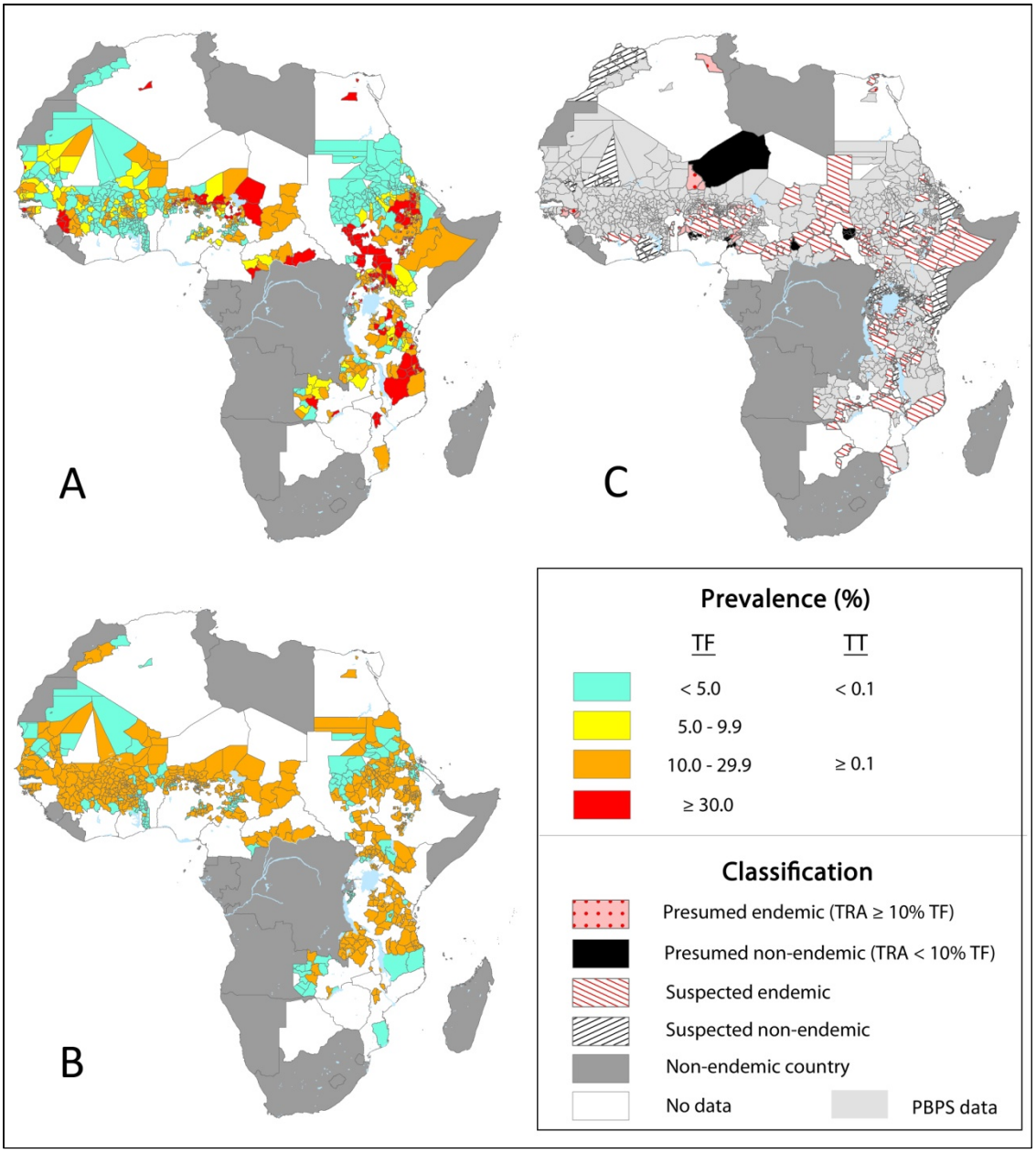



Figure 2.4 Empirical prevalence of A) trachomatous inflammation–follicular (TF) and B) trachomatous trichiasis (TT) and C) areas of suspected and presumed endemicity in Africa between 1985-2012. Population based prevalence surveys generated data for 1095 districts and 24 regions, while trachoma rapid assessment (TRA) surveys provided information on endemicity for 101 additional districts.

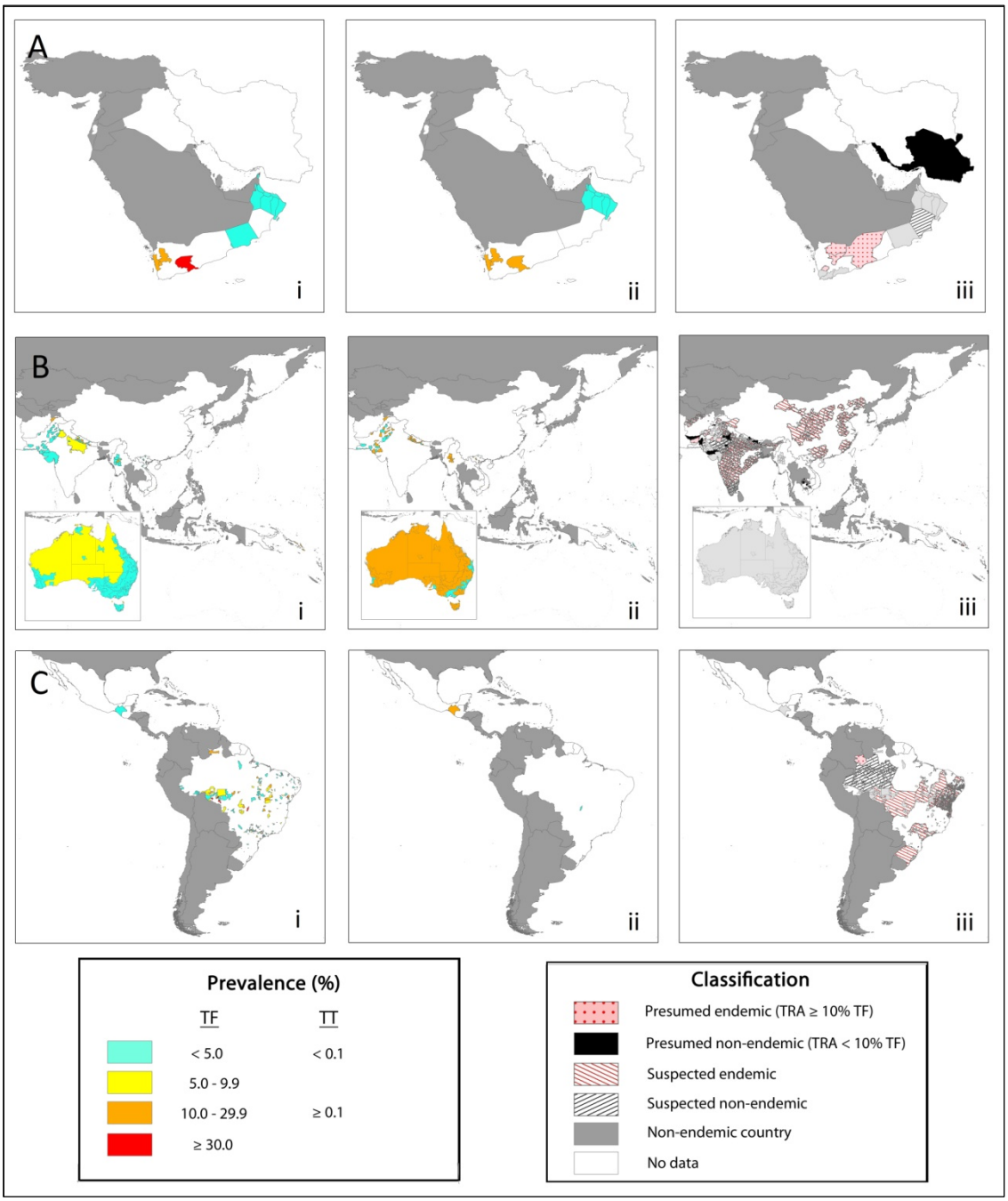



Figure 2.5 Empirical prevalence of i) trachomatous inflammation–follicular (TF), ii) trachomatous trichiasis (TT) and iii) areas of suspected and presumed endemicity in A) the Middle East, B) Asia and Western Pacific and C) Latin America between 1985-2012. Population based prevalence surveys generated data for 417 health districts and 3 regions in India, while trachoma rapid assessments (TRAs) and surveys in specific sites provided information on endemicity for 169 additional districts.

Table 2.4 Population in endemic categories of trachomatous inflammation–follicular (TF) and availability of district-level data from population-based prevalence surveys (PBPS) in Africa in children aged 1-9 years

| | | | | | | | | Prevalence of TF from PBPS[b] | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Suspected Endemic | | | <5% | | | 5-9.9% | | | 10-29.9% | | | >30% | | |
| | Total Number | Total Pop | Total Surveyed districts | | Districts | | Pop | Districts | | Pop | Districts | | Pop | Districts | | Pop | Districts | | Pop |
| Country | districts | (000s) | n | (%) | n | (%)[a] | (000s) | n | (%) | (000s) | n | (%) | (000s) | n | (%) | (000s) | n | (%) | (000s) |
| Algeria | 1,592 | 36,507 | 1 | (0.1) | | | | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 1 | (100) | 33 |
| Benin | 77 | 9,307 | 0 | (0.0) | 6 | (7.8) | 2,192 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| Botswana | 25 | 1,877 | 0 | (0.0) | 3 | (12.0) | 338 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| Burkina Faso | 63 | 16,806 | 63 | (100) | 0 | (0.0) | 0 | 24 | (38.1) | 7,497 | 16 | (25.4) | 3,805 | 23 | (36.5) | 5,504 | 0 | (0.0) | 0 |
| Burundi | 139 | 9,681 | 23 | (16.5) | 0 | (0.0) | 0 | 11 | (47.8) | 2,263 | 8 | (34.8) | 1,210 | 4 | (17.4) | 965 | 0 | (0.0) | 0 |
| Cameroon | 178 | 20,416 | 41 | (23.0) | 8 | (5.8) | 948 | 20 | (48.8) | 1,157 | 4 | (9.8) | 434 | 15 | (36.6) | 1,544 | 2 | (4.9) | 72 |
| CAR[c] | 17 | 4,540 | 8 | (47.1) | 1 | (11.1) | 194 | 0 | (0.0) | 0 | 2 | (25.0) | 1,853 | 3 | (37.5) | 749 | 3 | (37.5) | 596 |
| Chad[c] | 14 | 12,113 | 8 | (57.1) | 5 | (83.3) | 4,320 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 5 | (62.5) | 3,958 | 3 | (37.5) | 3,533 |
| Cote d' Ivoire | 58 | 19,790 | 6 | (10.3) | | | | 5 | (83.3) | 1,594 | 1 | (16.7) | 306 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| Djibouti | 11 | 791 | 0 | (0.0) | 0 | 0.0 | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| Egypt[g] | 26 | 80,095 | 2 | (7.7) | 3 | (12.5) | 11,704 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 2 | (100) | 8,226 |
| Eritrea | 58 | 5,485 | 36 | (62.1) | 11 | (50.0) | 731 | 19 | (52.8) | 2,013 | 8 | (22.2) | 968 | 8 | (22.2) | 984 | 1 | (2.8) | 92 |
| Ethiopia[d] | 928 | 86,132 | 229[f] | (24.8) | 470 | (67.3) | 32,586 | 2 | (0.9) | 257 | 4 | (1.7) | 562 | 78 | (33.9) | 11,212 | 145 | (63.5) | 18,840 |
| Ghana | 143 | 25,305 | 35 | (24.5) | 0 | (0.0) | 0 | 35 | (100.0) | 4,366 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| Guinea | 38 | 10,957 | 10 | (26.3) | 5 | (17.9) | 785 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 5 | (50.0) | 1,577 | 5 | (50.0) | 1,029 |
| Guinea Bissau[c] | 9 | 1,646 | 9 | (100) | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 1 | (11.1) | 229 | 7 | (77.8) | 1,152 | 1 | (11.1) | 213 |
| Kenya | 75 | 38,862 | 25 | (33.3) | 0 | (0.0) | 0 | 6 | (24.0) | 1,344 | 6 | (24.0) | 1,855 | 10 | (40.0) | 1,991 | 3 | (12.0) | 346 |
| Malawi | 32 | 14,460 | 3 | (9.4) | 5 | (17.2) | 2,609 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 3 | (100) | 1,290 | 0 | (0.0) | 0 |
| Mali | 55 | 15,864 | 53 | (96.4) | 0 | (0.0) | 0 | 32 | (60.4) | 9,147 | 11 | (20.8) | 3,454 | 10 | (18.9) | 1,317 | 0 | (0.0) | 0 |
| Mauritania | 46 | 4,260 | 31 | (67.4) | 0 | (0.0) | 0 | 20 | (64.5) | 1,137 | 8 | (25.8) | 299 | 2 | (6.5) | 20 | 1 | (3.2) | 764 |
| Morocco | 46[e] | 31,954 | 4 | (8.7) | 0 | (0.0) | 0 | 4 | (100) | 1,719 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| Mozambique | 132 | 22,467 | 0[f] | (0.0) | 106 | (80.3) | 16,580 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| Niger | 43 | 16,196 | 33 | (76.7) | 5 | (50.0) | 1,233 | 10 | (30.3) | 4,375 | 3 | (9.1) | 1,111 | 14 | (42.4) | 5,942 | 6 | (18.2) | 3,097 |
| Nigeria | 774 | 160,067 | 176 | (22.7) | 224 | (37.5) | 46,132 | 53 | (30.1) | 10,039 | 39 | (22.2) | 6,945 | 66 | (37.5) | 14,644 | 18 | (10.2) | 3,569 |
| Senegal | 44 | 12,034 | 0[f] | (0.0) | | | | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| Somalia | 74 | 8,958 | 0 | (0.0) | | | | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| South Sudan | 99 | 9,606 | 23 | (23.2) | 31 | (40.8) | 3,523 | 3 | (13.0) | 194 | 0 | (0.0) | 0 | 1 | (4.3) | 150 | 19 | (82.6) | 1,931 |
| Sudan | 142 | 32,376 | 88 | (62.0) | 39 | (72.2) | 7,266 | 73 | (83.0) | 18,634 | 12 | (13.6) | 3,206 | 3 | (3.4) | 381 | 0 | (0.0) | 0 |
| Tanzania | 120 | 43,494 | 54 | (45.0) | 45 | (68.2) | 17,976 | 6 | (11.1) | 2,445 | 6 | (11.1) | 2,196 | 25 | (46.3) | 7,464 | 17 | (31.5) | 4,027 |
| The Gambia | 43 | 1,719 | 41 | (95.3) | 0 | (0.0) | 0 | 21 | (51.2) | 617 | 13 | (31.7) | 569 | 7 | (17.5) | 139 | 0 | (0.0) | 0 |
| Togo | 30 | 5,944 | 28 | (93.3) | 0 | (0.0) | 0 | 28 | (100) | 5,649 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| Uganda | 112 | 32,415 | 38 | (33.9) | 8 | (10.8) | 1,899 | 1 | (2.6) | 1,753 | 3 | (7.9) | 1,289 | 20 | (52.6) | 5,450 | 14 | (36.8) | 3,848 |
| Zambia | 65 | 12,004 | 26 | (40.0) | 28 | (71.8) | 4,072 | 5 | (19.2) | 522 | 5 | (19.2) | 781 | 14 | (53.8) | 1,697 | 2 | (7.7) | 304 |
| **Total** | **5,308** | **804,128** | **1095** | **(20.6)** | **1,003** | **23.8** | **155,086** | **378** | **(34.5)** | **76,722** | **150** | **(13.7)** | **31,072** | **323** | **(29.5)** | **62,596** | **244** | **(22.3)** | **35,749** |

[a] Proportion of unsurveyed districts that are suspected endemic [b] proportion of known endemic districts falling into each category of endemicity [c] Unit of implementation (health district) is defined as the first administrative level [d] Third administrative level (wereda) is the implementation unit, but some zonal data are included in this table and used to inform SAFE implementation [e] Five districts were historically endemic in Morocco [f] Regional data available in Table 5 [g] Data in Egypt were collected at the governorate (regional) level, there have been no recent surveys at finer spatial scales and no alternative public health districts have been defined

Table 2.5 Population in each endemic category of trachomatous inflammation–follicular (TF) and availability of current district level data from population based prevalence surveys (PBPS) in the Middle East, Asia and Western Pacific and Latin America, in children aged 1-9 years.

| | | | | | Suspected Endemic | | | Prevalence of TF from PBPS[b] | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | <5% | | | 5-9.9% | | | 10-29.9% | | | >30% | | |
| Country | Total Number districts | Total Pop (000s) | Total Surveyed districts n | (%) | Districts n | (%)[a] | Pop (000s) | Districts n | (%) | Pop (000s) | Districts n | (%) | Pop (000s) | Districts n | (%) | Pop (000s) | Districts n | (%) | Pop (000s) |
| **Middle East** | | | | | | | | | | | | | | | | | | | |
| Iran[c,i] | 25 | 82,370 | 0 | (0.0) | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| Oman[c,h,i] | 12 | 3,561 | 12 | (100) | 0 | (0.0) | 0 | 12 | (100) | 3,556 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| Yemen[c] | 18 | 28,330 | 4 | (22.2) | 4 | (28.6) | 6,003 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 2 | (50.0) | 882 | 2 | (50.0) | 759 |
| **Total** | **55** | **114,261** | **16** | **(29.0)** | **4** | **(10.3)** | **6,003** | **12** | **(75.0)** | **3,556** | **0** | **(0.0)** | **0** | **2** | **(12.5)** | **882** | **2** | **(12.5)** | **759** |
| **Asia and Western Pacific** | | | | | | | | | | | | | | | | | | | |
| Afghanistan | 329 | 33,296 | 0 | (0.0) | 60 | (9.1) | 6,339 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| Australia[j,e] | 5 | 519 | 5 | (100) | 0 | (0.0) | 0 | 4 | (80.0) | 436 | 1 | (20.0) | 83 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| Cambodia | 187 | 17,067 | 0 | (0.0) | 59 | (31.6) | 5394 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| China | 344 | 1,352,277 | 0 | (0.0) | 134 | (39.0) | 51923 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| Fiji | 15 | 838 | 0 | (0.0) | 3 | (20.0) | 341 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| India | 565 | 1,185,067 | 1 | (0.2) | 414 | (25.0) | 980,185 | 1 | (100) | 3,395 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| Kiribati[k] | - | 101 | 0 | (0.0) | 1 | (100) | 101 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| Myanmar[j] | 62 | 53,102 | 11 | (17.7) | 0 | (0.0) | 0 | 9 | (81.8) | 12,469 | 1 | (9.0) | 822 | 1 | (9.0) | 2,022 | 0 | (0.0) | 0 |
| Nepal | 75 | 30,865 | 27 | (36.0) | 15 | (43.8) | 4,551 | 14 | (51.9) | 3330 | 5 | (18.5) | 1037 | 7 | (25.9) | 2738 | 1 | (3.7) | 284 |
| Pakistan | 131 | 192,888 | 29 | (22.1) | 21 | (20.5) | 44,984 | 28 | (96.6) | 52,518 | 0 | (0.0) | 0 | 1 | (3.4) | 619 | 0 | (0.0) | 0 |
| Papua New Guinea | 86 | 7,350 | 0 | (0.0) | | | | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| Solomon Islands[c] | 8 | 567 | 3 | (37.5) | 3 | (60.0) | 131 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 3 | (0.0) | 242 | 0 | (0.0) | 0 |
| Vanuatu | 6 | 229 | 0 | (0.0) | 2 | (33.3) | 115 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| Viet Nam | 635 | 88,412 | 42 | (6.6) | | | | 31 | (73.8) | 5,089 | 10 | (23.8) | 1,737 | 1 | (2.4) | 174 | 0 | (0.0) | 0 |
| **Total** | **2,448** | **2,962,578** | **118** | **(4.8)** | **712** | **(30.6)** | **1,094,064** | **87** | **(73.7)** | **77237** | **8** | **(6.8)** | **3,679** | **13** | **(11.0)** | **5795** | **1** | **(0.8)** | **284** |
| **Americas** | | | | | | | | | | | | | | | | | | | |
| Brazil[h] | 5,229 | 844 | 282 | (5.4) | 530 | (10.7) | 93 | 174 | (61.7) | 71 | 61 | (21.6) | 12 | 37 | (13.1) | 13 | 10 | (3.5) | 3 |
| Guatemala | 354 | 15,386 | 0 | (0.0) | 40 | (11.3) | 861 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| Mexico[c] | 33 | 115,068 | 1[d] | (0.0) | 0 | (0.0) | 0 | 1 | (100) | 4,730 | 0 | (0.0) | 0 | 0 | (0.0) | 0 | 0 | (0.0) | 0 |
| **Total** | **5,616** | **131,298** | **283** | **(5.0)** | **570** | **(10.7)** | **954** | **175** | **(61.8)** | **4,801** | **61** | **(21.6)** | **12** | **37** | **(13.1)** | **13** | **10** | **(3.5)** | **3** |

[a] Proportion of unsurveyed districts that are suspected endemic [b] proportion of known endemic districts falling into each category of endemicity [c] Unit of implementation (health district) is defined as the first administrative level [d] Number of districts considered to be endemic [e] National survey [g] Data in Egypt were collected at the governorate (regional) level, there have been no recent surveys at finer spatial scales and no alternative public health districts have been defined [h] Surveillance data in primary school children and prevalence surveys in focal, indigenous communities. Population data correspond to indigenous population figures calculated from the 2010 census (regional proportions) [I] WHO have certified as reaching elimination targets [j] Districts represent remoteness areas, corresponding to prevalence data from the National Indigenous Eye Health Survey [204], and indigenous population figures calculated from the 2006 census (regional proportions)

Table 2.6 Availability of current region level trachomatous inflammation–follicular (TF) data from population based prevalence surveys (PBPS) in children aged 1-9 years.

| Country | Total number regions | Total Pop (000s) | Regions with current regional-level PBPS | | Prevalence of TF from PBPS | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | <5% | | 5-9.9% | | 10-29.9% | | >30% | |
| | | | Regions *n (%)* | Pop (000s) | Regions *n (%)* | Pop (000s) | Regions *n (%)* | Pop (000s) | Regions *n (%)* | Pop (000s) | Regions *n (%)* | Pop (000s) |
| **Africa** | | | | | | | | | | | | |
| Ethiopia | 11 | 86,132 | 11 (100) | 82,835 | 5 (45.5) | 6,400 | 0 (0.0) | 0 | 5 (45.5) | 55,192 | 1 (9.0) | 21,243 |
| Mozambique[a] | 11 | 22,467 | 3 (27.3) | 4,435 | 0 (0.0) | 0 | 0 (0.0) | 0 | 2 (66.7) | 3,141 | 1[b] (33.3) | 1,294 |
| Senegal | 14 | 12,034 | 10 (71.4) | 11,107 | 3 (30.0) | 3,560 | 4 (40.0) | 3,632 | 3 (30.0) | 3,916 | 0 (0.0) | 0 |
| **Asia and Western Pacific** | | | | | | | | | | | | |
| India | 35 | 1,185,067 | 3 | 282,776 | 1 | 58,859 | 2 | 223,916 | 0 | 0 | 0 | 0 |

a Two surveys in Mozambique were conducted in "super" districts consisting of larger, aggregated geographical areas. An average value has been used for this analysis.

Table 2.7 District estimates in each endemic category of trichaisis (TT) and availability of current district level data from population based prevalence surveys (PBPS) in Africa in adults aged greater than 15 years.

| Country | Total Number districts | Total Pop (000s) | Total Surveyed districts n | (%) | Suspected Endemic Districts *n* | *(%)*[a] | Pop (000s) | Surveyed by PBPS[b] <0.1% Districts *n* | *(%)* | ≥ 0.1% Districts *n* | *(%)* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Algeria | 1,592 | 36,507 | 0 | (0.0) | 1 | (0.1) | 33 | 0 | (0.0) | 0 | (0.0) |
| Benin | 77 | 9,307 | 0 | (0.0) | 6 | (7.8) | 2,192 | 0 | (0.0) | 0 | (0.0) |
| Botswana | 25 | 1,877 | 0 | (0.0) | 3 | (12.0) | 338 | 0 | (0.0) | 0 | (0.0) |
| Burkina Faso | 63 | 16,806 | 63 | (100) | 0 | (0.0) | 0 | 6 | (9.5) | 57 | (90.5) |
| Burundi | 139 | 9,681 | 0 | (0.0) | 4 | (2.9) | 965 | 0 | (0.0) | 0 | (0.0) |
| Cameroon | 178 | 20,416 | 41 | (23.0) | 8 | (5.8) | 948 | 15 | (36.6) | 26 | (63.4) |
| CAR[c] | 17 | 4,540 | 9 | (52.9) | 1 | (12.5) | 194 | 1 | (11.1) | 8 | (88.9) |
| Chad[c] | 14 | 12,113 | 8 | (57.1) | 5 | (83.3) | 4,320 | 0 | (0.0) | 8 | (100) |
| Cote d' Ivoire | 58 | 19,790 | 6 | (10.3) | | | | 4 | (66.7) | 2 | (33.3) |
| Djibouti[g] | 11 | 791 | 0 | (0.0) | 4 | (36.4) | 580 | 0 | (0.0) | 0 | (0.0) |
| Egypt[h] | 26 | 80,095 | 2 | (7.7) | 3 | (12.5) | 11,704 | 0 | (0.0) | 2 | (100) |
| Eritrea | 58 | 5,485 | 36 | (62.1) | 11 | (50.0) | 731 | 14 | (38.9) | 22 | (61.1) |
| Ethiopia[d] | 928 | 86,132 | 202 | (21.8) | 470 | (64.7) | 32,586 | 1 | (0.5) | 201 | (99.5) |
| Ghana | 143 | 25,305 | 35 | (24.5) | 0 | (0.0) | 0 | 15 | (42.9) | 20 | (57.1) |
| Guinea | 38 | 10,957 | 15 | (39.5) | 0 | (0.0) | 0 | 0 | (0.0) | 15 | (100) |
| Guinea Bissau[c] | 9 | 1,646 | 9 | (100) | 0 | (0.0) | 0 | 0 | (0.0) | 9 | (100) |
| Kenya | 75 | 38,862 | 13 | (17.3) | 7 | (11.3) | 258 | 0 | (0.0) | 13 | (100) |
| Malawi | 32 | 14,460 | 3 | (9.4) | 5 | (17.2) | 2,609 | 0 | (0.0) | 3 | (100) |
| Mali | 55 | 15,864 | 53 | (96.4) | 0 | (0.0) | 0 | 2 | (3.8) | 51 | (96.2) |
| Mauritania | 46 | 4,260 | 31 | (67.4) | 0 | (0.0) | 0 | 12 | (38.7) | 19 | (61.3) |
| Morocco[e] | 46 | 31,954 | 5 | (10.9) | 0 | (0.0) | 0 | 0 | (0.0) | 5 | (100) |
| Mozambique | 132 | 22,467 | 0 | (0.0) | 106 | (85.3) | 16,580 | 0 | (0.0) | 0 | (0.0) |
| Niger | 43 | 16,196 | 33 | (76.7) | 2 | (20.0) | 137 | 6 | (18.2) | 27 | (81.8) |
| Nigeria | 774 | 160,067 | 175 | (22.6) | 230 | (38.4) | 47,072 | 26 | (14.9) | 149 | (85.1) |
| Senegal | 44 | 12,034 | 0 | (0.0) | 1 | (2.3) | 283 | 0 | (0.0) | 0 | (0.0) |
| Somalia | 74 | 8,958 | 0 | (0.0) | | | | 0 | (0.0) | 0 | (0.0) |
| South Sudan | 99 | 9,606 | 17 | (17.2) | 40 | (48.8) | 4,160 | 0 | (0.0) | 17 | (100) |
| Sudan | 142 | 32,376 | 87 | (61.3) | 39 | (70.9) | 7,266 | 23 | (26.4) | 64 | (73.6) |
| Tanzania | 120 | 43,494 | 55 | (45.8) | 45 | (69.2) | 17,976 | 3 | (5.5) | 52 | (94.5) |
| The Gambia | 43 | 1,719 | 39 | (90.7) | 0 | (0.0) | 0 | 38 | (97.4) | 1 | (0.0) |
| Togo | 30 | 5,944 | 28 | (93.3) | 0 | (0.0) | 0 | 25 | (89.3) | 3 | (10.7) |
| Uganda | 112 | 32,415 | 35 | (31.3) | 9 | (11.7) | 2,113 | 0 | (0.0) | 35 | (100) |
| Zambia | 65 | 12,004 | 24 | (36.9) | 27 | (65.9) | 3,871 | 4 | (16.7) | 20 | (83.3) |
| **Total** | **5,308** | **804,128** | **1,024** | **(19.3)** | **1,027** | **(24.0)** | **156,915** | **195** | **(20.1)** | **829** | **(88.6)** |

[a] Proportion of unsurveyed districts that are suspected endemic [b] Proportion of known endemic districts falling into each category of endemicity [c] Unit of implementation (health district) is defined as the first administrative level [d] Third administrative level (wereda) is the implementation unit, but some zonal data are included in this table and used to inform SAFE implementation [e] National survey [f] Regional data available in Table 6 [g] TT estimates are in the whole population (0-99 years)[h] Data in Egypt were collected at the governorate (regional) level, there have been no recent surveys at finer spatial scales and no alternative public health districts have been defined.

Table 2.8 District estimates in each endemic category of trichaisis (TT) and availability of current district level data from population based prevalence surveys (PBPS) in Middle East, Asia and Western Pacific and Latin America in adults aged greater than 15 years.

| Country | Total Number districts | Total Pop (000s) | Total Surveyed districts n | (%) | Suspected Endemic Districts *n* | (%) | Pop (000s) | Surveyed by PBPS[b] <0.1% Districts *n* | (%) | ≥ 0.1% Districts *n* | (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Middle East** | | | | | | | | | | | |
| Iran [c,d] | 25 | 82,370 | 0 | (0.0) | 0 | (0.0) | 0 | 0 | (0.0) | 0 | (0.0) |
| Oman [c,d,e] | 12 | 3,561 | 8[f] | (100) | 0 | (0.0) | 0 | 8 | (100) | 0 | (0.0) |
| Yemen [c,] | 18 | 28,330 | 4 | (22.2) | 4 | (28.6) | 6,003 | 0 | (0.0) | 4 | (100) |
| **Total** | **55** | **114,261** | **12** | **(21.8)** | **4** | **(9.3)** | **6,003** | **8** | **(66.7)** | **4** | **(33.3)** |
| **Asia and Western Pacific** | | | | | | | | | | | |
| Afghanistan | 329 | 33,296 | 0 | (0.0) | 57 | (17.3) | 5880 | 0 | (0.0) | 0 | (0.0) |
| Australia[h] | 5 | 519 | 5 | (100) | 0 | (0.0) | 0 | 2 | | 3 | |
| Cambodia | 187 | 17,067 | 0 | (0.0) | 43 | (23.0) | 3754 | 0 | (0.0) | 0 | (0.0) |
| China | 344 | 1,352,277 | 0 | (0.0) | 134 | (39.0) | 51923 | 0 | (0.0) | 0 | (0.0) |
| Fiji | 15 | 838 | 0 | (0.0) | | | | 0 | (0.0) | 0 | (0.0) |
| India | 565 | 1,185,067 | 0 | (0.0) | 343 | (60.7) | 760,437 | 0 | (0.0) | 0 | (0.0) |
| Kiribati | - | 101 | 0 | (0.0) | 1 | (100) | 101 | 0 | (0.0) | 0 | (0.0) |
| Myanmar | 62 | 53,102 | 6 | (9.7) | 0 | (0.0) | 0 | 0 | (0.0) | 6 | (100) |
| Nepal | 75 | 30,865 | 26 | (34.7) | 0 | (0.0) | 0 | 3 | (11.5) | 23 | (88.5) |
| Pakistan | 131 | 192,888 | 29 | (22.1) | | | | 27 | (93.1) | 2 | (6.9) |
| Papua New | 86 | 7,350 | 0 | (0.0) | | | | 0 | (0.0) | 0 | (0.0) |
| Solomon Islands[c] | 8 | 567 | 1 | (12.5) | 4 | (57.1) | 251 | 1 | (100) | 0 | (0.0) |
| Vanuatu | 6 | 229 | 0 | (0.0) | 2 | (33.3) | 115 | 0 | (0.0) | 0 | (0.0) |
| Viet Nam | 635 | 88,412 | 28 | (4.4) | | | | 3 | (10.7) | 25 | (89.3) |
| **Total** | **2,448** | **2,962,578** | **95** | **(3.9)** | **584** | **(24.8)** | **822,461** | **36** | **(37.9)** | **59** | **(62.1)** |
| **Americas** | | | | | | | | | | | |
| Brazil[g,i] | 5229 | 844 | 2 | (0.0) | 530 | (10.1) | 93 | 1 | (0.0) | 1 | (0.0) |
| Guatemala | 354 | 15,386 | 0 | (0.0) | 40 | (0.0) | 861 | 0 | (0.0) | 0 | (0.0) |
| Mexico[c] | 33 | 115,068 | 1[f] | (100) | 0 | (0.0) | 0 | 1 | (100) | 0 | (0.0) |
| **Total** | **5616** | **131,298** | **3** | **(0.0)** | **570** | **(10.2)** | **954** | **2** | **(66.7)** | **1** | **(33.3)** |

[a] Proportion of unsurveyed districts that are suspected endemic [b] Proportion of known endemic districts falling into each category of endemicity [c] Unit of implementation (health district) is defined as the first administrative level [d] WHO have certified as reaching elimination targets [e] TT cases refer to unmanaged/new cases [f] Number of districts considered to be endemic [g] Surveys restricted to indigenous populations, of varying representativeness [h] Districts represent remoteness areas, corresponding to prevalence data from the National Indigenous Eye Health Survey [204], and indigenous population figures calculated from the 2006 census [i] TT estimates are in the whole population (0-99 years)

Table 2.9 Availability of current region level trachomatous trichiasis (TT) data from population based prevalence surveys (PBPS) in adults aged greater than 15 years.

| Country | Total number regions | Total Pop (000s) | Regions with current regional-level PBPS Regions *n* *(%)* | Pop (000s) | Prevalence of TT from PBPS <0.1% Regions *n* | *(%)* | ≥0.1% Regions *n* | *(%)* |
|---|---|---|---|---|---|---|---|---|
| **Africa** | | | | | | | | |
| Senegal | 14 | 12,034 | 9 (64.3) | 11,108 | 0 | (0.0) | 9 | (100) |

### 2.5.3 Future mapping needs

Available data suggest that most highly endemic countries are in Africa, and mapping efforts should target gaps in these countries in order to scale up SAFE interventions in time to reach elimination targets. In many countries, particularly Central African Republic, Ethiopia, South Sudan, Tanzania and Zambia, there remain a large proportion of unmapped districts that are suspected to be endemic based on higher-level (i.e. first administrative level) prevalence surveys, health systems data or rapid assessments. Based on median TF prevalence in 1-9 year-olds of >20% in surveyed districts (where more time may be needed for control activities to reduce disease prevalence to below elimination thresholds), countries which should be prioritized to finish mapping include Chad, Egypt (based on limited and outdated data at the regional level), Ethiopia, Guinea, Mozambique, Nigeria, South Sudan, Tanzania, Uganda and Zambia. Several other countries with ongoing control programmes, including Cameroon, Kenya, and Malawi, have few remaining unmapped districts that are suspected to be endemic for trachoma and mapping could be completed within a shorter time frame (Table 2.4).

Globally, approximately a fourth (25.7%) of surveyed districts fall in the 5-10% TF prevalence category which indicates that they may require higher resolution mapping at the subdistrict level to target MDA to disease foci. As trachoma has a patchier distribution in many Asian and Latin American countries, higher resolution mapping of areas suspected to be endemic in these countries may be warranted. In addition, those countries in Asia, the Western Pacific and Latin America with a more focal distribution of trachoma may present an opportunity to eliminate trachoma in a shorter time frame.

## 2.6 Discussion

With prevalence estimates available for at least parts of 40 trachoma endemic countries globally and for 12.6% of all districts in these countries, the GAT represents the most comprehensive resource on the geographical distribution of trachoma and an important planning tool for efforts to finalise the global trachoma map and target interventions to priority districts. Globally, trachoma continues to be a significant public health problem in many parts of the world, with important regional differences in data availability and endemicity patterns. Variation in risk of trachoma both within and between countries and sub-national administrative areas has been linked to socioeconomic factors that are associated with transmission through hygienic behaviours and sanitation, as well as varying climatic conditions [152,157,158,219]. Based on the current data and population estimates, an estimated 141 million people live in trachoma endemic areas (TF prevalence greater than 5% in children) and a further 1.3 billion in areas suspected to be endemic. Current data from the GAT confirm that countries with the highest burden of active trachoma and trichiasis remain in the Sahel and Savannah areas of Africa, and the majority of individuals at risk (129.4 million known and 155 million suspected) are in Africa. These figures correspond to 98 million people who live in areas of Africa where the prevalence of active trachoma is known to be greater than 10% and currently require access to the SAFE strategy including annual MDA with azithromycin, and a further 31 million people where treatment may need to be targeted at the subdistrict level (Table 2.4).

The reduction of trachoma in a number of countries in the past decade has been attributed both to successful control programmes, as well as rapid socioeconomic development. Well established control programmes in several west and north African countries and outside of Africa are likely to have had an impact on the burden of trachoma in the last decade, with successes in control activities documented in Burkina Faso, The Gambia [220], Ghana [221,222], Mali [223-225], Mauritania [202], Mexico [208], Morocco [226], Myanmar [209], and Oman [227]. The Gambia, Ghana, Iran, Morocco, Myanmar and Oman have now

reported achievement of trachoma elimination targets and trachoma is generally believed to be no longer a public health concern in these countries [203]. While the global map of trachoma endemicity may be shrinking, information continues to be scarce in many countries which are assumed to have a lower burden of disease (such as in China, India, Latin America, and the Pacific Islands) and entirely lacking for others (Botswana, Djibouti and Somalia). While endemicity is highest in Africa and more widespread, trachoma continues to be reported in marginalised populations and small foci in many Asian and Latin American countries experiencing rapid development, such as China, Brazil and India. The absence of reliable PBPS data continues to limit efforts to describe the geographical distribution of trachoma and accurately estimate population at risk in these countries. Areas suspected to be endemic in India and China alone adds as many as 1 billion individuals to the population suspected to be at risk, although the numbers affected are likely to be far lower due to the focal distribution of trachoma reported in these countries. In addition, estimates of population at risk presented in this chapter do not include populations of countries currently classified as endemic, but for which no data are currently available (i.e. Botswana, Djibouti, Guatemala, Iran and Somalia).

It should be recognised that data included in GAT vary in quality and methodology, which limit the comparability of the data. The methods used to collect data (sample size, age groups and sampling method) vary and data are collected over a range of years, in which potential socioeconomic changes could introduce further variation. While differences in the age groups surveyed for TT have been adjusted for, older data may not represent current levels of endemicity where mass antibiotic treatments, TT surgery campaigns and secular trends have had an impact on the prevalence of trachoma. Information on treatment and maps of antibiotic and surgical interventions are available on a partner website developed by the International Coalition for Trachoma Control (http://www.trachomacoalition.org/). In practice, these detailed data are assessed contextually and used alongside treatment data to make mapping decisions within a country. Prevalence estimates are rarely reported with confidence intervals, limiting our

ability to assess their precision. Generally, precision for TT prevalence estimates is likely to be low as surveys are usually powered only to provide estimates for active trachoma. While much of the available survey data have helped to inform trachoma control activities, some survey data have not been used to inform control due to limited resources, outdated prevalence data or use of unreliable sampling methodologies. Where prevalence data are felt to be unusable because of their age or the methods used for their collection, the corresponding areas will need to be resurveyed. The wide prevalence bands used to display these data minimizes the effect of this imprecision and of variation in survey methodologies.

Variation in the geographical scale at which surveys are conducted introduces a further level of complexity. While the unit of implementation is defined by WHO as the district (which generally corresponds to the second administrative level), in some cases the region (first administrative level) is used instead. Recent recommendations allow data from larger geographic areas (e.g. regions) to justify programme launch in areas where local knowledge or higher level data demonstrate that trachoma is widespread and highly endemic, as was the case in Unity state of South Sudan and Amhara region in Ethiopia [228,229]. Much historical data in west Africa are representative at regional level and thus not directly comparable to district level data. Future work could include methods, such as small area estimation, to estimate uncertainty and provide realistic confidence intervals for population estimates [230].

A closely related issue is the focal distribution of trachoma, which varies with individual and community-level risk factors [127,152,179,231]. Displaying data aggregated at higher administrative levels belies the small scale spatial heterogeneity of clinical disease and the true population at risk, implying that higher resolution mapping is required to accurately capture patterns of risk. However, geographical analyses are currently lacking to demonstrate: i) how the spatial heterogeneity of trachoma changes over the course of a control programme, ii) the relative importance of hotspots for elimination of blinding

trachoma and iii) the optimal spatial resolution of survey data to inform decision-making. It is likely that some areas of countries currently regarded as non-endemic may have small pockets of transmission occurring, such as areas in DRC bordering CAR, South Sudan and Zambia, and areas bordering Brazil (Bolivia, Colombia, Peru and Venezuela). In addition, the inclusion of urban areas into estimates of the population at risk is contentious, as they are commonly perceived as having lower risk of trachoma and are generally excluded from the sampling frame of population-based prevalence surveys. However, these urban populations are typically defined locally and thus vary between countries and districts. Urban populations were included in estimates of the population at risk, due to i) a lack of reliable evidence that there is no risk of trachoma in urban areas and ii) the absence of a comparable definition of urban with which to identify these populations. However, in contexts where non-surveyed urban populations have a different risk of trachoma, this decision will result in an under- or over-estimation of the population at risk.

Information from this analysis highlights a number of important next steps for defining the burden of trachoma to inform programmatic action. First, a number of countries have both a high prevalence of active trachoma in mapped areas and a large proportion of unmapped districts that are suspected to be endemic. These countries include Central African Republic, Ethiopia, Nigeria, South Sudan and Tanzania. Second, Chad, Guinea, Mozambique, and Uganda are likely to have sizeable areas of high endemicity contributing to the current magnitude of the burden of trachoma in Africa. Generation of baseline data where required, and commencement of interventions in these countries should be accelerated. Third, prioritising countries that have large populations in highly endemic areas, such as Ethiopia, and Nigeria, will have a greater impact on the overall burden of disease within the programmatic timeframe. Egypt also may be prioritised, based on this rationale, due to the high endemicity of trachoma found in populous areas by earlier regional surveys and a lack of data excluding other geographical areas. Targeting future survey activities to areas which are likely to be highly endemic will allow the initiation of control activities in those areas in which control of trachoma is likely to take the longest.

In addition, 13.7% of surveyed districts lie in the 5-10% prevalence category and may require higher resolution mapping at the subdistrict level to target MDA to disease foci (Table 3). Countries in Latin America may represent a easy elimination target where it is found in pockets of endemicity and indigenous populations, however the full extent of distribution of disease must be better defined [232]. In these and other countries reported to have a focal distribution of disease, such as China and India, strategies must be decided to reliably define the distribution of disease. This presents significant challenges due to the geographical and population size of these countries. Increasingly, data collected in these and other developing countries come from surveillance in school children and disease reporting systems. While countries with established and equitable health care and education systems are able to provide more sensitive reports of trachoma endemicity, other countries may struggle to provide conclusive evidence on the absence of disease.

Finally, scaling up surgical interventions for TT alongside MDA poses an important challenge in reducing the burden of disease and is increasingly perceived as a limiting factor in meeting UIG targets. There is a substantial backlog of surgeries in countries with historically high endemicity rates. While the incidence of TT will decrease over time along with the number of cases of active disease, the reduction of TT cases should be the first main goal of control programmes and necessitates scaling up of surgical services in order to meet UIG targets. This presents a number of logistical challenges and demands on human resources; requiring considerable investment in health infrastructure and training in order to identify TT cases and optimise surgical outcomes in order to achieve a sustainable impact.

In conclusion, the data presented in this chapter highlight the heterogeneous global distribution of trachoma both within countries and between countries and regions. This variation will have profound impact on projected estimates of the burden of trachoma, which typically are based on very limited data on the prevalence of trachomatous blindness and heavily extrapolated within geographic regions. Therefore, the following

chapter explores the use of TT data presented in this chapter to inform more robust estimates of the actual burden of disease. Subsequent chapters will investigate the correlates of TT and TF at country scales and the potential of risk mapping.

# Chapter 3: The disease burden of trachoma in Africa

## 3.1 Overview

The previous chapter described the current global distribution of trachoma using data from the Global Atlas of Trachoma (GAT). As well as demonstrating a high prevalence of active and chronic stages of disease in Africa, the GAT highlights substantial increases in the availability of epidemiological data over the last decade. While detailed estimates of the prevalence of trachoma and population living in areas known to be endemic are essential for targeting control activities, summary estimates at national scales are important for large-scale planning. Comparable estimates of the burden of disease which incorporate the disability associated with disease processes may be used to prioritise countries, provide a comparable measure to other diseases, track changes in burden over time and to justify requests for and allocation of resources against competing priorities.

This chapter presents work on a new methodology using data from the GAT to estimate the burden of trachoma in Africa, where current initiatives are being focused and information is needed most. The GAT represents a rapidly growing source of information on the distribution and prevalence of trichiasis, which is a direct mechanical cause of visual impairment. I conducted the literature review and was responsible for the analysis presented in this chapter. Estimates of trachomatous blindness and low vision generated using this methodology were used by the Vision Loss Expert Group to inform the Global Burden of Disease 2010 project [71].

## 3.2 Introduction

While the burden of a disease is often measured by the numbers of individuals in a particular health state, such indicators only provide information on the magnitude of the

health problem without quantifying the additional burden due to associated functional limitations. Summary measures that combine information on resultant morbidity, mortality and economic losses have the benefit of capturing different aspects of the disability associated with a disease process in a single indicator, thus providing a framework to compare the burden of different disease states between countries. Disability-adjusted life years (DALYs) are one such summary measure that combines information on morbidity and mortality in order to quantify the years of full health lost to a disease at a given time point. Health metrics such as these are increasingly required by donors and implementing organisations for improved decision making and to demonstrate quantifiable results and programme performance [233,234].

Estimation of the burden of disease due to trachoma is defined by the numbers of cases of blindness and low vision attributed to trachoma and the resulting disability associated with this vision loss [79-82,235]. In later stages of clinical trachoma, trichiasis develops and the lashes turn inwards to touch the cornea. This abrasion provides an important mechanical cause of progressive vision loss, although other factors such as secondary bacterial and fungal infections and ocular dryness are speculated to contribute to this process [18,19]. While blindness and low vision are disabling sequelae with well recognized physical, social and economic repercussions, any functional limitations and discomfort due to trichiasis are not well defined. Trichiasis, occurring with or without concurrent visual impairment, has not been previously been included as a cause of disability in burden of disease estimates [79-82,235]. However, a study by Frick et al. (2001) in rural Tanzania provided evidence that trichiasis was associated with excess functional limitations in daily activities of afflicted individuals, regardless of vision status [83]. In addition, ocular pain and photophobia associated with this condition may cause further mental distress and social consequences not captured by the disability assigned to vision loss. The higher prevalence of trichiasis and younger age distribution of this condition suggests that it could have an important contribution to the overall disease burden of trachoma.

A seminal effort to define the global prevalence of trachomatous blindness was provided by Ranson and Evans (1996) for the 1990 GBD study [79,235], which has informed subsequent estimation strategies [80-82]. Unfortunately, epidemiological data on the prevalence of trachomatous blindness are geographically and temporally limited and national surveys on causes of blindness are even scarcer. This lack of data has been overcome by either i) extrapolating the results of existing national causal blindness surveys to all endemic countries within a world region [80,235] or ii) modeling national data using information on Gross Domestic Product (GDP) to provide estimates for countries that are lacking blindness data and those without data at multiple time points [82].

In order to capture all available information on the burden of trachoma, this chapter presents a comprehensive framework to model the prevalence of trachomatous blindness, using population-based estimates of the prevalence of trichiasis within countries where national blindness surveys were not available. In addition, the burden contributed by trichaisis was evaluated and a sensitivity analysis conducted to estimate the impact of different assumptions used in the modelling process.

## 3.3 Methods for estimating the burden of trachoma in Africa

### 3.3.1 Overview

Data on the prevalence of trichiasis in Africa were extracted from the GAT, as summarised in the previous chapter. Age-standardised and matched data were used to model the relationship between trichiasis and trachomatous blindess, and trachomatous blindness and low vision. Trichiasis prevalence data were then used to predict the burden of trachomatous blindness where national prevalence estimates were lacking as outlined in Figure 3.1. Models were developed within a Bayesian framework that incorporated age-prevalence curves for each health state and published estimates of the difference in risk

between males and females. Age and sex-specific estimates of the burden of trachomatous trichiasis, trachomatous low vision and trachomatous blindness were estimated in each endemic country and corresponding DALYs for each condition. A sensitivity analysis was conducted in order to vary assumptions around the method of extrapolating data where within-country geographical coverage was low, and to estimate the impact of the disability weight assigned to disease sequelae.

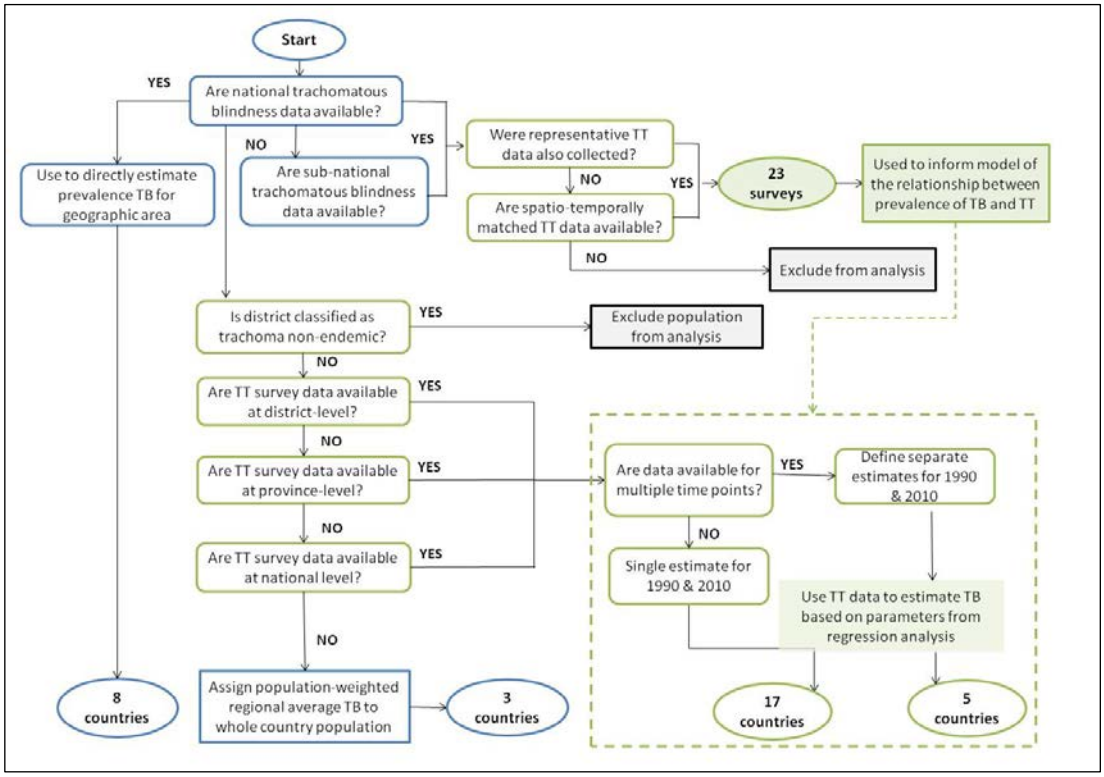



Figure 3.1 Decision algorithm used to determine the prevalence of trachomatous blindness for each district in suspected endemic countries. NB: Trichiasis data from Chad, Morocco and The Gambia provided separate estimates for 2010 in addition to those 5 countries shown here.
TB: trachomatous blindness; TT: trichiasis

### 3.3.2 Disease definition and health states

Based on the natural history of disease (reviewed in Chapter 1.2.2), the disabling sequelae of trachoma include low vision and blindness. In causal blindness surveys these two states are attributed to trachoma based on an ophthalmic examination of individuals with

impaired vision (typically defined as a Snellen visual acuity below 6/18). The operational definition of trachoma as a cause of blindness are typically based on the WHO/PBL coding instructions for eye examination, which diagnose the condition based on the presence of central scarring with trichiasis/entropian, conjunctival scarring, pannus or Herbert's pits [77]. In practice, an element of subjectivity is present in identifying the principal cause of vision loss, particularly where multiple causes of blindness are present or diagnostic criteria vary between studies. While the WHO guidelines are most commonly used, studies may extend diagnostic criteria for trachomatous low vision to include a history of trichiasis [236].

Table 3.1 lists the case definitions set out by the WHO's International Statistical Classification of Diseases and corresponding disability weights for low vision and blindness proposed in the recent GBD Disability Weights Study [75,237]. In this analysis, we explore the additional burden caused by trichiasis in Africa, with or without concurrent loss of vision. The disability weight for trichiasis was set as 0.034, which is half of the weight proposed in a recent study by Gouda et al. (2012), in reference to earlier weights for vision loss used in WHO's 2004 Global Burden of Disease Study [238].

Table 3.1 GBD cause category, disabling sequelae, case definition and disability weights for trachoma (adapted from Salomon et al. (2013) [237] & Gouda et al. (2012) [238].

| **GBD cause/sequelae** | **Case definition** | **Disability weight** |
|---|---|---|
| ***Trachoma (W031)*** | | |
| Blindness | Best corrected visual acuity in the better eye of less than 3/60 due to corneal opacity as a result of trachoma. | 0.195 |
| Low vision | Best corrected visual acuity in the better eye of less than 6/18 but better than or equal to 3/60 due to corneal opacity as a result of trachoma. | 0.112 |
| Trichiasis | At least one eyelash rubs on the eyeball or evidence of recent removal of in-turned eyelashes | 0.034 |

### 3.3.4 Data Assembly

*Literature review*

A full systematic literature review was carried out in August 2011 and updated in July 2013 to identify published and unpublished blindness studies providing estimates of the prevalence of trachomatous blindness. Searches were focused on the list of 33 African countries classified as known or suspected endemic for trachoma, as outlined in the previous chapter. Sources included i) electronic bibliographic databases (including PubMED and EMBASE), ii) references from earlier global burden reviews [79-82,235] and iii) the global vision database [239]. Electronic search terms used in combination with the name of the country included either of the two keywords "trachoma" or "cause" and one of the following keywords: "low vision", "blindness" and "visual impairment". References were screened and retained where i) surveys were conducted after 1980, ii) papers used the definitions of blindness and low vision presented in Table 3.1 and iii) prevalence estimates were derived from population-based samples. Surveys that met the initial inclusion criteria included population based causal blindness studies and rapid assessment of avoidable blindness (RAAB) surveys [74].

Only studies which estimated the prevalence of trachomatous blindness using a national sampling frame were included in the prevalence calculations (Appendix 3.1). Sub-national surveys reporting both trachomatous blindness and trachomatous low vision, or trachomatous blindness and trichiasis, were also retained for modeling these associations (Appendix 3.2). Data on the level of the survey (national, sub-national, local), location, year, age range of surveyed individuals, numbers of individuals examined, numbers positive for each outcome and mean prevalence were extracted from each included study. For studies where only older populations were sampled, estimates were adjusted to include all age groups using the age-weighting method described in Appendix 3.3 and country- and age-specific demographics.

*Prevalence data*

Endemic countries were classified according to the 5 epidemiological regions used by the GBD 2010 within Africa [71]. A summary of all available data on blindness and trichiasis from population-based prevalence surveys (PBPS) that were used to calculate estimates are detailed by country in Appendix 3.4. Empirical estimates of low vision and blindness attributed to trachoma are available from a limited number of national causal blindness surveys (Appendix 3.1). Where these data were available at a country level, they were adjusted if necessary and applied to country-level population figures to directly inform estimates of burden. A single national blindness survey was available for 8 of the 33 countries classed as endemic (Benin, Botswana, Chad, Eritrea, Ethiopia, Morocco, Nigeria, The Gambia).

Where national blindness estimates were not identified, provincial or district estimates of trichiasis from the GAT informed the analysis and were applied to the district-level population, as outlined in the algorithm provided in Figure 3.1. Available data in Africa used for these estimates (also reviewed in the previous chapter) and country-specific decisions are documented in Appendix 3.4. Data on the prevalence of trichiasis were used to estimate trachomatous blindness in 22 countries lacking national blindness data. Trichiasis data were also used to model the prevalence of trachomatous blindness in 2010 in Chad, The Gambia, and Morocco, where more recent trichiasis survey data were available.

*Extrapolation of data in time and space*

Burden estimates were defined for two time periods: 1990 and 2010. As multiple causal blindness and trachoma surveys are typically only available in the context of national trachoma control programmes, data were limited in both time and space for many countries (Appendix 3.4). Separate estimates for the two time periods could be calculated

for eight countries: three of which had an older national blindness survey and more recent (or post-intervention) trichiasis data (Chad, The Gambia, Morocco) and five which had comprehensive pre and post intervention data by PBPS at the district level (Burkina Faso, Ghana, Mali, Mauritania, Niger). Two further countries (Kenya and Tanzania) had limited post-intervention data for some districts and so had slightly different data for the two time periods. Prevalence estimates for the remaining countries are based on all available data.

Where population-based prevalence estimates were lacking, the population of a district was excluded if trachoma rapid assessments did not find trachoma in the area or it was suspected to be non-endemic based on hospital data. Districts where trachoma had been found by rapid assessment or were lacking information were assigned a population-weighted average prevalence value following the protocol for the base case outlined in Table 3.2. Where geographic coverage of population-based prevalence data was high (>50%) within a country, the median prevalence within countries was extrapolated to areas lacking data. In countries with less than 50% geographical coverage of data, regional prevalence averages were extrapolated to areas with no data. Figure 3.2 summarizes the proportion of districts for each country for which prevalence data were available, were classed as non-endemic or were based on extrapolated data.

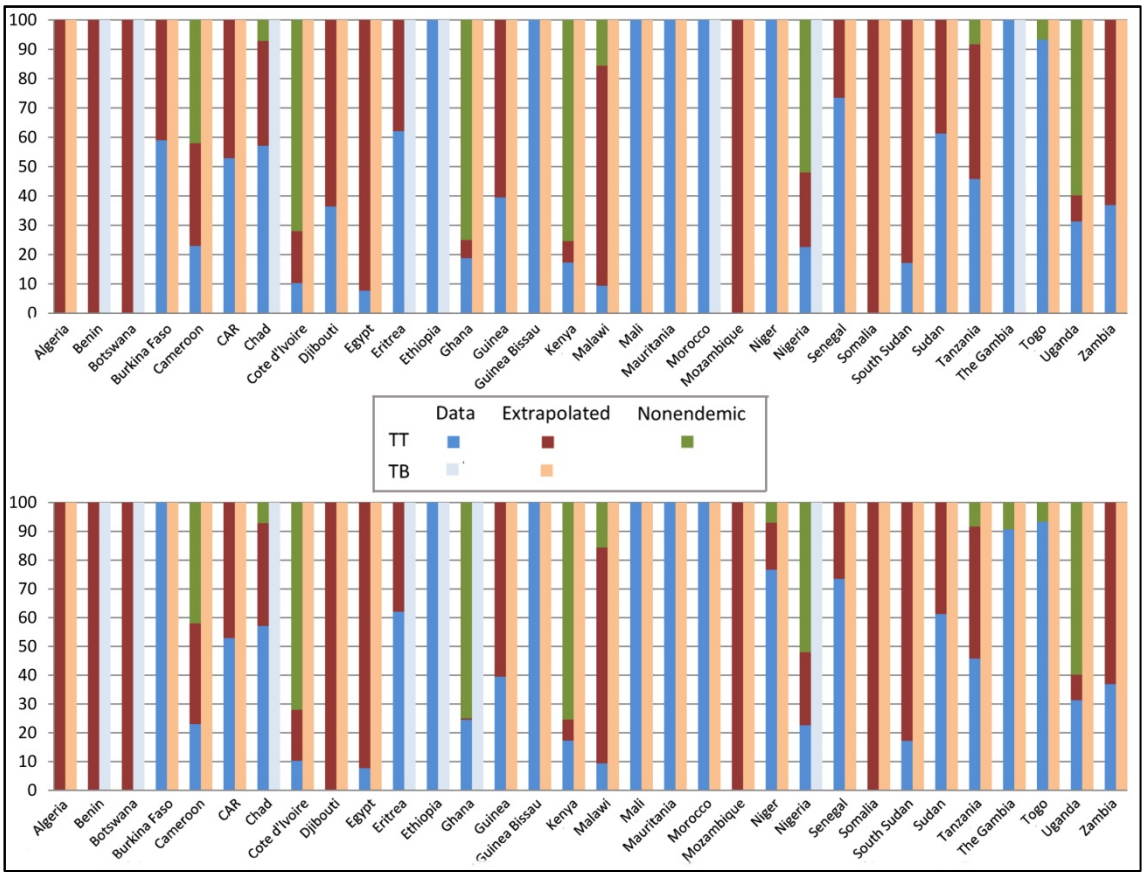



Figure 3.2 Bar plot of proportion of districts where trachomatous trichiasis (TT) data are available in 1990 (top) and 2010 (bottom), classed as non-endemic or extrapolated using the base case scenario described in Table 3.2. Data used are fully described in Appendix 3.4 and "extrapolated" trachomatous blindness (TB) data are modelled from TT data. NB: Burundi was excluded from the graph as it was assigned a prevalence of zero based on evidence that it is non-endemic.

In this chapter, estimates are based solely on available empirical data and were not modeled further at the country level to incorporate temporal trends. Although most of these countries have not implemented large-scale control, some North African countries in particular may have experienced a secular decline. A temporal trend was not included based on the rationale that 1) secular changes captured by changes in country-level indicators (such as Gross Domestic Product (GDP) or Joint-Monitoring Programme indicators) are likely to be inequitable within African countries and may overestimate changes in endemic areas and 2) trachomatous blindness (even more than trichiasis) results from repeated episodes of active disease experienced over 20-30 years and these sequelae tend to "backlog". In addition, surgical intervention for trichiasis has made slow gains compared to MDA of antibiotics and post-operative recurrence of eyelash

malposition is common, with some evidence that there will continue to be incident cases even after ocular chlamydial infection has been eliminated [240].

*Demographic data*

District-level population figures were derived from the Afripop project, which provided a continental 1 km gridded population map produced using projected population census data for 2010 and settlement extents (www.afripop.org). These estimates were adjusted by a correction factor to be consistent with the official population estimates for 1990 and 2010 provided by the Institute of Health Metrics (IHME) for the GBD 2010. Detailed demographic ratios on the proportion of the population within each age category were also derived from this source.

### 3.3.4 Modelling

*Age- and sex- prevalence of TT and TB*

Clear age-prevalence relationships are observed both for trachomatous blindness and trichiasis. Even in hyper-endemic areas, trachomatous blindness is rarely observed in individuals under the age of 15 years and the burden of disability increases with age [241]. Ranson and Evans (1996) estimated approximately 80% of the burden in adults over 60 years of age, 18% in individuals aged 45-59 years and 2% in those aged 15-44 years [79]. Subsequent efforts to quantify the burden of trachoma have assumed the same distribution. For this analysis, disaggregated data from published and unpublished studies were used to model the age-distribution of the burden of trichiasis (12 studies), trachomatous blindness (4 studies) and low vision (2 studies). These models were used to standardize survey data using methods detailed in Appendix 3.3, and calculate age-specific prevalence estimates within the Bayesian framework. Although the age-prevalence profile is expected to vary with trachoma endemicity, there were insufficient data to characterize

age-prevalence profiles of trachomatous blindness in different endemic contexts. Estimates were set to zero for low vision in ages 0-4 years and for trachomatous blindness in ages 0-4 and 5-14 years. This was done to ensure that no positive cases were predicted, based on the near-absence of these conditions in these age groups even in highly endemic areas [241].

Previous estimates have assumed that women account for 75% of trachomatous trichiasis and used a male-to-female ratio of 1:3 [82]. More recently, a comprehensive meta-analysis by Cromwell et al. (2009) [113] reviewed this assumption, and reported an overall odds ratio of trichiasis in women compared with men of 1.82 (95% CI 1.61-2.07). This ratio was used to inform sex-specific prevalence estimates for trachomatous blindness in the following analysis, by incorporating an informative prior. Males and females were assumed to have the same age-distribution of each sequelae.

*Prevalence of TT and trachomatous blindness*

The relationship between trichiasis and trachomatous blindness was quantified using 23 blindness surveys that were spatio-temporally matched at the corresponding administrative level (usually district) to i) trichiasis data collected during the same survey or ii) data available in the GAT (Appendix 3.2).

The mean prevalence of trachomatous blindness was modeled using a generalized linear Poisson regression model, using a log link, within a Bayesian framework. By including an offset term equal to the population, Poisson models are appropriate for modelling rates and particularly suitable where the underlying mean is low (allowing a skewed distribution) and events to take positive, integer values. A number of covariates were evaluated as possible explanatory factors for observed variation, including: GBD region, GDP, mortality rate, the presence of an established trachoma control programme and the year of the survey. Unfortunately, the majority of these variables are only available at the national level and thus may not be well matched to the study. The Poisson distribution

assumes the variance of response to be equal to its mean, whereas data are frequently found to be overdispersed (more variable) than this allows. The presence of overdispersion was formally tested using a goodness of fit chi-squared statistic based on the residual deviance divided by the degrees of freedom. The final model included trichaisis prevalence as a log-transformed continuous fixed-effect and a random effect ($RE$) added to allow for observed extra-Poisson variation between studies. The mean prevalence of blindness ($tb\_\gamma_i$) was predicted as follows from counts ($Y_i$) arising from examined individuals ($offset_i$) in each of *i* districts, where $\alpha$ is the intercept and $\beta_1$ is the coefficient for trichiasis:

$$(i) \qquad Y_i \sim Pois(tb\_\gamma_i)$$

$$(ii) \qquad \log(tb\_\gamma_i) = \alpha + \beta_1 \text{ x } \log(tt\_\gamma_i) + \log(offset_i) + RE_i$$

Predictive models were based on the mean district prevalence ($tt\_\gamma_i$) in each of the *i* districts, which incorporated uncertainty associated with sampling error around the mean estimate by modeling it within a binomial distribution based on the number of examined individuals ($Sample$) and the number of TT positive individuals ($tt\_pos$).

$$(iii) \qquad tt\_pos_i \sim Binomial(tt\_\gamma_i, Sample_i)$$

Where the sample size was unknown, district estimates from PBPS were assigned a higher sample size (the average sample size within a given country) and a progressively lower weight was assigned to areas where data were extrapolated, reflecting our growing uncertainty in areas with no PBPS data. Areas known to be trachoma endemic by TRA or classified non-endemic based on hospital data were assigned a sample size of 500 and areas with no information a sample size of 100.

*Trachomatous blindness and low vision*

Low vision was modeled similarly, using the observed relationship between trachomatous blindness and low vision. This was quantified using 29 studies which had measured the prevalence of both trachomatous blindness and associated low vision, also detailed in Appendix 3.2. Estimates were age-standardized and modeled with a generalized linear fixed-effects Poisson regression model, using a log link, within a Bayesian framework.

$$(i) \qquad Y_i \sim Pois(lv\_\gamma_i)$$

$$(ii) \qquad \log(lv\_\gamma_i) = \alpha + \beta_1 \text{ x } \log(TB_i) + \beta_2 \text{ x } \log(TB_i)\text{^}2$$

Where ($Y_i$) is the count of individuals with low vision in each of *i* districts, $\alpha$ is the intercept, $\beta_1$ is the coefficient for the logged prevalence of trachomatous blindness ($\log(TB_i)$), and $\beta_2$ is the coefficient for the squared term. Unlike the previous model, the Poisson distribution provided an adequate fit to variation in these data.

*Country-level estimates*

Median estimates of the prevalence of trachomatous blindness and low vision and their corresponding confidence intervals were summarized using either the national blindness survey data or the predicted prevalence of trachomatous blindness as detailed in Appendix 3.4. Prevalence estimates for each age group were calculated using (i) age weightings derived from age-prevalence models described above and (ii) country level demographic profiles (Appendix 3.3). In additional to overall prevalence estimates (as presented by country in Tables 3.5-3.6), estimates were calculated using a standard based on the overall demographic structure in Africa in 1990 to facilitate comparisons between time points.

All Bayesian analyses described above were conducted in WinBUGS version 1.4 (MRC Biostatistics Unit, Cambridge and Imperial College, London, UK) using the "RtoWinBUGs" package in R version 2,10,1.

### 3.3.5 Estimation of DALYs

DALYs are calculated from adding years of life lost to premature mortality (YLLs) to a weighted estimate of years of life lived with disability (YLDs) [71]. As trachoma does not have any YLLs attributed to it, it is simply calculated using YLDs. The most recent GBD 2010 study revised their previous methodology to compute YLDs as the number of prevalent cases of each sequela multiplied by the relevant disability weight, without any age weighting or discounting. The same prevalence based approach was used in this analysis, and DALYs were calculated using the equation:

$$PYLD = DW\ x\ p$$

Where PYLD is the prevalence YLD for a specific disease state, DW is the disability weight and p is the number of prevalent cases per capita.

### 3.3.5 Sensitivity analysis

The impact of extrapolating data within countries and regions was varied in a sensitivity analysis, summarised in Table 3.2. Scenario 1 includes only known PBPS data in all countries, so that areas lacking data were assumed to have a trichiasis prevalence (and blindness prevalence) of zero. All countries with no data and areas within countries with less than 50% geographical coverage (including areas classified as non-endemic) were assigned a regional average of trichiasis prevalence data in Scenario 2 or a country average in Scenario 3.

Table 3.2 Sensitivity analyses varying method of data extrapolation to areas lacking data in i) countries with no data, ii) countries with limited geographical coverage including non-endemic areas (less than 50%) and iii) those with geographical coverage over 50%

| | | Sensitivity Analyses | | |
|---|---|---|---|---|
| **Country** | **Base case[a]** | **Scenario 1** | **Scenario 2[a]** | **Scenario 3[a]** |
| **Countries with no data** | | | | |
| Algeria | Average: NA/ME | Assumed 0 | Average: NA/ME | Average: NA/ME |
| Djibouti (2010) | Average: NA/ME | Assumed 0 | Average: NA/ME | Average: NA/ME |
| Mozambique | Average: SSAS | Assumed 0 | Average: SSAS | Average: SSAS |
| Somalia | Average: SSAE | Assumed 0 | Average: SSAE | Average: SSAE |
| **Countries with < 50% geographical coverage** | | | | |
| Djibouti | Average: NA/ME | Assumed 0 | Average: NA/ME | Country Average |
| Egypt | Average: NA/ME | Assumed 0 | Average: NA/ME | Country Average |
| Guinea | Average: SSAW | Assumed 0 | Average: SSAW | Country Average |
| Malawi | Average: SSAE | Assumed 0 | Average: SSAE | Country Average |
| South Sudan | Average: SSAE | Assumed 0 | Average: SSAE | Country Average |
| Zambia | Average: SSAE | Assumed 0 | Average: SSAE | Country Average |
| **Countries with ≥ 50% geographical coverage** | | | | |
| Sub-national areas w/no data | Country Average | Assumed 0 | Corresponding Regional Average | Country Average |

[a] Regional averages: North Africa/Middle East (NA/ME); Sub-Saharan Africa West (SSAW); Sub-Saharan Africa East (SSAE); Sub-Saharan Africa Central (SSAC); Sub-Saharan Africa Southern (SSAS),

Disability weights were also varied when calculating DALYs (Table 3.3). The initial value used corresponded to that assigned in the disability weights measurement study for the GBD 2010 [237]. These are approximately a third of weights assigned in previous iterations of the GBD, and thus were used as a lower bound to evaluate the impact of changing weights. Upper estimates for trachomatous blindness and low vision weights correspond to those used 2000 GBD study [82]. Trichiasis was assigned an upper disability weight corresponding to that recently suggested by Gouda et al. (2012) in proportion to the GBD 2000 weights [238].

Table 3.3 Disability weights associated with visual impairment attributed to trachoma and trichiasis, derived from the Global Burden of Disease (GBD) study 2010 and upper estimates corresponding to the GBD 2000 study [82,237,238].

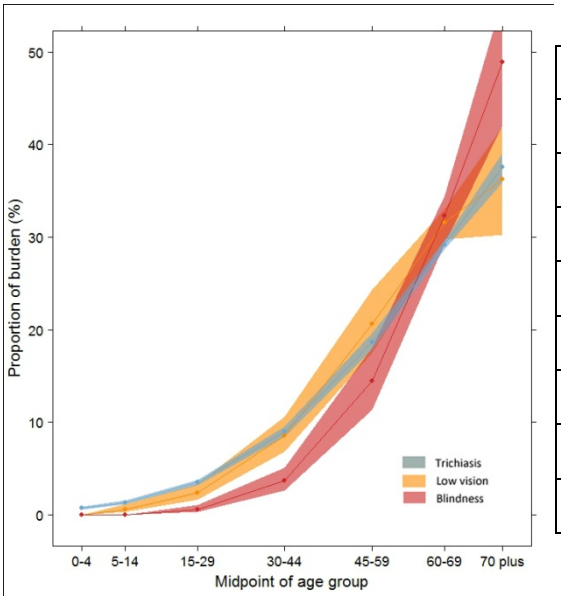
| | Disability Weights | | |
|---|---|---|---|
| Sequelae | GBD 2010 | Current analysis | Sensitivity Analysis |
| Blindness | 0.195 | 0.195 | 0.6 |
| Low Vision[a] | 0.112 | 0.112 | 0.245 |
| Trichiasis | - | 0.034 | 0.068 |

[a] Mean of weights for moderate and severe low vision

## 3.4 Results

*Age-prevalence curves*

Results from models of the age distribution of the burden of trichiasis, trachomatous blindness and low vision were consistent with our understanding of disease progression (Figure 3.3). While the burden of all disease sequelae increased with age, the risk of trachomatous blindness was relatively low in individuals under age 30 years and increased steeply after the age of 40 years, compared to trichiasis and low vision. This resulted in a higher proportion of the burden of blindness in older age groups compared to earlier stages of disease (Figure 3.3; Table 3.4). The relatively narrow confidence intervals for trichiasis reflect the greater availability of age-stratified prevalence data for this condition compared to trachomatous blindness and, particularly, low vision.




| Age Group (years) | Proportion of burden (%) | | |
|---|---|---|---|
| | TT | LV | TB |
| 0-4 | 0.72 (0.59, 0.89) | 0 | 0 |
| 5-14 | 1.36 (1.18, 1.58) | 0.52 (0.24, 0.95) | 0 |
| 15-29 | 3.50 (3.22, 3.81) | 2.24 (1.50, 3.16) | 0.61 (0.35, 1.0) |
| 30-44 | 8.95 (8.42, 9.50) | 8.63 (6.73, 10.74) | 3.64 (2.60, 4.90) |
| 45-59 | 18.61 (17.78, 19.46) | 21.23 (17.87, 25.31) | 14.36 (11.38, 17.49) |
| 60-69 | 29.14 (28.68, 29.58) | 31.74 (30.23, 33.60) | 32.18 (29.20, 34.19) |
| 70 plus | 37.7 (36.18, 39.31) | 35.26 (28.70, 41.45) | 49.17 (43.13, 55.67) |
| Total | 100 | 100 | 100 |

Figure 3.3 and Table 3.4 Calculated estimates of the proportion of the burden of trichiasis (TT), trachomatous low vision (LV) and trachomatous blindness (TB) in each age group, with corresponding confidence intervals around each point, connected by lines.

*Models*

The prevalence of trachomatous blindness increased non-linearly with the (log-transformed) prevalence of trichiasis (Figure 3.4). There was some indication that areas with an established control programme had a lower risk of trachomatous blindness associated with a given level of trichiasis. However, this association dropped out after allowing for extra-Poisson variability observed in the model, by addition of a normally distributed random effect.

Similarly, the prevalence of low vision increased non-linearly with the (log-transformed) prevalence of trachomatous blindness so that there was a low vision prevalence of 1.4 per 1000 when the prevalence of blindness was 1 per 1000, but was much higher in areas where blindness was highly prevalent (>2 per 1000) (Figure 3.4). There was substantial variation in this relationship however, which was not found to be associated with region, presence of an established control programme or year the survey was conducted.

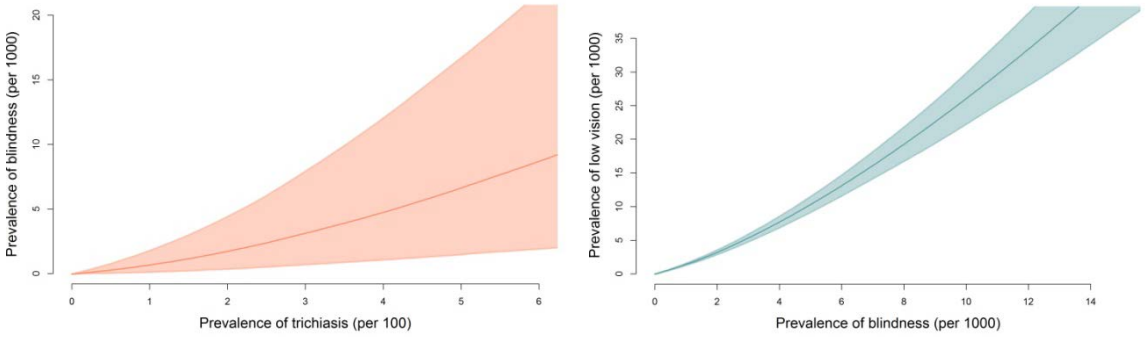



Figure 3.4 Poisson regression models of the prevalence of trachomatous blindness (left) from trichiasis data and low vision (right) from blindness data. Lines represent the mean estimate, while shaded areas depict associated uncertainty in the model. Input data are detailed in Appendix 3.2.

*Burden of disease*

The age-standardised prevalence of trichiasis, trachomatous low vision and trachomatous blindness in Africa was estimated as 0.72 (per 100), 1.44 (per 1000) and 0.68 (per 1000) in 1990. In 2010, these estimates had decreased to 0.65 (per 100), 1.12 (per 1000) and 0.55 (per 1000). The changes in prevalence may in part reflect the impact of established control programmes, particularly those West African countries with pre- and post-intervention trichiasis data or blindness surveys available. These countries include Burkina Faso, Chad, Ghana, Mali, Mauritania, Niger, Morocco and The Gambia, which in turn influence regional estimates extrapolated to countries with low geographical coverage of data. Figure 3.5 highlights changes in the prevalence of the three age-standardised sequelae between different regions and heterogeneity between endemic countries.

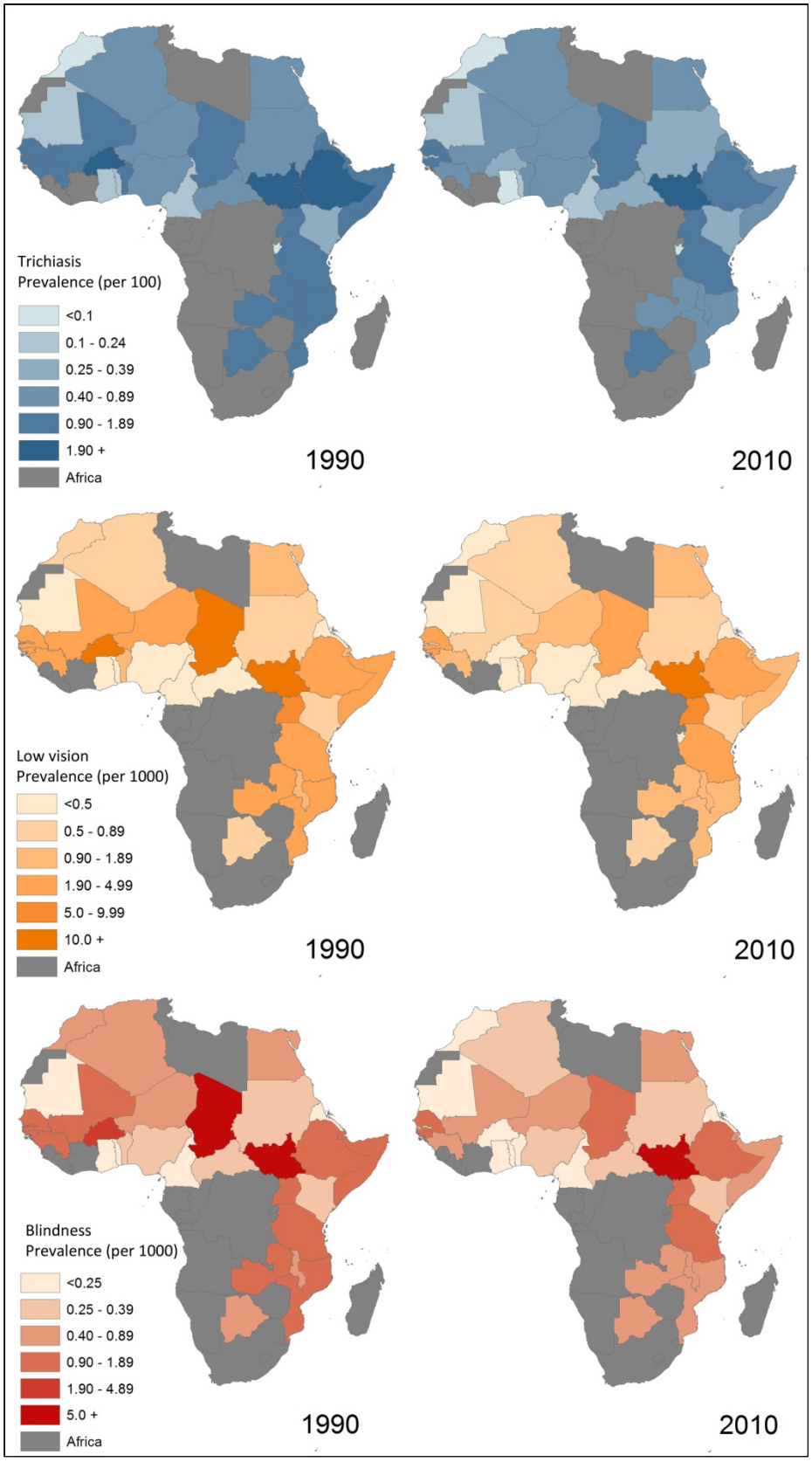



Figure 3.5 Age-standardised prevalence rates for trichiasis, trachomatous low vision and trachomatous blindness in trachoma-endemic countries in Africa for 1990 and 2010. Results particularly highlight a decline in prevalence in West African countries, which may in part be due to well-established trachoma control programmes in these countries.

The un-standardised all-ages prevalence estimates, presented in Tables 3.5-3.6, strongly reflect underlying demographic changes within countries between the two periods. The estimated 20-34% drop in the prevalence of these conditions by 2010 partly reflects real changes in empirical trichiasis and trachomatous blindness data, as discussed above. However, in the majority of countries, changes in prevalence are simply due to differences in the proportion of the population in older age categories and/or gender composition. While the population in Africa has aged overall, some countries (including many highly endemic countries) have experienced large population increases in lower-prevalence age groups which reduces the overall prevalence estimate (Figure 3.6). In addition, changes in the gender ratio in different age groups may reflect trends in mortality or employment migration.

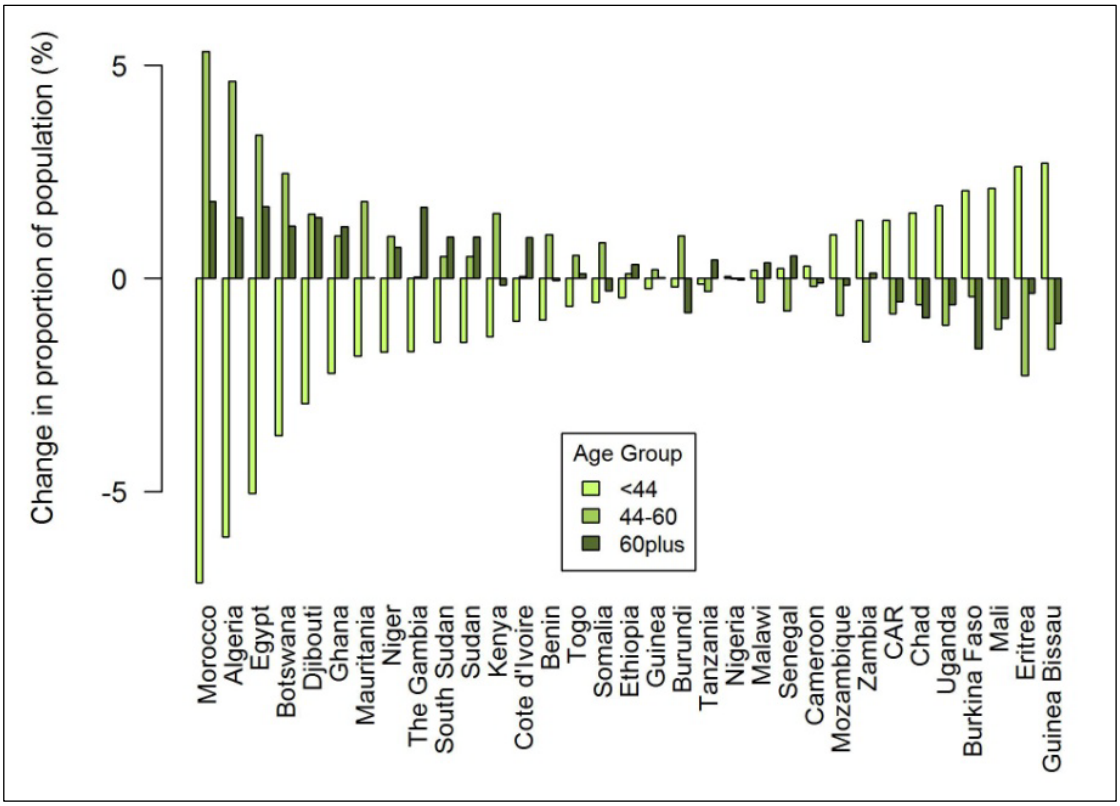



Figure 3.6 Changes in the proportion of the population in three major age categories between 1990 and 2010

Estimates of numbers affected and associated DALYs have increased from 1990 to 2010, due to the rapid population growth in much of Africa. There were an estimated 383,000 cases of trachomatous blindness and 642,000 cases of low vision in Africa due to trachoma in 1990, compared to 405,000 cases of trachomatous blindness and 655,000 cases of low vision in 2010. DALY estimates mirror these patterns (as they are calculated directly from prevalent cases), as detailed by country for 1990 (Table 3.5) and 2010 (Table 3.6). Again, while the total estimated DALYs have increased from 269,000 in 1990 to 310,000 in 2010, this masks a decrease in DALYs in those countries with established control programmes. Approximately half (45%) of estimated DALYs in this analysis are attributed to trichaisis, with the higher prevalence and increased burden in more populous younger age groups outweighing the low disability weight assigned to this condition.

Table 3.5 Country-level estimates of the prevalence, numbers affected and disability-adjusted life years (DALYs) for each disease sequelae: trachomatous trichiasis (TT), trachomatous blindness (TB) and low vision (LV) in 1990

| Country | Pop (000s) | Median Prevalence[a] (95% BCI) | | | Numbers (000s) (95% BCI) | | | DALYs (000s) (95% BCI) | | | Total DALYs | % Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | TT (per 100) | LV (per 1000) | TB (per 1000) | TT | LV | TB | TT | LV | TB | | |
| **North Africa** | | | | | | | | | | | | |
| Algeria | 25283 | 0.48 (0.41,0.55) | 0.86 (0.49, 2.09) | 0.44 (0.29, 0.750) | 121 (105, 139) | 22 (12, 53) | 11 (7, 19) | 4.1 (3.6, 4.7) | 2.4 (1.4, 5.9) | 2.2 (1.4, 3.7) | 8.8 (6.6, 13.8) | 3.3 |
| Egypt[b] | 55137 | 0.81 (0.54, 1.18) | 2.01 (0.69, 8.92) | 0.88 (0.38, 2.31) | 447 (295, 650) | 111 (38, 492) | 49 (21, 127) | 15.2 (10.0, 22.1) | 12.4 (4.3, 55.1) | 9.5 (4.1, 25.0) | 38.0 (20.2, 95.7) | 14.1 |
| Morocco | 24808 | 0.10 (0.01,0.23) | 0.08 (0.02, 0.22) | 0.46 (0.14, 1.09) | 25 (3, 57) | 2 (1, 6) | 11 (3, 27) | 0.9 (0.1, 1.9) | 0.2 (0.1, 0.6) | 2.2 (0.7, 5.3) | 3.4 (1.4, 6.8) | 1.3 |
| **Total[c]** | **109812** | **0.54 (0.37, 0.77)** | **1.22 (0.46, 5.01)** | **0.65 (0.29, 1.58)** | **593 (403, 845)** | **134 (51, 550)** | **71 (32, 173)** | **20.2 (13.7, 28.7)** | **15.1 (5.7, 61.6)** | **13.9 (6.2, 34.0)** | **50.1 (28.2, 116.2)** | **18.6** |
| **Sub-Saharan Africa, Central** | | | | | | | | | | | | |
| CAR | 3008 | 0.45 (0.34, 0.61) | 0.43 (0.21, 1.46) | 0.28 (0.15, 0.64) | 14 (10, 18) | 1 (1, 4) | 1 (1, 1.9) | 0.5 (0.3, 0.6) | 0.1 (0.1, 0.5) | 0.2 (0.1, 0.4) | 0.8 (0.5, 1.5) | 0.3 |
| **Total[c]** | **55164** | **0.03 (0.02, 0.03)** | **0.02 (0.02, 0.07)** | **0.02 (0.02, 0.04)** | **14 (10, 18)** | **1 (1, 4)** | **1 (1, 2)** | **0.5 (0.3, 0.6)** | **0.1 (0.1, 0.5)** | **0.2 (0.1, 0.4)** | **0.8 (0.5, 1.5)** | **0.3** |
| **Sub-Saharan Africa, East** | | | | | | | | | | | | |
| Burundi | 5692 | 0.00 (0.00, 0.00) | 0.00 (0.00, 0.04) | 0.00 (0.00, 0.00) | 0 (0, 0) | 0 (0, 0) | 0 (0, 0) | 0.0 (0.0, 0.0) | 0.0 (0.0, 0.0) | 0.0 (0.0, 0.0) | 0.0 (0.0, 0.0) | 0.0 |
| Djibouti[b] | 561 | 0.70 (0.42, 1.05) | 1.02 (0.29, 5.45) | 0.55 (0.20, 1.98) | 4 (2, 6) | 1 (0, 3) | 0 (0, 1) | 0.1 (0.1, 0.2) | 0.1 (0.0, 0.3) | 0.1 (0.0, 0.2) | 0.3 (0.1, 0.7) | 0.1 |
| Eritrea | 3158 | 0.62 (0.40, 0.88) | 0.12 (0.03, 0.36) | 0.14 (0.04, 0.34) | 19 (13, 28) | 0 (0, 1) | 0 (0, 1) | 0.7 (0.4, 0.9) | 0.0 (0.0, 0.1) | 0.1 (0.0, 0.2) | 0.8 (0.6, 1.1) | 0.3 |
| Ethiopia | 51148 | 1.35 (1.16, 1.54) | 0.43 (0.30, 0.64) | 1.63 (1.27, 2.08) | 689 (595, 789) | 22 (15, 33) | 84 (65, 106) | 23.4 (20.2, 26.8) | 2.4 (1.7, 3.6) | 16.4 (12.7, 20.9) | 42.4 (37.1, 48.0) | 15.7 |
| Kenya | 23447 | 0.17 (0.16, 0.19) | 0.53 (0.20, 2.65) | 0.20 (0.11, 0.61) | 41 (37, 46) | 13 (5, 62) | 5 (3, 14) | 1.4 (1.3, 1.5) | 1.4 (0.5, 7.0) | 0.9 (0.5, 2.8) | 3.7 (2.4, 11.2) | 1.4 |
| Malawi[b] | 9446 | 0.65 (0.37, 1.01) | 1.27 (0.39, 5.61) | 0.61 (0.25, 1.71) | 62 (35, 95) | 12 (4, 53) | 6 (2, 16) | 2.1 (1.2, 3.2) | 1.3 (0.4, 5.9) | 1.1 (0.5, 3.2) | 4.6 (2.3, 11.8) | 1.7 |
| Mozambique[d] | 13544 | 0.92 (0.78, 1.15) | 1.58 (0.83, 4.96) | 0.83 (0.51, 168) | 126 (105, 156) | 21 (11, 67) | 11 (7, 23) | 4.2 (3.6, 5.3) | 2.4 (1.3, 7.5) | 2.2 (1.4, 4.5) | 8.9 (6.5, 16.4) | 3.3 |
| Somalia[d] | 6717 | 0.84 (0.57, 1.19) | 2.03 (0.84, 7.50) | 0.88 (0.45, 2.10) | 57 (38, 80) | 13 (6, 50) | 6 (3, 14) | 1.9 (1.3, 2.7) | 1.5 (0.6, 5.6) | 1.2 (0.6, 2.8) | 4.7 (2.6, 10.5) | 1.7 |
| South Sudan[b] | 4755 | 1.82 (1.65, 2.01) | 10.66 (4.1, 32.3) | 2.92 (1.50, 8.25) | 87 (79, 96) | 51 (20, 153) | 14 (7, 39) | 2.9 (2.7, 3.3) | 5.7 (2.2, 17.2) | 2.7 (1.4, 7.7) | 11.4 (6.6, 27.1) | 4.2 |
| Sudan | 21178 | 0.36 (0.31, 0.44) | 0.54 (0.26, 1.91) | 0.28 (0.17, 0.59) | 77 (65, 93) | 12 (6, 41) | 6 (4, 13) | 2.6 (2.2, 3.2) | 1.3 (0.6, 4.5) | 1.2 (0.7, 2.5) | 5.2 (3.7, 9.8) | 1.9 |

Table 3.5 continued

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Tanzania | 25494 | 0.80 (0.74, 0.86) | 2.00 (1.01, 6.13) | 0.84 (0.53, 1.77) | 204 (189, 220) | 51 (26, 156) | 22 (14, 45) | 6.9 (6.4, 7.5) | 5.7 (2.9, 17.5) | 4.2 (2.6, 8.8) | 16.9 (12.4, 32.8) | 6.3 |
| Uganda | 17841 | 0.73 (0.59, 0.90) | 4.37 (1.42, 13.71) | 1.19 (0.54, 3.73) | 131 (105, 160) | 78 (25, 245) | 21 (10, 67) | 4.4 (3.6, 5.4) | 8.7 (2.8, 27.4) | 4.2 (1.9, 13.0) | 17.5 (8.9, 45.4) | 6.5 |
| Zambia[b] | 8122 | 0.73 (0.54, 0.97) | 1.41 (0.65, 5.23) | 0.69 (0.38, 1.59) | 60 (44, 78) | 11 (5, 43) | 6 (3, 13) | 2.0 (1.5, 2.7) | 1.3 (0.6, 4.8) | 1.1 (0.6,2.5) | 4.4 (2.9, 9.4) | 1.6 |
| **Total[c]** | **212013** | **0.73 (0.62, 0.87)** | **1.34 (0.58, 4.28)** | **0.85 (0.55, 1.66)** | **1554 (1307,1846)** | **285 (122, 907)** | **180 (117, 352)** | **52.8 (44.4, 62.8)** | **31.9 (13.7, 101.6)** | **35.3 (22.9, 69.1)** | **120.8 (86.0, 224.3)** | **44.9** |
| **Sub-Saharan Africa, Southern** | | | | | | | | | | | | |
| Botswana | 1367 | 1.21 (0.40, 3.80) | 0.11 (0.04, 0.22) | 0.44 (0.20, 0.78) | 17 (6, 52) | 0 (0, 0) | 1 (0, 1) | 0.6 (0.2, 1.8) | 0.0 (0.0, 0.0) | 0.1 (0.1, 0.2) | 0.7 (0.3, 1.9) | 0.3 |
| **Total[c]** | **52315** | **0.03 (0.01, 0.10)** | **0.00 (0.00, 0.00)** | **0.02 (0.00, 0.02)** | **17 (6, 52)** | **0 (0, 0)** | **1 (0, 1)** | **0.6 (0.2, 1.8)** | **0.0 (0.0, 0.0)** | **0.1 (0.1, 0.2)** | **0.7 (0.3, 1.9)** | **0.3** |
| **Sub-Saharan Africa, West** | | | | | | | | | | | | |
| Benin | 5179 | 0.99 (0.28, 3.35 ) | 0.16 (0.07, 0.34) | 0.21 (0.03, 0.69) | 51 (15, 174) | 1 (0, 2) | 1 (0, 4) | 1.7 (0.5, 5.9) | 0.1 (0.0, 0.2) | 0.2 (0.0, 0.7) | 2.1 (0.7, 6.5) | 0.8 |
| Burkina Faso | 8871 | 3.24 (3.07, 3.41) | 13.54 (5.8, 43.3) | 4.52 (2.51, 10.40) | 287 (273, 302) | 120 (52, 385) | 40 (22, 92) | 9.8 (9.3, 10.3) | 13.4 (5.8, 43.2) | 7.9 (4.4, 18.1) | 31.1 (19.9, 72.6) | 11.6 |
| Cameroon | 12239 | 0.12 (0.10, 0.13) | 0.23 (0.09, 1.12) | 0.11 (0.05, 0.29) | 14 (13, 16) | 3 (1, 14) | 1 (1, 3) | 0.5 (0.4, 0.5) | 0.3 (0.1, 1.5) | 0.3 (0.1, 0.7) | 1.1 (0.7, 2.7) | 0.4 |
| Chad | 6113 | 1.32 (0.87, 1.98) | 1.79 (1.04, 3.09) | 5.21 (3.62, 7.45) | 80 (53, 121) | 11 (6, 19) | 32 (22, 45) | 2.7 (1.8, 4.1) | 1.2 (0.7, 2.1) | 6.2 (4.3, 8.9) | 10.3 (7.6, 14.0) | 3.8 |
| Cote d'Ivoire[b] | 12780 | 0.02 (0.01, 0.03) | 0.01 (0.00, 0.04) | 0.01 (0.0, 0.02) | 2 (1, 4) | 0 (0, 1) | 0 (0, 0) | 0.1 (0.0, 0.1) | 0.0 (0.0, 0.1) | 0.0 (0.0, 0.1) | 0.1 (0.1, 0.2) | 0.0 |
| Ghana | 15579 | 0.11 (0.10, 0.12) | 0.24 (0.09, 1.30) | 0.10 (0.05, 0.29) | 17 (15, 19) | 4 (1, 20) | 2 (1, 5) | 0.6 (0.5, 0.6) | 0.4 (0.2,2.3) | 0.3 (0.2, 0.9) | 1.3 (0.9, 3.8) | 0.5 |
| Guinea | 6033 | 0.87 (0.61, 1.18) | 2.13 (0.86, 9.17) | 0.93 (0.47, 2.46) | 53 (37, 71) | 13 (5, 55) | 6 (3, 15) | 1.8 (1.3, 2.4) | 1.4 (0.6, 6.2) | 1.1 (0.6, 2.9) | 4.4 (2.5, 11.4) | 1.6 |
| Guinea Bissau | 1017 | 1.14 (0.81, 1.86) | 2.88 (1.18, 14.39) | 1.21 (0.61, 3.96) | 12 (8, 19) | 3 (1,15) | 1 (1, 4) | 0.4 (0.3, 0.6) | 0.3 (0.1, 1.6) | 0.2 (0.1, 0.8) | 1.0 (0.6, 3.0) | 0.4 |
| Mali | 7669 | 1.49 (1.34, 1.65) | 2.95 (1.47, 8.81) | 1.43 (0.86, 2.73) | 114 (103, 126) | 23 (11, 68) | 11 (7, 21) | 3.9 (3.5, 4.3) | 2.5 (1.3, 7.6) | 2.2 (1.3, 4.1) | 8.6 (6.3, 15.8) | 3.2 |
| Mauritania | 1945 | 0.10 (0.03, 0.24) | 0.16 (0.03, 1.02) | 0.09 (0.02, 0.30) | 2 (1, 5) | 0 (0, 2) | 0 (0, 1) | 0.1 (0.0, 0.2) | 0.0 (0.0, 0.2) | 0.0 (0.0, 0.1) | 0.1 (0.0, 0.5) | 0.1 |
| Niger | 7822 | 0.54 (0.33, 0.78) | 1.40 (0.41, 6.88) | 0.58 (0.24, 1.66) | 42 (26, 61) | 11 (3, 54) | 5 (2, 13) | 1.4 (0.9, 2.1) | 1.2 (0.4, 6.0) | 0.9 (0.4, 2.5) | 3.6 (1.7, 10.2) | 1.3 |

Table 3.5 continued

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Nigeria | 94454 | 0.61 (0.58, 0.63) | 0.14 (0.07, 0.29) | 0.23 (0.15, 0.33) | 575 (552, 597) | 13 (6, 25) | 22 (14, 31) | 19.6 (18.8, 20.3) | 1.5 (0.7, 2.8) | 4.2 (2.7, 6.1) | 25.4 (23.4, 27.5) | 9.4 |
| Senegal | 7896 | 1.24 (1.08, 1.43) | 2.44 (1.09, 8.20) | 1.18 (0.66, 2.44) | 98 (85, 113) | 19 (9, 65) | 9 (5, 19) | 3.3 (2.9, 3.8) | 2.2 (1.0, 7.2) | 1.8 (1.0, 3.8) | 7.4 (5.1, 14.4) | 2.7 |
| The Gambia | 962 | 0.70 (0.38, 1.23) | 0.31 (0.14, 0.63) | 1.29 (0.67, 2.13) | 7 (4, 12) | 0 (0, 1) | 1 (1, 2) | 0.2 (0.1, 0.4) | 0.0 (0.0, 0.1) | 0.2 (0.1, 0.4) | 0.5 (0.3, 0.8) | 0.2 |
| Togo | 3961 | 0.08 (0.05, 0.12) | 0.03 (0.01, 0.09) | 0.02 (0.01, 0.07) | 3 (2, 5) | 0 (0, 0) | 0 (0, 0) | 0.1 (0.1, 0.2) | 0.0 (0.0, 0.03) | 0.0 (0.0, 0.1) | 0.1 (0.1, 0.2) | 0.1 |
| **Total[c]** | **199217** | **0.68 (0.60, 0.82)** | **1.11 (0.49, 3.64)** | **0.66 (0.39, 1.29)** | **1357 (1187, 1643)** | **221 (97, 725)** | **131 (78, 256)** | **46.1 (40.4, 55.9)** | **24.8 (10.9, 81.2)** | **25.6 (15.3, 50.1)** | **97.1 (69.8, 183.6)** | **36.1** |
| **Grand Total** | **628521** | **0.56 (0.46, 0.70)** | **1.02 (0.43 3.48)** | **0.61 (0.36, 1.25)** | **3534 (2912, 4405)** | **642 (271, 2187)** | **383 (227, 785)** | **120.2 (99.0, 149.8)** | **71.9 (30.3, 244.9)** | **75.1 (44.5, 153.8)** | **269 (184.9, 527.5)** | **100** |

[a] Prevalence rates are per 100 (TT) or per 1000 (TB and LV) [b] Areas lacking data were assigned the regional mean (as opposed to country average) [c] Regional totals correspond to all countries in the region, assuming a prevalence of zero in all non-endemic countries [d] No TT data, presented estimates correspond to regional average from existing survey data

Table 3.6 Country-level estimates of the prevalence, numbers affected and disability-adjusted life years (DALYs) for each disease sequelae: trachomatous trichiasis (TT), trachomatous blindness (TB) and low vision (LV) in 2010

| Country | Pop (000s) | Median Prevalence[a] (95% BCI) | | | Numbers (000s) (95% BCI) | | | DALYs (000s) (95% BCI) | | | Total DALYs | % Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | TT (per 100) | LV (per 1000) | TB (per 1000) | TT | LV | TB | TT | LV | TB | | |
| **North Africa** | | | | | | | | | | | | |
| Algeria | 35423 | 0.68 (0.63, 0.75) | 0.89 (0.52, 1.86) | 0.52 (0.35, 0.89) | 242 (221, 264) | 32 (18, 66) | 19 (12, 31) | 8.2 (7.5, 9.0) | 3.5 (2.1, 7.4) | 3.6 (2.4, 6.1) | 15.4 (12.5, 22.4) | 5.0 |
| Egypt | 79537 | 1.05 (0.76, 1.39) | 2.13 (0.75, 8.77) | 1.01 (0.46, 2.53) | 838 (602, 1102) | 169 (60, 698) | 80 (37, 201) | 28.5 (20.5, 37.5) | 18.9 (6.7, 78.1) | 15.7 (7.2, 39.4) | 63.6 (38.0, 143.7) | 20.5 |
| Morocco | 32381 | 0.05 (0.00, 0.11) | 0.05 (0.00, 0.37) | 0.03 (0.00, 0.16) | 15 (1, 36) | 2 (0, 12) | 1 (0, 5) | 0.5 (0.0, 1.2) | 0.2 (0.0, 1.3) | 0.2 (0.0, 1.0) | 0.9 (0.1, 3.3) | 0.3 |
| **Total[b]** | **154401** | **0.71 (0.53, 0.91)** | **1.31 (0.51, 5.03)** | **0.64 (0.32, 1.54)** | **1095 (824, 1402)** | **202 (79, 776)** | **99 (49, 237)** | **37.2 (28.0, 47.7)** | **22.6 (8.7, 86.9)** | **19.5 (9.6, 46.5)** | **79.9 (50.6, 169.4)** | **25.8** |
| **Sub-Saharan Africa, Central** | | | | | | | | | | | | |
| CAR | 4592 | 0.34 (0.263, 0.43) | 0.34 (0.17, 0.95) | 0.22 (0.12, 0.46) | 16 (12, 20) | 2 (1, 4) | 1 (1, 2) | 0.5 (0.4, 0.7) | 0.2 (0.1, 0.5) | 0.2 (0.1, 0.4) | 0.9 (0.6, 1.5) | 0.3 |
| **Total[b]** | **98041** | **0.02 (0.01, 0.02)** | **0.02 (0.01, 0.04)** | **0.01 (0.01, 0.02)** | **16 (12, 20)** | **2 (1, 4)** | **1 (1, 2)** | **0.5 (0.4, 0.7)** | **0.2 (0.1, 0.5)** | **0.2 (0.1, 0.4)** | **0.9 (0.6, 1.5)** | **0.3** |
| **Sub-Saharan Africa, East** | | | | | | | | | | | | |
| Burundi | 9553 | 0.00 (0.00, 0.00) | 0.00 (0.00, 0.01) | 0.00 (0.00, 0.00) | 0 (0, 0) | 0 (0, 0) | 0 (0, 0) | 0.0 (0.0, 0.0) | 0.0 (0.0, 0.0) | 0.0 (0.0, 0.0) | 0.0 (0.0, 0.0) | 0.0 |
| Djibouti | 877 | 0.66 (0.28, 1.19) | 0.75 (0.21, 4.27) | 0.46 (0.14, 1.90) | 6 (3, 10) | 1 (0, 4) | 0 (0, 2) | 0.2 (0.1, 0.4) | 0.1 (0.0, 0.4) | 0.1 (0.0, 0.3) | 0.4 (0.1,1.0) | 0.1 |
| Eritrea | 5323 | 0.43 (0.32, 0.55) | 0.12 (0.02, 0.32) | 0.07 (0.02, 0.17) | 23 (17, 29) | 1 (0, 2) | 0 (0, 1) | 0.8 (0.6, 1.0) | 0.1 (0.0, 0.2) | 0.1 (0.0, 0.2) | 0.9 (0.7, 1.2) | 0.3 |
| Ethiopia | 89566 | 1.05 (0.89, 1.23) | 0.44 (0.31, 0.68) | 1.25 (0.96, 1.60) | 936 (800, 1100) | 40 (27, 61) | 112 (86, 143) | 31.8 (27.3, 37.4) | 4.5 (3.1, 6.9) | 22.0 (16.9, 28.0) | 58.8 (51.5, 66.8) | 19.0 |
| Kenya | 40645 | 0.14 (0.12, 0.17) | 0.37 (0.13, 1.58) | 0.15 (0.07, 0.40) | 59 (50, 67) | 15 (5, 64) | 6 (3, 16) | 2.0 (1.7, 2.3) | 1.7 (0.6, 7.2) | 1.2 (0.6, 3.2) | 4.9 (3.0, 12.5) | 1.6 |
| Malawi | 15037 | 0.47 (0.32, 0.69) | 0.65 (0.26, 2.60) | 0.37 (0.17, 0.98) | 70 (48, 104) | 10 (4, 40) | 6 (3, 15) | 2.4 (1.6, 3.5) | 1.1 (0.4, 4.4) | 1.1 (0.5, 2.9) | 4.6 (2.6, 10.3) | 1.5 |
| Mozambique[c] | 22635 | 0.63 (0.52, 0.77) | 0.93 (0.49, 2.23) | 0.51 (0.32, 0.96) | 143 (118, 173) | 21 (11, 50) | 12 (7, 22) | 4.8 (4.0, 5.9) | 2.4 (1.2, 5.6) | 2.3 (1.4, 4.3) | 9.5 (7.1, 15.5) | 3.1 |

Table 3.6 continued

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Somalia[c] | 9486 | 0.48 (0.36, 0.63) | 0.76 (0.37, 2.23) | 0.41 (0.23, 0.87) | 46 (34, 60) | 7 (4, 21) | 4 (2, 8) | 1.6 (1.2, 2.0) | 0.8 (0.4, 2.4) | 0.8 (0.4, 1.6) | 3.2 (2.1, 5.7) | 1.0 |
| South Sudan | 7561 | 1.76 (1.64, 1.89) | 9.03 (2.91, 24.02) | 2.71 (1.25, 6.44) | 133 (124, 143) | 68 (22, 182) | 21 (9, 49) | 4.5 (4.2, 4.8) | 7.6 (2.5, 20.3) | 4.0 (1.9, 9.5) | 16.2 (8.9, 34.3) | 5.2 |
| Sudan | 33670 | 0.37 (0.31, 0.43) | 0.53 (0.26, 1.56) | 0.28 (0.17, 0.61) | 124 (103, 146) | 17 (9, 52) | 10 (6, 21) | 4.2 (3.5, 4.9) | 2.0 (1.0, 5.9) | 1.9 (1.1, 4.0) | 8.1 (6.0, 14.1) | 2.6 |
| Tanzania | 43542 | 0.67 (0.62, 0.73) | 1.43 (0.70, 3.61) | 0.65 (0.40, 1.21) | 292 (269, 319) | 62 (31, 157) | 28 (17, 53) | 9.9 (9.1, 10.9) | 7.0 (3.4, 17.6) | 5.6 (3.4, 10.3) | 22.6 (16.5, 37.9) | 7.3 |
| Uganda | 34040 | 0.52 (0.44, 0.60) | 2.49 (0.82, 7.17) | 0.77 (0.35, 2.09) | 176 (151, 203) | 85 (28, 244) | 26 (12, 71) | 6.0 (5.1, 6.9) | 9.5 (3.4, 17.6) | 5.1 (2.3, 13.9) | 20.7 (11.1, 46.4) | 6.7 |
| Zambia | 12625 | 0.51 (0.41, 0.64) | 0.74 (0.34, 2.13) | 0.41 (0.23, 0.84) | 65 (51, 81) | 9 (4, 27) | 5 (3, 11) | 2.2 (1.7, 2.8) | 1.0 (0.5, 3.0) | 1.0 (0.6, 2.1) | 4.3 (2.9, 7.7) | 1.4 |
| **Total[b]** | **358652** | **0.58 (0.49, 0.68)** | **0.93 (0.40, 2.52)** | **0.64 (0.42, 1.14)** | **2071 (1768, 2435)** | **337 (145, 904)** | **230 (149, 410)** | **70.4 (60.1, 82.8)** | **37.7 (16.3, 101.2)** | **45.1 (29.2, 80.3)** | **154.2 (112.7, 253.5)** | **49.7** |
| **Sub-Saharan Africa, Southern** | | | | | | | | | | | | |
| Botswana | 1953 | 1.38 (0.43, 4.75) | 0.18 (0.08, 0.41) | 0.52 (0.25, 0.97) | 27 (8, 93) | 0 (0, 1) | 1 (0, 2) | 0.9 (0.3, 3.2) | 0.0 (0.0, 0.1) | 0.2 (0.1, 0.4) | 1.2 (0.5, 3.4) | 0.4 |
| **Total[b]** | **70352** | **0.04 (0.01, 0.13)** | **0.00 (0.00, 0.01)** | **0.01 (0.00, 0.02)** | **27 (8, 93)** | **0 (0, 1)** | **1 (0, 2)** | **0.9 (0.3, 3.2)** | **0.0 (0.0, 0.1)** | **0.2 (0.1, 0.4)** | **1.2 (0.5, 3.4)** | **0.4** |
| **Sub-Saharan Africa, West** | | | | | | | | | | | | |
| Benin | 9872 | 0.99 (0.27, 3.13) | 0.17 (0.07, 0.34) | 0.16 (0.03, 0.50) | 97 (27, 309) | 2 (1, 3) | 2 (0, 5) | 3.3 (0.9, 10.5) | 0.2 (0.1, 0.4) | 0.3 (0.0, 1.0) | 3.9 (1.3, 11.1) | 1.2 |
| Burkina Faso | 16097 | 0.17 (0.15, 0.18) | 0.13 (0.07, 0.31) | 0.09 (0.05, 0.18) | 27 (25, 30) | 1 (1, 5) | 1 (1, 3) | 0.9 (0.3, 3.2) | 0.2 (0.1, 0.6) | 0.3 (0.2, 0.6) | 1.4 (1.2, 2.0) | 0.5 |
| Cameroon | 19662 | 0.09 (0.07, 0.10) | 0.17 (0.06, 0.75) | 0.08 (0.04, 0.22) | 17 (15, 20) | 3 (1, 15) | 2 (1, 4) | 0.6 (0.8, 1.0) | 0.4 (0.1, 1.7) | 0.3 (0.1, 0.8) | 1.3 (0.8, 3.1) | 0.4 |
| Chad | 11715 | 0.84 (0.61, 1.12) | 1.73 (0.65, 6.30) | 0.81 (0.38, 2.01) | 99 (72, 132) | 20 (8, 74) | 10 (4, 24) | 3.3 (2.4, 4.5) | 2.3 (0.9, 8.3) | 1.9 (0.9, 4.6) | 7.5 (4.4, 16.8) | 2.4 |
| Cote d'Ivoire | 20375 | 0.02 (0.01, 0.04) | 0.01 (0.00, 0.06) | 0.01 (0.00, 0.03) | 4 (2, 8) | 0 (0, 1) | 0 (0, 1) | 0.1 (0.0, 0.3) | 0.0 (0.0, 0.1) | 0.0 (0.0, 0.1) | 0.2 (0.1, 0.5) | 0.1 |
| Ghana | 24890 | 0.02 (0.01, 0.02) | 0.01 (0.00, 0.03) | 0.01 (0.00, 0.02) | 4 (4, 5) | 0 (0, 1) | 0 (0, 0) | 0.1 (0.1, 0.2) | 0.0 (0.0, 0.1) | 0.0 (0.0, 0.1) | 0.2 (0.2, 0.3) | 0.1 |
| Guinea | 10028 | 0.53 (0.39, 0.68) | 0.88 (0.38, 2.85) | 0.45 (0.24, 1.01) | 53 (39, 68) | 9 (4, 29) | 4 (3, 4) | 1.8 (1.3, 2.3) | 1.0 (0.4, 3.2) | 0.9 (0.5, 2.0) | 3.7 (2.4, 7.2) | 1.2 |
| Guinea Bissau | 1853 | 0.73 (0.53, 1.05) | 1.60 (0.65, 5.50) | 0.71 (0.36, 1.82) | 14 (10, 19) | 3 (1, 10) | 1 (1, 3) | 0.5 (0.3, 0.7) | 0.3 (0.1, 1.1) | 0.3 (0.1, 0.7) | 1.1 (0.7, 2.4) | 0.3 |

Table 3.6 continued

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mali | 13506 | 0.49 (0.44, 0.56) | 0.63 (0.29, 1.96) | 0.36 (0.21, 0.76) | 67 (59, 76) | 9 (4, 27) | 5 (3, 10) | 2.3 (2.0, 2.6) | 1.0 (0.4, 3.0) | 1.0 (0.5, 2.0) | 4.2 (3.1, 7.3) | 1.4 |
| Mauritania | 3363 | 0.10 (0.05, 0.20) | 0.10 (0.03, 0.34) | 0.06 (0.02, 0.17) | 4 (2, 7) | 0 (0, 1) | 0 (0, 1) | 0.1 (0.1, 0.2) | 0.0 (0.0, 0.1) | 0.0 (0.0, 0.1) | 0.2 (0.1, 0.4) | 0.1 |
| Niger | 15791 | 0.40 (0.27, 0.55) | 0.82 (0.29, 3.55) | 0.38 (0.17, 1.01) | 64 (43, 87) | 13 (5, 56) | 6 (3, 16) | 2.2 (1.4, 3.0) | 1.5 (0.5, 6.3) | 1.2 (0.5, 3.1) | 4.9 (2.7, 11.6) | 1.6 |
| Nigeria | 158313 | 0.47 (0.45, 0.49) | 0.15 (0.07, 0.27) | 0.17 (0.11, 0.24) | 748 (714, 780) | 23 (12, 43) | 27 (18, 39) | 25.4 (24, 27) | 2.6 (1.3, 4.8) | 5.2 (3.5, 7.6) | 33.4 (30.7, 36.6) | 10.8 |
| Senegal | 13311 | 1.22 (1.07, 1.37) | 2.24 (1.06, 5.96) | 1.12 (0.64, 2.18) | 162 (143, 182) | 30 (14, 79) | 15 (9, 29) | 5.5 (4.9, 6.2) | 3.3 (1.6, 8.9) | 2.9 (1.7, 5.7) | 11.7 (8.5, 20.1) | 3.8 |
| The Gambia | 1845 | 0.01 (0.00, 0.12) | 0.01 (0.00, 0.09) | 0.00 (0.00, 0.05) | 0 (0, 2) | 0 (0, 0) | 0 (0, 0) | 0.0 (0.0, 0.1) | 0.0 (0.0, 0.0) | 0.0 (0.0, 0.0) | 0.0 (0.0, 0.0) | 0.0 |
| Togo | 7122 | 0.07 (0.04, 0.10) | 0.03 (0.01, 0.09) | 0.02 (0.01, 0.07) | 5 (3, 7) | 0 (0, 1) | 0 (0, 1) | 0.2 (0.1, 0.3) | 0.0 (0.0, 0.1) | 0.0 (0.0, 0.1) | 0.2 (0.1, 0.4) | 0.1 |
| **Total[b]** | **338971** | **0.40 (0.34, 0.51)** | **0.33 (0.15, 1.02)** | **0.22 (0.12, 0.43)** | **1363 (1153, 1730)** | **114 (50, 345)** | **73 (41, 145)** | **46.3 (39.2, 58.8)** | **12.8 (5.6, 38.6)** | **14. 4 (8.1, 28.5)** | **73.9 (56.2, 119.9)** | **23.8** |
| **Grand Total** | **1020417** | **0.45 (0.37, 0.56)** | **0.64 (0.27, 1.99)** | **0.40 (0.31, 0.60)** | **4572 (3766, 5680)** | **655 (274, 2029)** | **412 (312, 612)** | **155.5 (128.0,** | **73.4 (30.7, 227.3)** | **79.4 (47.1, 156.1)** | **310 (220.6, 547.9)** | **100** |

[a] Prevalence rates are per 100 (TT) or per 1000 (TB and LV) [b] Areas lacking data were assigned the regional mean (as opposed to country average) [c] Regional totals correspond to all countries in the region, assuming a prevalence of zero in all non-endemic countries [d] No TT data, presented estimates correspond to regional average from existing survey data

As illustrated in Figure 3.7, a substantial proportion of the estimated burden of disease is in Egypt, Ethiopia and Nigeria. While data in the latter two countries are robust, data from Egypt are geographically limited, and much of the projected burden is due to a greater life expectancy and increased populations in older age groups. As a result, these estimates are particularly sensitive to the method of extrapolation, as presented in the next section.

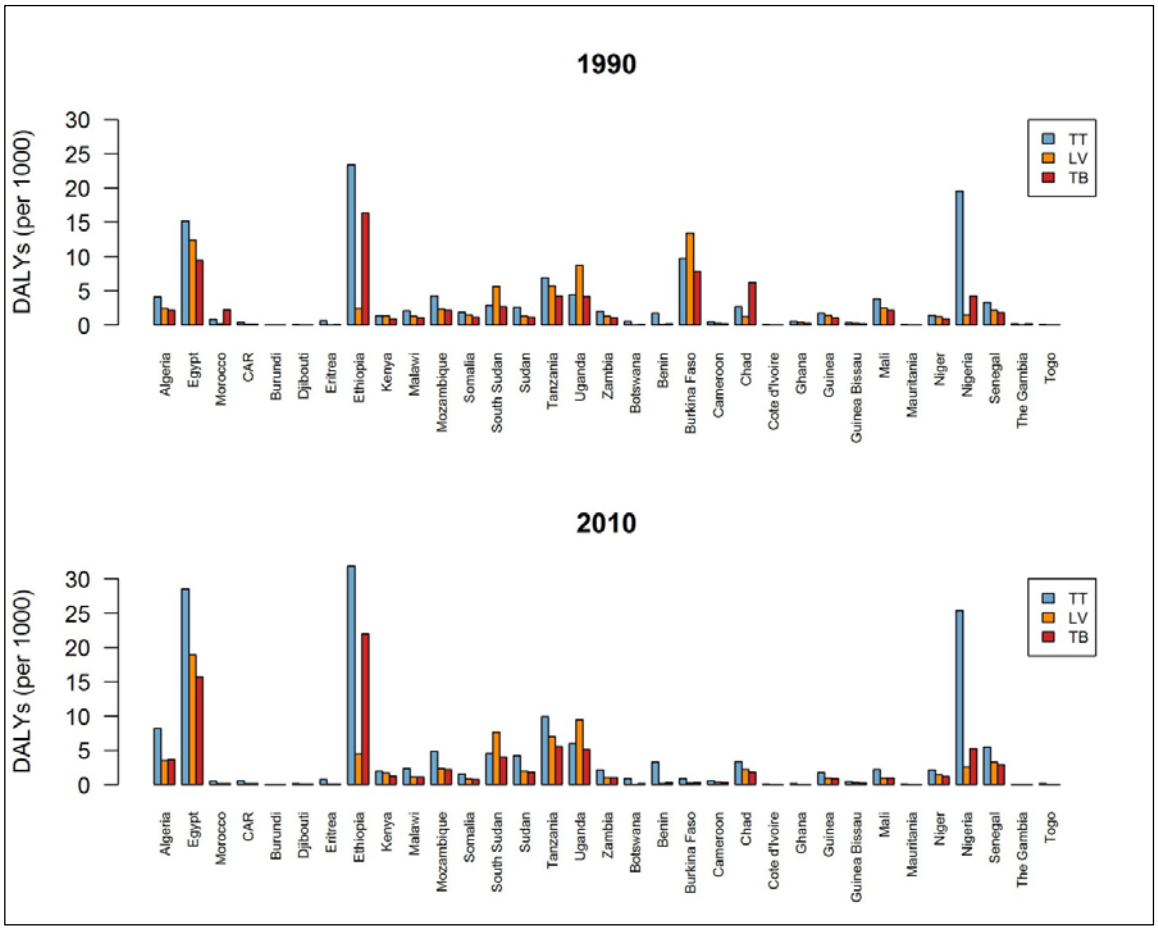



Figure 3.7 Disability-adjusted life year (DALY) estimates in 1990 and 2010 in 33 African countries for trachomatous trichiasis (TT), trachomatous low vision (LV) and trachomatous blindness (TB)

Finally, age- and sex-specific estimates are presented by region in Tables 3.7-3.8. A higher proportion of DALYs (66%) are in females and age-specific burden estimates clearly reflect both the modelled age-prevalence distribution of trachomatous sequelae and differences in the underlying population distribution between regions.

Table 3.7 Regional estimates of the population and total disability-adjusted life years (DALYs) attributed to trachoma by gender and age in 1990

| | North Africa | | | Sub-Saharan Africa, Central | | | Sub-Saharan Africa, East | | | Sub-Saharan Africa, Southern | | | Sub-Saharan Africa, West | | | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Factor** | **Pop (000)** | **DALYs (000)** | **Prop. DALY** | **Pop (000)** | **DALYs (000)** | **Prop. DALY** | **Pop (000)** | **DALYs (000)** | **Prop. DALY** | **Pop (000)** | **DALYs (000)** | **Prop. DALY** | **Pop (000)** | **DALYs (000)** | **Prop. DALY** | **DALYs (000)** |
| **Gender** | | | | | | | | | | | | | | | | |
| Male | 55288 | 17.1 (10.0, 40.2) | 35 | 27156 | 0.2 (0.2, 0.5) | 32 | 104997 | 41.7 (31.5, 71.0) | 34 | 25833 | 0.2 (0.1, 0.7) | 32 | 99293 | 33.6 (25.3, 57.7) | 34 | 92.8 (67.0, 170.0) |
| Female | 54524 | 32.4 (18.9, 76.4) | 65 | 28008 | 0.5 (0.3, 1.0) | 68 | 107016 | 81.3 (62.1, 136.0) | 66 | 26482 | 0.5 (0.2, 1.4) | 68 | 99924 | 65.5 (50.2, 109.5) | 66 | 180.2 (131.7, 324.2) |
| **Age** | | | | | | | | | | | | | | | | |
| 0-4 | 16158 | 0.3 (0.2, 0.5) | 1 | 10445 | 0.0 (0.0, 0.0) | 1 | 39040 | 1.3 (1.0, 1.6) | 1 | 7595 | 0.0 (0.0, 0.0) | 2 | 36978 | 1.1 (0.9, 1.4) | 1 | 2.7 (2.0, 3.5) |
| 5-14 | 29097 | 1.5 (1.0, 2.8) | 3 | 15216 | 0.0 (0.0, 0.0) | 4 | 58932 | 4.5 (3.6, 6.7) | 4 | 13726 | 0.0 (0.0, 0.1) | 6 | 54628 | 3.8 (3.0, 5.8) | 4 | 9.9 (7.6, 15.6) |
| 15-29 | 29731 | 4.4 (3.1, 8.0) | 9 | 14372 | 0.1 (0.1, 0.1) | 10 | 56269 | 12.9 (10.7, 18.4) | 11 | 14727 | 0.1 (0.0, 0.3) | 15 | 51422 | 10.3 (8.7, 15.1) | 10 | 27.1 (22.5, 42.0) |
| 30-44 | 18175 | 10.2 (6.5, 22.4) | 21 | 8060 | 0.2 (0.1, 0.3) | 20 | 31309 | 25.9 (20.4, 43.8) | 21 | 8853 | 0.2 (0.1, 0.5) | 24 | 29619 | 19.2 (15.3, 31.2) | 19 | 55.6 (42.4, 98.2) |
| 45-59 | 10125 | 14.9 (9.1, 34.2) | 30 | 4548 | 0.2 (0.1, 0.4) | 29 | 16921 | 37.0 (28.5, 62.0) | 30 | 4782 | 0.2 (0.1, 0.5) | 26 | 16859 | 27.9 (21.7, 45.8) | 28 | 80.2 (59.5, 142.9) |
| 60-69 | 4111 | 10.5 (6.4, 23.6) | 21 | 1650 | 0.2 (0.1, 0.3) | 20 | 6239 | 23.6 (18.4, 40.2) | 19 | 1687 | 0.1 (0.1, 0.3) | 16 | 6238 | 19.1 (14.8, 32.2) | 19 | 53.5 (39.7, 96.6) |
| 70plus | 2415 | 8.0 (4.9, 17.5) | 16 | 873 | 0.1 (0.1, 0.2) | 16 | 3303 | 16.7 (12.7, 27.8) | 14 | 945 | 0.1 (0.1, 0.2) | 11 | 3473 | 16.7 (12.6, 30.0) | 17 | 41.6 (30.3, 75.8) |
| | | | | | | | | | | | | | | | | |
| **Total** | **109812** | **49.8 (31.4, 107.8)** | **100** | **55164** | **0.8 (0.5, 1.4)** | **100** | **212013** | **121.5 (97,198)** | **100** | **52315** | **0.7 (0.3, 1.9)** | **100** | **199217** | **98.4 (78.9, 158.8)** | **100** | **271 (209, 468)** |

Table 3. 8 Regional estimates of the population and total disability-adjusted life years (DALYs) attributed to trachoma by gender and age in 2010

| | North Africa | | | Sub-Saharan Africa, Central | | | Sub-Saharan Africa, East | | | Sub-Saharan Africa, Southern | | | Sub-Saharan Africa, West | | | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Factor** | **Pop (000)** | **DALYs (000)** | **Prop. DALY** | **Pop (000)** | **DALYs (000)** | **Prop. DALY** | **Pop (000)** | **DALYs (000)** | **Prop. DALY** | **Pop (000)** | **DALYs (000)** | **Prop. DALY** | **Pop (000)** | **DALYs (000)** | **Prop. DALY** | **DALYs (000)** |
| **Gender** | | | | | | | | | | | | | | | | |
| Male | 77207 | 27.2 (16.9, 61.0) | 34 | 48552 | 0.3 (0.2, 0.5) | 32 | 178459 | 53.1 (39.5, 78.9) | 34 | 34729 | 0.4 (0.1, 1.2) | 33 | 169797 | 25.9 (20.8, 36.8) | 34 | 106.9 (77.6, 178.3) |
| Female | 77194 | 52.6 (32.5, 114.9) | 66 | 49489 | 0.62 (0.4, 1.1) | 68 | 180193 | 103.8 (79.5,151.6) | 66 | 35623 | 0.8 (0.3, 2.4) | 67 | 169174 | 49.3 (40.7, 69.7) | 66 | 207.0 (153.4, 339.6) |
| **Age** | | | | | | | | | | | | | | | | |
| 0-4 | 16252 | 0.4 (0.2, 0.5) | 0 | 18500 | 0.0 (0.0, 0.0) | 1 | 58986 | 1.3 (1.1, 1.9) | 1 | 7765 | 0.0 (0.0, 0.0) | 1 | 54793 | 0.9 (0.7, 1.3) | 1 | 2.6 (2.0, 3.8) |
| 5-14 | 29949 | 1.6 (1.1, 2.6) | 2 | 27060 | 0.0 (0.0, 0.0) | 4 | 96420 | 5.5 (4.4, 7.4) | 4 | 15361 | 0.0 (0.0, 0.1) | 3 | 89968 | 3.2 (2.6, 4.1) | 4 | 10.3 (8.2, 14.3) |
| 15-29 | 44658 | 6.5 (4.7, 10.6) | 8 | 26709 | 0.1 (0.1, 0.2) | 12 | 102008 | 16.9 (14.2, 22.1) | 11 | 21521 | 0.2 (0.1, 0.5) | 14 | 96304 | 9.4 (8.1, 11.4) | 12 | 33.0 (27.2, 44.8) |
| 30-44 | 31226 | 15.5 (10.5, 31.8) | 19 | 14358 | 0.2 (0.1, 0.3) | 20 | 55577 | 33.6 (26.7, 49.0) | 22 | 12938 | 0.3 (0.1, 0.8) | 22 | 52614 | 16.3 (13.9, 21.0) | 22 | 65.8 (51.3, 102.8) |
| 45-59 | 20555 | 26.4 (17.0, 57.4) | 33 | 7325 | 0.3 (0.2, 0.4) | 28 | 29017 | 45.4 (35.6, 66.3) | 29 | 7904 | 0.3 (0.1, 1.0) | 29 | 28714 | 21.5 (17.9, 28.8) | 29 | 93.9 (70.9, 153.9) |
| 60-69 | 6897 | 15.6 (9.9, 33.9) | 20 | 2607 | 0.2 (0.1, 0.3) | 18 | 10359 | 29.3 (23.0, 42.6) | 19 | 2963 | 0.2 (0.1, 0.5) | 16 | 10223 | 13.0 (10.8, 17.5) | 17 | 58.2 (43.9, 94.8) |
| 70plus | 4864 | 13.6 (8.8, 28.4) | 17 | 1482 | 0.2 (0.1, 0.3) | 17 | 6285 | 23.0 (17.2, 33.7) | 15 | 1900 | 0.2 (0.1, 0.4) | 14 | 6357 | 11.1 (9.0, 14.9) | 15 | 48.0 (35.3, 77.7) |
| | | | | | | | | | | | | | | | | |
| **Total** | **109812** | **79.6 (53.3, 160.7)** | **100** | **55164** | **0.9 (0.6, 1.5)** | **100** | **212013** | **155.1 (126, 219)** | **100** | **52315** | **1.2 (0.5, 3.4)** | **100** | **199217** | **75.3 (64.8, 97.4)** | **100** | **312 (245, 482)** |

*Sensitivity analyses*

Although Bayesian credible intervals reported alongside mean estimates capture underlying uncertainty in sample sizes and modelled relationships, they do not provide information on how extrapolation methods impact resulting estimates. The results from the sensitivity analyses highlighted the importance of this issue and the unreliability of estimates based off country averages where geographical coverage is low and/or likely to be unrepresentative. Scenario 3 extrapolated existing data within all countries, as opposed to using regional averages where data were geographically limited, and effectively doubled estimates (Table 3.9). This problem was particularly obvious in certain countries, such as Egypt, where subnational data are limited, highly endemic and the population is rapidly aging.

The choice of disability weights also had a clear impact on DALY estimates, particularly for trichiasis and low vision due to their higher prevalence in younger (and more populous) age groups. Using disability weights from the previous GBD 2000 study increased the estimated total DALYs more than two-fold, from 273,000 to 735,000 in 1990 and 314,000 to 796,000 in 2010.

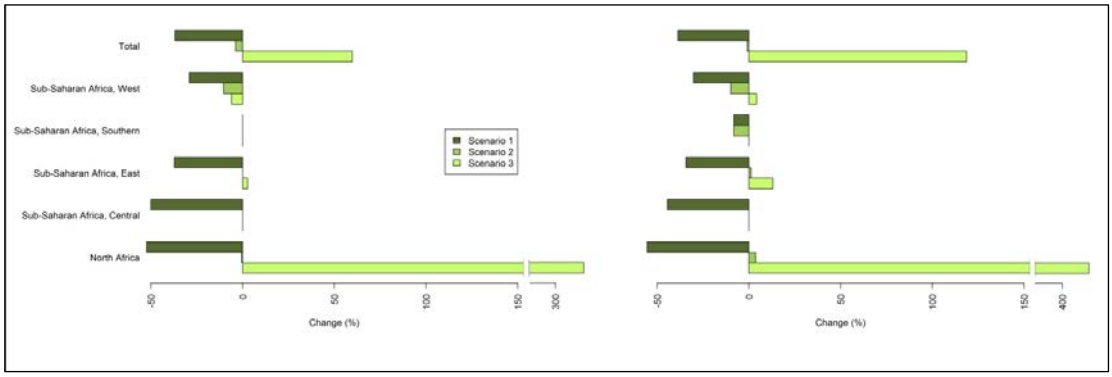



Figure 3.8 Percent change in disability-adjusted life year (DALY) estimates from base case using different methods of extrapolation in 1990 (A) and 2010 (B)

Table 3.9 Estimated disability-adjusted life years (DALYs) by region using existing data only (Scenario 1), extrapolation of regional averages to all subnational areas lacking data (Scenario 2), and extrapolation of country averages to all subnational areas (Scenario 3)

| | 1990 | | | | | | | | 2010 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Base Case | | Sensitivity Analysis | | | | | | Base Case | | Sensitivity Analysis | | | | | |
| | | | Scenario 1 | | Scenario 2 | | Scenario 3 | | | | Scenario 1 | | Scenario 2 | | Scenario 3 | |
| Region | DALYs | % total | DALYs | % total | DALYs | % total | DALYs | % total | DALYs | % total | DALYs | % total | DALYs | % total | DALYs | % total |
| North Africa | 49.8 (31.4, 107.8) | 16.9 | 23.7 (12.2, 82.7) | 13.9 | 49.5 (28.1, 117.8) | 19.0 | 214.0 (106.8, 568.0) | 49.4 | 79.6 (53.3, 160.7) | 24.3 | 35.4 (16.6, 150.0) | 18.5 | 82.5 (51.6, 212.8) | 26.7 | 426.5 (231, 1201) | 62.5 |
| Sub-Saharan Africa, Central | 0.8 (0.5, 1.4) | 0.3 | 0.4 (0.3, 0.8) | 0.3 | 0.8 (0.5, 1.5) | 0.3 | 0.8 (0.5, 1.5) | 0.2 | 0.9 (0.6, 1.5) | 0.3 | 0.5 (0.4, 0.9) | 0.3 | 0.9 (0.6, 1.6) | 0.3 | 0.9 (0.6, 1.6) | 0.1 |
| Sub-Saharan Africa, East | 121.5 (98,198) | 44.8 | 76.2 (52.4, 174.4) | 44.6 | 121.4 (89.6, 186.8) | 46.7 | 124.7 (89.9, 215.8) | 28.8 | 155.1 (126, 219) | 51.4 | 101.8 (69.1, 205.1) | 53.3 | 156.9 (118, 278) | 50.8 | 175.3 (128, 315) | 25.7 |
| Sub-Saharan Africa, Southern | 0.7 (0.3, 1.9) | 0.3 | 0.7 (0.2, 2.2) | 0.4 | 0.7 (0.2, 2.1) | 0.3 | 0.7 (0.2, 2.2) | 0.2 | 1.2 (0.5, 3.4) | 0.5 | 1.1 (0.4, 3.4) | 0.6 | 1.1 (0.4, 3.5) | 0.4 | 1.2 (0.4, 3.1) | 0.2 |
| Sub-Saharan Africa, West | 98.4 (78.9, 158.8) | 36.3 | 69.7 (51.5, 147.4) | 40.8 | 88.1 (67.9, 139.3) | 33.9 | 92.4 (71.0, 143.8) | 21.3 | 75.3 (64.8, 97.4) | 23.6 | 52.5 (41.0, 86.3) | 27.5 | 67.8 (54.5, 103.3) | 22.0 | 78.4 (62.9, 122.3) | 11.5 |
| **Total** | **271 (208, 467)** | **100** | **170.8 (117, 407)** | **100** | **260.5 (186, 447)** | **100** | **433 (269, 931)** | **100** | **312 (245, 482)** | **100** | **191.3 (127, 445)** | **100** | **309 (225, 599)** | **100** | **682.2 (422,1642)** | **100** |

## 3.6 Discussion

Based on the methodology presented this chapter, trachomatous blindness and low vision were responsible for an estimated 147,000 DALYs in Africa in 1990 and 152,800 in 2010. The additional burden contributed by trichiasis only, regardless of vision status, add a further 155,500 DALYs to the 2010 estimates. The majority of the DALYs associated with trachoma (63%) are, as expected, among individuals aged over 45 years, due to disease sequelae arising as a result of cumulative episodes of active infection over time. The relatively large burden accounted for by trichiasis and low vision, despite their lower disability weight, may be attributed to the younger age-prevalence profiles of these conditions compared to trachomatous blindness. A higher proportion (66%) of DALYs were experienced by females, who tend to make up a larger proportion of the population (particularly in older age groups) and who experience a higher risk of trachoma [113]. The key advantages of the methodology used in this chapter are fourfold: i) it provides a substantially broader evidence base from which to generate estimates, ii) it explicitly allows within and between country variation in risk to be incorporated into overall burden estimates, iii) estimates can be easily refined as new data become available with scale up of trachoma mapping activities and post-intervention surveys and iv) it provides a framework in which sub-national trichiasis data might, in the future, be predicted from socioeconomic or environmental covariates.

DALY estimates are similar to those presented in the GBD 2010, which calculated 50,000 DALYs for trachomatous blindness within Africa in 1990 (compared to 75,100) and 68,000 in 2010 (compared to 79,400) [71]. DALY estimates for low vision are higher in this analysis, due to the different age distribution used to calculate prevalent cases. The methodology used for the GBD 2010 estimates was based on that presented here. However, country-level prevalence estimates were further modelled as causal blindness fractions by the Vision Loss Expert Group to ensure that cause-specific estimates fit within the all-cause blindness prevalence “envelope”. Furthermore, country-level models were

fitted with a regional temporal trend, informed by the average year that surveys were conducted, in order to predict separate estimates for all countries at both time points. On the whole, earlier estimates generated by Ranson & Evans (1996) [235], Frick et al. (2003) [80] and the 2000 GBD study [82] have generated much higher estimates of the number of cases and associated DALYs, in the order of 1-2 million DALYs. However, in addition to a higher disability weight used for these calculations, these studies have relied on rather broad modelling assumptions based on limited empirical data, mainly generated in the 1980s. These methods included i) extrapolating very few national estimates to all countries within a region or ii) modelling regional estimates for different time points using national GDP indicators and very limited national and sub-national blindness survey data.

As might be expected, the rapidly changing demographic profiles in many African countries have an important impact on overall prevalence of these sequelae and burden estimates. In some countries, particularly North Africa and Southern Africa, the population structure has aged between 1990 and 2010. In these contexts, the overall prevalence and burden will increase assuming a constant age-specific prevalence. In other countries where the population has grown with minimal gains in life expectancy, the overall prevalence will decrease while the burden will increase, in terms of numbers affected and DALYs. An important consideration, is whether younger populations in these countries will be at risk of developing disease sequelae in 20-30 years time. The answer is probably yes, in the absence of successful interventions, in countries which remain undeveloped or have significant inequities in development, and where endemicity of active disease remains high. Unfortunately, information tends to be scarcer in these contexts than in countries which are rapidly experiencing development or that have established control programmes (such as Oman, Morocco, Mali, and Burkina Faso). Other countries, however, are making progress in reducing the burden of trachomatous sequlae which is to some degree masked by these demographic changes and better seen by looking at age-specific or age-standardised rates. The estimates presented in this chapter suggest a decline in estimated DALYs in west Sub-Saharan Africa from 98,000 to 75,000 between 1990 and

2010 (Tables 3.5 & 3.6), which may be attributed to large-scale implementation of control in many of these countries.

A key finding of this analysis was the sensitivity of estimates to the method of extrapolation and choice of disability weight. While relatively minor changes were observed by basing estimates on known data or regional extrapolations, the sensitivity analysis highlighted the unreliability of estimates based on country averages where geographical coverage is low and/or likely to be unrepresentative. Future work should focus on identifying optimal methods of extrapolating reliable data, possibly modelling trichiasis data using sub-national socioeconomic or environmental data. Also, as no discounting was applied in this analysis the reported DALYs are directly proportional to the number of prevalent cases and the disability weight. Since discounting, which devalues future health states, was not included into the 2010 GDB study, the choice of disability weight and number of prevalent cases at a given time point are the sole inputs into DALYs for trachoma and have an important impact on resulting estimates. Assigned weights for blindness and low vision from the GBD 2010 were half of those used previously, which has already generated substantial debate [72,242]. Valuation of these health states pose a challenge, however, and developing countries are generally underrepresented in valuation studies [237]. Finally, the DALYs attributed to trachoma are only calculated from YLDs, despite an increased risk of mortality associated with blindness and low vision. A study in Africa found the standardized mortality rate to be 3.8 times higher among females who were blind and 2.5 times higher in blind males compared to those with normal vision [243]. Low vision was associated with 1.5 times the risk of mortality in females and 1.4 times higher in males.

It is important to recognize a number of limitations around the data used to inform this analysis as well as model assumptions, which will have an important impact on resulting estimates. First, variation in risk of trachoma (and disabling sequelae) between and within countries is a hallmark of the disease. This will limit the representativeness of available

data, from both causal blindness surveys within Africa and sub-national trachoma surveys, when extrapolated to the broader region or country. National causal blindness surveys may be targeted to areas that are either expected to bear a high burden of these conditions or conducted in countries with well-established blindness programmes (such as Morocco, the Gambia, and Botswana). It is also likely that areas expected to be trachoma endemic are surveyed first by control programmes and may introduce a positive bias when there is poor geographical coverage within a country, overestimating the prevalence of trichiasis and trachomatous blindness. This bias was minimized by using information from TRA to exclude the population in districts where trachoma has not been found to be a public health problem or are suspected to be nonendemic based on hospital data. Second, data which are derived from causal blindness surveys may provide relatively imprecise estimates for trachomatous blindness as sample size estimation is based on the total expected prevalence of blindness at the national level. Third, trachomatous corneal opacity (CO) often coexists with other conditions (glaucoma for example), particularly in elderly populations. Where the cause of blindness is assigned to another coexisting condition, estimates will underestimate the visual impairment due to trachoma. Fourth, there are limited data to inform models and assess the impact of potential explanatory factors on observed relationships. Longitudinal studies in the Gambia suggest conjunctival inflammation (with or without associated infection), frequency of epilation, ocular dryness and secondary infections may influence rates of progressive TT [18,19]. Thus, improved models might incorporate information on endemicity of active infection and associated contextual factors which may influence disease progression and explain variation between studies. Fifth, it is likely that urban populations will have a lower risk of trichiasis and trachomatous blindness compared to rural populations. While the sampling strategy for national causal blindness surveys typically use a probability proportional to size design, allowing a representative national sample, trachoma surveys often exclude urban populations from the sampling frame and so may overestimate the risk in urban areas.

Finally, the geographical level of input data has an important impact on resulting prevalence estimates, due to the non-linearity of these models. Countries with data at a smaller scale (i.e. district) will provide more information on variation in risk within a country than those countries which have an average value assigned to all districts within a larger area. As a consequence, countries with more extreme district-level TT values from the positive tail of the distribution (such as South Sudan or Uganda) will predict very high estimates of district-level trachomatous blindness and result in higher overall estimates compared to countries with a more moderate distribution of risk (Zambia) or where data are at the region (Chad) or national (Morocco) level. Ideally, when estimates are available only for higher geographical areas (such as province or country) it would be preferable to assume that district-level prevalence estimates arise from a defined distribution. Unfortunately, without incorporating this within an informative spatial (such as a conditional-autoregressive model) or multilevel modelling framework that is able to draw information from surrounding districts, assigning these estimates at random is likely to increase uncertainty in the resulting estimates due to variation in the underlying population distribution.

Unlike previous estimates, this approach did not incorporate a temporal trend or assume that changes in the GDP over time will be reflected in the prevalence of trachomatous blindness. While secular changes and socioeconomic improvements are certainly associated with a decline in the burden of trachoma, national changes in GDP may occur more quickly than a corresponding improvement in socioeconomic status and sanitation in populations at risk or subsequent impact on trachomatous sequelae, which tend to "backlog". Empirically, the relationship between GDP and national trachomatous blindness estimates was found to be

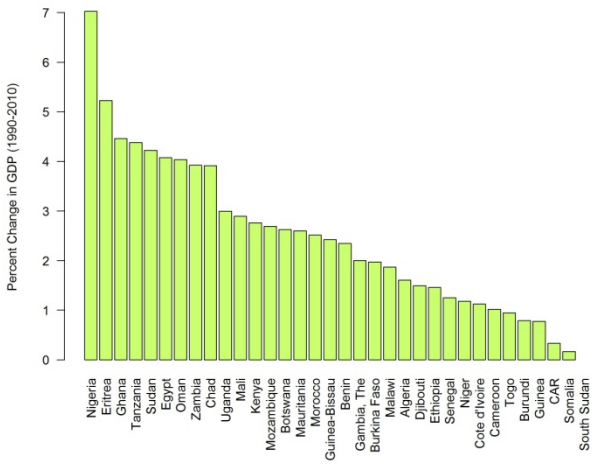



Figure 3.9 Percent change in national GDP between 1990 and 2010 in trachoma endemic countries. NB: no data available for South Sudan

highly variable and, as illustrated in Figure 3.9, changes in GDP between 1990 and 2010 do not correspond well to demographic changes observed in Figure 3.6 or heterogeneity between countries in terms of control activities, water availability and personal hygiene in trachoma endemic areas. While there has been strong economic growth in Africa over the last decade, there are a number of reasons why it might be a poor indicator for poverty reduction. One reason is that GDP includes sectors outside of household income, such as government and external balances, which usually grow faster. Second, corresponding increases are not necessarily seen all along the income distribution due to high levels of inequality within countries, which in some cases have increased along with national GDP [244].

Implementation of the A, F and E components of the SAFE strategy should decrease rates of blindness over a longer time period, as transmission decreases and the cumulative effect of repeated infection on the incidence of trichiasis and visual impairment is relieved. Surgical correction of trichiasis will have a direct impact on trachomatous blindness rates, however surgical intervention for trichiasis has made slow gains compared to MDA of antibiotics and up to 40% of trichaisis recurs within four years of primary correction [245]. Development of more sophisticated modelling techniques, possibly incorporating mathematical modelling, may allow for estimation of the variable impact of control activities and secular changes on transmission of *C trachomatis* and disease progression.

This chapter proposes a framework to utilise available trichiasis data from the GAT to estimate the burden of trachoma where national trachomatous blindness estimates are lacking. This approach is particularly timely as the global mapping of trachoma is scaled up and high resolution pre- and post-intervention data are increasingly available to monitor the changing burden of disease. As discussed above, models should continue to be refined in order to capture the variable impact of control and, importantly, alternative extrapolation methods should be explored to address gaps in geographical coverage of existing data. The next chapter will explore associations between climatic and

environmental factors and large-scale patterns of trichiasis risk that may inform potential risk mapping to refine future estimates of the burden of trachoma.

# Chapter 4: Multilevel analysis of trachomatous trichiasis and corneal opacity in Nigeria: the role of environmental risk factors on the distribution of disease

## 4.1 Overview

As highlighted in chapters 2 and 3, a continuing limitation of efforts to quantify the burden of trachoma in terms of long term disease sequelae, including trachomatous trichiasis (TT) and vision loss, is the sparseness of data in space and time. The method by which existing data are extrapolated has a profound impact on resulting estimates. The distribution of trachoma is noted to be spatially heterogeneous, with large-scale trends observed across countries (Chapter 2; [194]) and more local variation within areas. However, relative contributions of individual and cluster-level risk factors to the geographic distribution of disease remain largely unknown. Better identification of determinants of TT within countries and geographic regions may provide a basis for more reliable extrapolation of data and refined estimation of the burden of disease.

This chapter uses multilevel modelling to quantify the relationship between climatic factors and TT and/or corneal opacity (CO) due to trachoma in Nigeria using data from the 2007 National Blindness and Visual Impairment Survey. This chapter aims to establish the importance of large scale geographical risk factors for later stages of trachoma, adjusting for individual-level risk factors, in order to highlight the potential use of risk mapping to improve estimation of the burden of disease at large scales. The next chapter will investigate geographic risk factors for active disease (TF), using nationwide data from Kenya.

## 4.2 Introduction

Although a number of studies have identified risk factors that are responsible for clustering of trachoma within villages, households [97,106,108,179] and individuals [246], only a few studies have quantified associations at larger scales. Anecdotally, trachoma is believed to be a greater public health risk in dry, dusty and hot settings. Climatic variables are postulated to indirectly influence the transmission of trachoma through the following mechanisms: low rainfall which leads to reduced access or use of water for washing faces; higher temperatures which may influence the distribution and activity of the putative vector *Musca sorbens*; and climatic conditions that favour drying of faeces, the fly's preferred breeding site [135,139,142,143]. In addition, there may be a potential role for ocular dryness or environmental irritants to contribute to progression of chronic disease, by aggravating scarring processes [18,19,154,247].

However, robust studies investigating relationships between detailed epidemiological observations and environmental determinants are scarce. Existing studies, recently reviewed in full by Ramesh et al. (2013) [161], provide some support for a role of temperature and rainfall in the distribution of trachoma [152,153,157,158], as well as altitude (which might be a proxy for temperature) [159,178,248]. However, most studies are limited by lack of control for individual level factors [158,159,249], and in particular variation in socioeconomic factors [250]. In practice, it is difficult to disentangle the effects of risk factors of trachoma at different spatial levels, due to a complex interplay between large-scale factors such as climate, and mediating factors at smaller scales, like water availability and sanitation at the household level and individual behaviours, including household water use and personal hygiene [131,132,178].

Bayesian hierarchical models (BHM) are a robust and well established methodology for modelling data that are naturally grouped and identifying risk factors at different scales [251]. This approach can be expanded to incorporate information on residual underlying spatial patterns, thus explicitly addressing any remaining spatial correlation between

observations [252]. Previous studies have used this approach to identify risk factors at multiple hierarchical levels for various tropical diseases, using data from school-based and community surveys, including malaria [253], soil-transmitted helminths [190,254], schistosomiasis [255,256] and trachoma [152]. A common application in multilevel models is to then apportion the variance in the response according to the different levels of the data, referred to as the variance partition coefficient (VPC) [257,258]. These methods offer a robust and flexible approach to modelling prevalence data routinely collected as part of disease control programmes in developing countries.

Nigeria is a populous country with over 160 million people, comprising approximately 20% of the total population in Africa [259]. There are diverse climatic conditions across the country, and three broad ecological zones: the southern rainforest zone, the central Guinea Savannah zone and the semi-arid northern Sudan Savannah [260]. Trachoma is a significant public health problem in the north of the country and currently only 43% of districts suspected to be endemic have been surveyed by population based prevalence surveys [261]. The 2007 National Blindness and Visual Impairment Survey was conducted in Nigeria to provide evidence on the prevalence and causes of blindness at the national level in order to inform policy and planning for the elimination of avoidable blindness [262]. During this survey, participants were assessed for presence of TT and CO, providing a unique opportunity to describe the distribution of later stages of trachoma in relation to underlying risk factors in Nigerian adults.

This chapter uses geostatistical BHMs to quantify the relationship between climatic factors and trachomatous trichiasis or corneal opacity (TT/CO) amongst adults in Nigeria, while accounting for the effects of risk factors at other levels and any residual spatial correlation.

## 4.3 Methods

### 4.3.1 Overview

Available data included field collected data at the individual level and remotely sensed or interpolated environmental variables at the cluster level. An exploratory principal components analysis was conducted on all climatic variables with a correlation coefficient ≥0.70, in order to explore covariance and variance between factors and ultimately inform dataset reduction and model building.

All field collected data were used with a reduced set of environmental covariates to build hierarchical multivariate regression models for the presence or absence of TT or CO (Table 4.1). Model-building took a spatially explicit approach and evaluated the addition of geostatistical random effects to account for any residual spatially-structured clustering.

### 4.3.2 Data

*National Blindness Survey*

Data were collected over a 30-month period from January 2005 to July 2007. The framework of this study has been fully described elsewhere [263]. Briefly, a multistage stratified cluster random sampling strategy with probability proportional to size was employed to generate a nationally representative sample of adults aged 40 years and above. A total of 50 adults were enumerated in each of 305 clusters, using a random walk procedure from the centre. Visual acuity (VA) was assessed and all participants had a basic eye examination by a qualified Nigerian ophthalmologist. The presence or absence of TT and CO were recorded based on diagnoses using the WHO simplified grading scheme [26]. In addition, a questionnaire was administered to collect demographic and socioeconomic indicators for each participant.

The majority (80%) of clusters had a specific longitude and latitude recorded by GPS during the survey. The remaining 20% of the clusters were geolocated to a specific location using a variety of electronic gazetters (7%) or the centroid of the corresponding LGA (13%). One cluster could not be geolocated to a unique location and was therefore excluded from the subsequent analysis.

*Environmental and Climatic Data*

Environmental variables were selected based on their potential relevance to active transmission, through water availability or the physiology and behaviour of *M sorbens*, or progression of disease. Gridded data were obtained from a variety of sources, fully detailed in Appendix 4.1. These variables included interpolated or satellite data on annual climate trends (mean annual precipitation, land surface temperature, mean annual temperature, annual aridity index and potential evapo-transpiration (PET)), enhanced vegetation index (EVI, sometimes used as a surrogate for rainfall) and extreme or potentially limiting climatic factors (maximum temperature in the warmest month, precipitation of driest month). Long-term averages of these indices were considered appropriate as later stages of trachoma represent the cumulative effects of repeated episodes of active disease over 20-30 years. Other environmental factors included altitude, urbanisation category, landcover type, population density, distance to nearest waterbody and livestock density. Information on gridded environmental and climatic variables was extracted for each point location in R version 2,10,1.

### 4.3.3 Data categorization

*Field Collected data*

Age was classified into ten year age bands, based on the nonlinear relationship observed with TT/CO and to minimise the effect of reporting biases. Occupation, literacy, water source and latrine type are all characteristics that capture various dimensions of an individual's socioeconomic status (SES) [264] and potentially influence transmission of *Chlamydia trachomatis*, through overcrowding, water availability and usage, waste disposal and hygienic behaviours. Occupational category was recoded to distinguish professionals, semi-skilled workers and unemployed. Literacy was kept as three categories: literate, semi-literate and illiterate. Presence of a latrine has been associated with lower density of *M sorbens* and fly-eye contact [97,143,144,147] and latrine type was categorized as flush toilet, pit latrine or bush for this analysis. Water source was recoded as a binary variable in two ways: 1) to reflect an individuals' access to an improved water source, using definitions provided by the Joint Monitoring Programme for Water Supply and Sanitation [265], and 2) to reflect distance and availability of water by categorising water sources located within the household or yard separately from wells, boreholes, bought and surface water.

*Environmental data*

Global land cover was recoded in this analysis to distinguish savannah and grassland areas, which have previously been associated with a higher risk of trachoma in Nigeria and South Sudan [158,249]. Categorical variables were generated for each environmental variable. Variables were first classified into quartiles and the relationship with the cluster-level prevalence of TT/CO observed using box-plots and scatter plots. Where there was a clear pattern in the risk of TT/CO across the factor values, variables were reclassified accordingly, otherwise categories were based off of quartiles. All continuous

environmental and climatic data were standardised to improve convergence of the models.

### 4.3.4 Exploratory principal components analysis

As noted in similar analyses, multicollinearity between environmental variables commonly presents a challenge in model-building [153]. A principal components analysis (PCA), presented in Appendix 4.2, was used to explore the underlying structure of climatic variables, in terms of variance and covariance, and inform reduction of the dimensionality of the dataset for subsequent model building strategies [266]. Principal components were not used directly in the model, as they are less interpretable. However, collinear pairs of climatic variables from each grouping identified in the PCA were added through sequential regression, which aims to create a new explanatory variable by removing common variation from variables deemed to be less important [267,268]. This approach involves determining a sequence of importance for the explanatory variables, which in this case was constructed from the literature and PCA. As the literature provides the strongest evidence base for an association between precipitation and trachoma, this was considered the principal climatic factor in the regression [152,158,161].

### 4.3.4 Modelling

*Model building*

Initially a non-spatial, frequentist approach was used to select candidate variables for Bayesian spatial models, using binomial logistic regression models with a cluster-level random effect. Univariate analyses of each field-collected variable and the reduced set of environmental variables were conducted to identify initial covariates associated with TT/CO and bivariate analysis used to explore relationships with potential confounders or

correlated variables. Univariate models were fitted with continuous and categorical variables in turn, and the variable with the lowest Akaike information criterion (AIC) retained for the modelling process. If included categorically, a model including the categorical variable was compared to one fitted with a quadratic term in addition to the continuous variable.

Initial covariate selection used a forward stepwise procedure for each of the two levels (individual and cluster) in order to develop a multivariate multilevel model, keeping variables with a p-value of 0.1 or less. As explained above, precipitation was the first climatic variable to be added into the model. After accounting for the common variation captured by this variable, collinear climatic variables were regressed against it and residuals included as new variables that are conditional on precipitation. Non-linear associations between environmental covariates and the outcome were explored by adding a squared term and assessing model fit. As the majority (79%) of households had only one (39%) or two (40%) individuals included in this study, resulting model instability led to the exclusion of any household level random effect.

*Bayesian models*

Final equivalent Bayesian models were then developed, incorporating a geostatistical random effect. Models took the form:

$$Y_{ij} \sim Bernoulli(p_{ij})$$

$$logit(p_{ij}) = \alpha + \sum_{g=1}^{n} \beta_g \times x_{ij} + \sum_{h=1}^{n} \beta_h \times x_j + v_j + u_j$$

Where $Y_{ij}$ is the infection status of individual $i$ in cluster $j$, $p_{ij}$ is the probability of a positive response, α is the intercept, $\sum_{g=1}^{n} \beta_g \times x_{ij}$ is a vector of $g$ independent variables at

the individual level measured in the field multiplied by their coefficient $\beta_g$, $\sum_{h=1}^{n} \beta_h \times x_j$ is a vector of $h$ independent variables at the cluster level multiplied by their coefficient $\beta_h$, $v_j$ is a non-spatial random effect (NSRE) and $u_j$ is a spatial random effect (SRE) at the cluster level. Non-informative priors were specified for the intercept (uniform prior with bounds $-\infty, \infty$) and the coefficients (normal prior with mean=0 and precision, the inverse of variance = $1 \times 10^{-6}$). NSREs were incorporated into all models, in order to allow residual variation to have uncorrelated and correlated components. The SRE models any residual correlation that is spatially structured using an isotropic, stationary exponential decay function: $f(d_{ab}; \phi) = \exp[-(\phi d_{ab})]$ where $d_{ab}$ is the straight-line distance between pairs of points $a$ and $b$, and $\phi$ is the rate of decline of spatial correlation. The NSRE had a non-informative priors imposed on its variance (uniform distribution with delimiting values of 0.01 and 100).

Non-spatial model residuals were used to construct semi-variograms, which are introduced in Appendix 4.3. Semi-variograms were visually inspected for spatial structure and non-stationarity, and spatially structured correlation was incorporated by inclusion of location-specific geostatistical random effects in the northern and southern states. A burn in of 20,000 iterations was allowed, followed by 10,000 iterations where values for monitored variables were stored and thinned by 10. Diagnostic tests for convergence of the monitored variables were undertaken, including visual examination of history and density plots. The runs were also assessed for evidence of autocorrelation. Model performance was assessed by comparing deviance information criteria (DIC).

Residuals were checked for normalcy and a sensitivity analyses was run, excluding sites which were geolocated to the local government area (LGA) centre and might introduce error. All analyses were run from R in WinBUGS 1.4 (MRC Biostatistics Unit, Cambridge and Imperial College London, UK), using the package 'R2WinBUGS'.

### 4.3.5 Variance components

There is no simple way to measure variance partition coefficient (VPC) for discrete response multilevel models, as the variance at the two levels are measured on different scales and dependent on individual level predictor variables. We used a simulation approach implemented in R 2,10,1 to estimate the VPC introduced by Goldstein et al. [257], which approximates the variance at each level from a large number of simulations based on the variance in the second-level random effect, beta values from the non-spatial model and average values for each coefficient.

## 4.4 Results

*Raw data*

Complete geolocated survey data were available for 304 clusters, corresponding to 8,621 households with 13,543 individuals aged 40 years and above with information on the absence/presence of TT and CO. Overall, 198 (Adjusted prevalence: 1.45%) individuals were diagnosed with either TT or CO, and only two individuals had clinical signs of CO without concurrent TT. Figure 4.1 presents the distribution of TT/CO among adults aged 40 years and above within clusters (ranging from 0 to 28.9%) and highlights the greater burden of trachoma in the northern areas of Nigeria.

Summary characteristics of the study population are described in Table 4.2 and reflect socioeconomic trends across the country. Overall, only 10% of participants used a flush toilet although 64% had access to a pit latrine, and over half (56%) of the participants could not read. Northern zones had higher illiteracy (62.5%) and unemployment (19.6%) compared to the south (49.3% and 12.6%). Although fewer people had access to a flush toilet in the northern zones (4.4%) than southern zones (17.5%), open defecation was also reported less in northern zones (20.9%) compared to the south (31.3%).

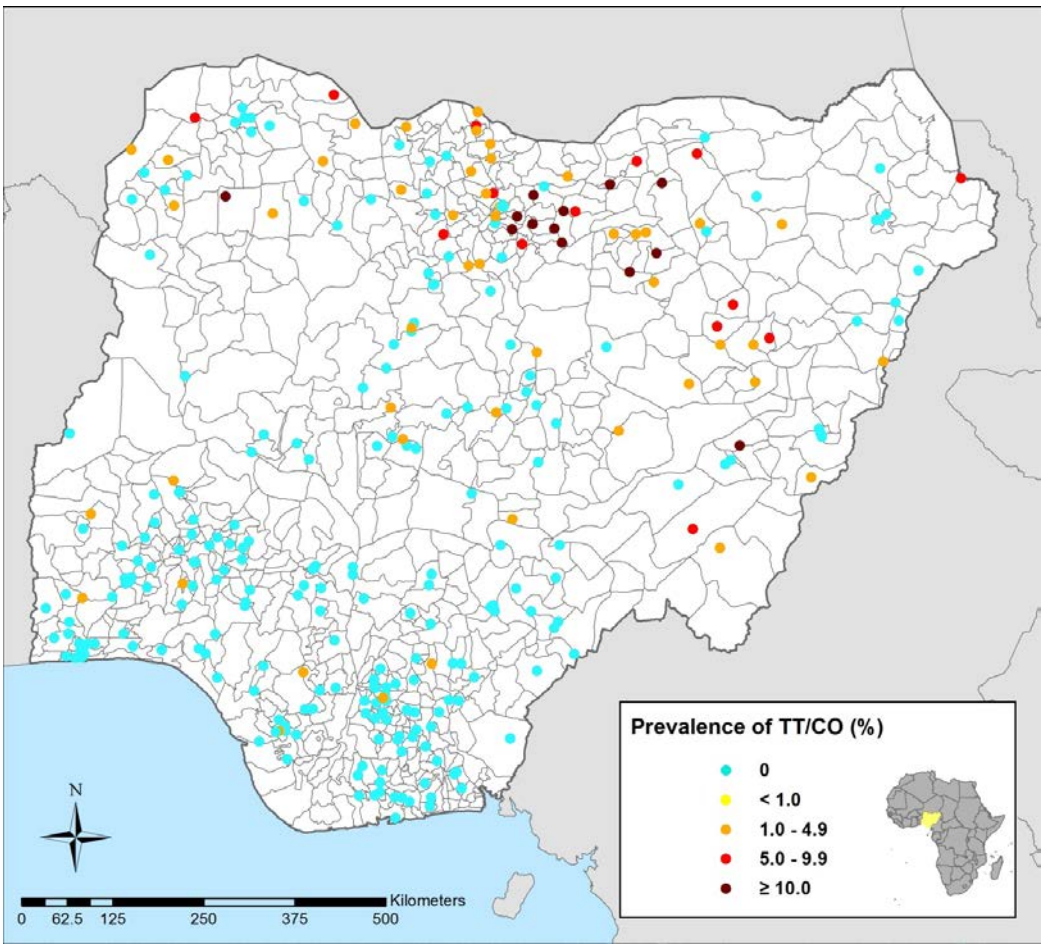



Figure 4.1 Prevalence of trichiasis (TT) or corneal opacity (CO) in adults over 40 years in Nigeria from 304 clusters, 2005-2007. Higher prevalence clusters are predominantly in northern areas of Nigeria.

Table 4.1 Summary statistics of the reduced set of climatic and environmental covariates included in model building

| Variable | Median (range) [a] | SD |
|---|---|---|
| **Climate** | | |
| Land surface temperature (LST) (°C) | 31.7 (22.5, 42.8) | 5.0 |
| Mean annual temperature (°C) | 26.4 (21.8, 28.7) | 1.0 |
| Mean annual precipitation (mm) | 1284.0 (407.0, 3833) | 639.2 |
| **Environmental** | | |
| Altitude (meters) | 270.5 (4.0, 1287.0) | 226.1 |
| Savannah or grasslands[b] | 17.0 % | - |
| Ruminant density (animals per 5km cell) | 68.1 (0, 1051.4) | 144.8 |
| Cost-distance to road network | 1387.5 (0, 22842.7) | 2983.3 |
| Urban classification | 27.3% urban | - |
| Population density | 285.0 (8.0, 27982.0) | 2570.4 |
| Enhanced vegetation index (EVI)[c] | 1218.0 (88.1, 1500.8) | 270.8 |

SE: standard deviation; °C: degrees Celsius; mm: millimetres; km: kilometer

[a] Proportion of sites for binary variables (Savannah/Grasslands and Urban classification);

[b] Reclassified from global land cover; [c] Fourier transformed data

Table 4.2 Descriptive statistics of the 13,543 individuals included in the 2007 National Blindness and Visual Impairment Survey

| **Statistic** | | **Number (%)** |
|---|---|---|
| Total number individuals | | 13,543 |
| Gender | Female | 7,317 (54.0) |
| Age group: | 40-50 years | 5,821 (43.0) |
| | 50-60 years | 3,371 (24.9) |
| | 60-70 years | 2,600 (19.2) |
| | 70-80 years | 1,312 (9.7) |
| | >80 years | 439 (3.2) |
| Literacy: | Easily | 2,925 (21.6) |
| | With difficulty | 2,983 (22.0) |
| | Not at all | 7,635 (56.4) |
| Occupation: | Professional | 1,317 (9.7) |
| | Semi Skilled | 10,013 (73.9) |
| | Unemployed | 2,213 (16.3) |
| Latrine type: | Flush | 1,415 (10.4) |
| | Pit Latrine | 8,648 (63.9) |
| | Bush | 3,480 (25.7) |
| Improved water source: | Unimproved | 3,802 (28.1) |
| | Improved | 9,741 (71.9) |
| Proximate water source: | Wells/boreholes/surface water | 8,188 (60.5) |
| | Within street or household | 5,355 (39.5) |
| TT and/or CO | | 198 (1.45) |

TT: trachomatous trichiasis; CO: trachomatous corneal opacity

*Exploratory analysis*

All field collected variables and environmental covariates were strongly associated with the presence of TT/CO in univariate logistic regression models, with the exception of mean annual temperature, as summarised in Table 4.3. Correlation was observed between a number of variables related to SES, including occupation, water availability, literacy and latrine type. Literacy was associated both with occupation ($p<0.0001$) and gender ($p<0.0001$). Women with a lower literacy status had a higher risk of TT/CO, partly accounting for the increased risk observed in illiterate individuals. A geographic north-south trend in risk of trichiasis was apparent across the country, and the unbounded semi-variogram for the raw TT/CO prevalence supported the presence of spatial autocorrelation in the distribution of disease (Figure 4.2A).

The results from the PCA, fully detailed in Appendix 4.2, identified five key groupings of variability in climatic covariates, from each of which a single variable was retained. Mean annual precipitation and land surface temperature were retained from the two contrasting

groups from the first component. Mean annual temperature and altitude (identified in the PCA as a second collinear pair contributing to climatic variation) and EVI were also retained for further analyses with all other uncorrelated environmental indices. Summary statistics for these variables are presented in Table 4.1. During model building, the residual variation in EVI was initially significant after accounting for collinear effects of precipitation and LST and the residual effect of LST, but dropped out after accounting for urban classification.

Table 4.3 Univariate associations with variables in 304 clusters of 13,543 individuals aged ≥40 in Nigeria, 2005-2007

| Variables | | Group | OR (p-value) |
|---|---|---|---|
| ***Field-collected variables*** | **Individual level** | | |
| | Age group: | 40-50 years | 1 |
| | | 50-60 years | 1.68 (0.02) |
| | | 60-70 years | 4.00 (<0.0001) |
| | | 70-80 years | 4.65 (<0.0001) |
| | | 80+ years | 5.30 (<0.0001) |
| | Gender | Female | 2.61 (<0.0001) |
| | Literacy | Easily | 1 |
| | | Difficult | 4.11 (<0.0001) |
| | | Not at all | 2.06 (0.03) |
| | Improved water source | Unimproved | 1 |
| | | Improved | 1.40 (0.421) |
| | Proximate water source | Village or further | 1 |
| | | Within street or household | 0.95 (0.86) |
| | Occupation | Professional | 1 |
| | | Semi Skilled | 11.44 (0.03) |
| | | Unemployed | 35.55 (0.002) |
| | **Cluster level** | | |
| ***Climatic covariates*** | Mean annual precipitation (mm) | | 0.997 (<0.0001) |
| | Monthly Average land surface temperature (LST) | | 1.358 (<0.0001) |
| | Mean annual temperature (°C) | | 1.01 (0.59) |
| ***Environmental covariates*** | Altitude (meters) | <200 | 1 |
| | | 200 – 499 | 14.41 (<0.0001) |
| | | 500 + | 5.64 (<0.0001) |
| | Enhanced vegetation index (EVI) | ≥ 0.35 | 1 |
| | | 0.25-0.34 | 6.66 (0.001) |
| | | 0.15-0.24 | 26.00 (<0.0001) |
| | | < 0.15 | 4.53 (0.01) |
| | Land cover type | Other | 1 |
| | | Savannah/ Grasslands | 2.70 (<0.0001) |
| | Urban classification | Rural | 1 |
| | | Urban | 0.27 (<0.0001) |
| | Distance to surface water (km) | | 1.18 (<0.0001) |
| | Population density (per 5km cell) | | 0.99 (<0.0001) |
| | Ruminant density (per 5km cell) | | 1.003 (0.02) |

All associations adjusted for clustering within villages

*Geostatistical risk model*

Model results are reported in Table 4.4 and retained both individual and cluster-level covariates. The final model reported is Model 4.2, which is non-spatial and includes age, gender, and occupation as well as mean annual precipitation, residual variation in LST, mean annual temperature and urban classification. Risk of TT/CO increased with age and was higher in women than men (OR 2.46, 95% BCI 1.82 – 3.39). Socioeconomic status, as measured by occupation, was also associated with an increased risk of trichiasis. Despite wide confidence intervals, there was evidence that individuals employed in a professional capacity had the lowest risk of trichiasis while unemployed were at highest risk (OR 16.71, 95% BCI 3.23 – 556.1).

Increased precipitation was associated with a lower risk of TT/CO in Nigeria (OR 0.17, 95% BCI 0.06 – 0.33), and higher residual LST was uniquely associated with an increased risk of TT/CO (OR 2.95 95% BCI 1.36 – 6.85) additional to any contribution made through variation that is shared with precipitation. Although not identified as a risk factor in the univariate analyses, increased mean annual air temperature was associated with lower risk of TT/CO after controlling for the effects of other environmental factors. This variable was kept in the model based on the lower DIC. Finally, the odds of TT/CO were lower in urban areas (OR 0.27 95% BCI 0.13 – 0.52), after controlling for individual level risk factors. Approximation of the VPC using a simulation approach suggested that 14% of the total variation (based on a null model) was attributed to the cluster level. After inclusion of both individual and cluster-level risk, 0.7% of the overall residual variation was at the higher level.

Although the results from the non-spatial model are reported here, there was evidence of large scale spatial trends as well as local clustering of TT/CO risk in Nigeria. The semi-variogram of the Pearson's residuals from Model 1 indicated that, compared to the null model, the addition of covariates decreased the proportion of variation that was spatially structured and controlled for large-scale trends (Figure 4.2). This spatial structure varied

between regions (non-stationarity), with a higher proportion of residual variation in the north showing spatial structure (Figure 4.3). Graphs and maps of the residuals from the non-spatial Model 2 suggested that residual variation was localised in a large cluster of higher risk in the north (Figure 4.4). These residuals are presumably due to presence of unknown, spatially structured risk factors that are not included in the model or spatially varying relationships with identified risk factors. Inclusion of separate random effects for the north and south had the effect of reducing overall residual error in the model, as indicated by the reduction in the non-spatial random effect and narrower confidence intervals (Model 3: Table 4.4 and Figure 4.4). However, addition of these terms also reduced observed associations with residual LST and mean annual air temperature, and widened their confidence intervals. This finding suggests that while these environmental factors may be associated with the distribution of risk in the north, they do not explain all observed clustering and are made redundant by inclusion of a spatial random effect. The range of spatial autocorrelation can be calculated by $3/\varphi$ and is thus 3.26 decimal degrees (approximately 365 km) in the north. Residual variation in the south was more likely to be aspatial and due to individual level factors.

Table 4.4 Random-effects models for trachomatous trichiasis or corneal opacity in adults over 40 years in Nigeria

| Variable | Null Model 1 [a] OR (95% BCI) | Null Model 2 [b] OR (95% BCI) | Model 1[a] OR (95% BCI) | Model 2 [a] OR (95% BCI) | Model 3[b] OR (95% BCI) |
|---|---|---|---|---|---|
| Age: 40-49 years | | | - | - | - |
| 50-59 years | | | 1.78 (1.27, 2.52)* | 1.86 (1.33, 2.71)* | 1.90 (1.30, 2.81)* |
| 60-69 years | | | 3.99 (2.87, 5.85)* | 4.48 (3.18,6.34)* | 4.54 (3.17, 6.55)* |
| 70-79 years | | | 4.30 (2.68, 6.61)* | 4.78 (3.13, 7.43)* | 5.17 (3.25, 8.05)* |
| 80 plus years | | | 3.70 (1.76,7.36)* | 5.24 (2.51, 10.67)* | 5.62 (2.63, 10.98)* |
| Gender: Male | | | - | - | - |
| Female | | | 2.23 (1.61, 3.01)* | 2.46 (1.82, 3.39)* | 2.55 (1.89, 3.51)* |
| Occupation: professionals | | | - | - | - |
| semi/skilled | | | 9.80 (2.09, 161.27)* | 10.83 (2.18, 359.4)* | 15.39 (3.18, 116.1)* |
| unemployed | | | 17.84 (3.89, 286.20)* | 16.71(3.23, 556.1)* | 21.63 (4.34, 173.2)* |
| Mean annual precipitation | | | | 0.17 (0.10, 0.26)* | 0.21 (0.11, 0.38)* |
| Residual LST | | | | 2.95 (1.36, 6.85)* | 1.91 (0.79, 5.08) |
| Squared term | | | | 0.21 (0.07,0.58)* | 0.19 (0.05, 0.63)* |
| Mean annual temperature | | | | 0.89 (0.69, 1.14) | 0.91 (0.62, 1.37) |
| Squared term | | | | 0.75 (0.62, 0.87)* | 0.83 (0.67, 1.02) |
| Urban classification | | | | 0.27 (0.13, 0.52)* | 0.34 (0.17, 0.61)* |
| | Beta coefficient (95% BCI) | Beta coefficient (95% BCI) | Beta coefficient (95% BCI) | Beta coefficient (95% BCI) | Beta coefficient (95% CI) |
| Intercept | -5.86 (-6.47, -5.36) | -5.43 (-6.58, -4.35) | -9.71(-12.54, -7.97) | -9.19 (-12.55, -7.47) | -9.90 (-12.1, -8.29) |
| $\sigma^2$ | 4.04 (2.56, 6.20) | 0.52 (0.23, 1.04) | 4.66 (3.35, 6.55) | 1.53 (0.98, 2.41) | 0.006 (0.00, 0.41) |
| Spatial southern $\varphi$ [range in km] | | 0.22 (0.05, 0.55) [1527] | | | 59.13 (22.2, 94.19) [6] |
| Spatial southern $\sigma^2$ | | 5.72 (1.13, 19.41) | | | 1.27 (0.24, 4.18) |
| Spatial northern $\varphi$ [range in km] | | | | | 0.92 (0.23, 2.42) [365] |
| Spatial northern $\sigma^2$ | | | | | 2.55 (1.18, 8.00) |
| DIC | 1658 | 1610 | 1533 | 1490 | 1483 |

OR: odds ratio; BCI: Bayesian credible interval; DIC: deviance information criterion (smaller DICs indicate better model fit); $\varphi$ = rate of decay of spatial correlation; $\sigma^2$ = variance of random effect [a] Non-spatial random effect only; [b] Including separate spatial random effect for the north and the south of Nigeria

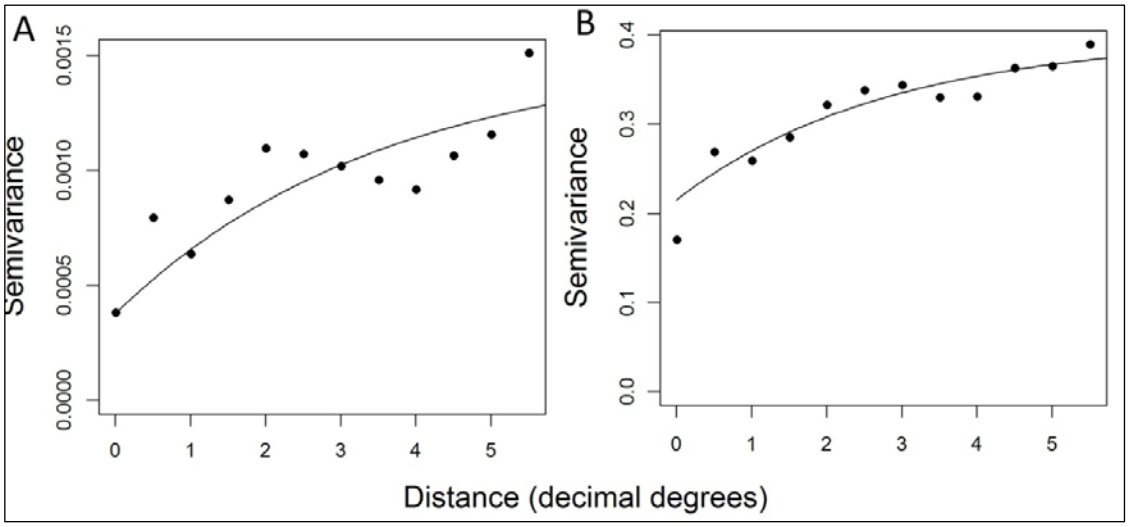



Figure 4.2 Semi-variograms related to risk mapping models for presence of trachomatous trichiasis (TT) or corneal opacity (CO) in adults over 40 years, Nigeria, 2005-2007. There is evidence of spatial structure and the suggestion of large scale trends in risk of TT/CO both in the raw data (A) and graphs of Pearson's residuals from model 2 including individual and cluster-level covariates (B).

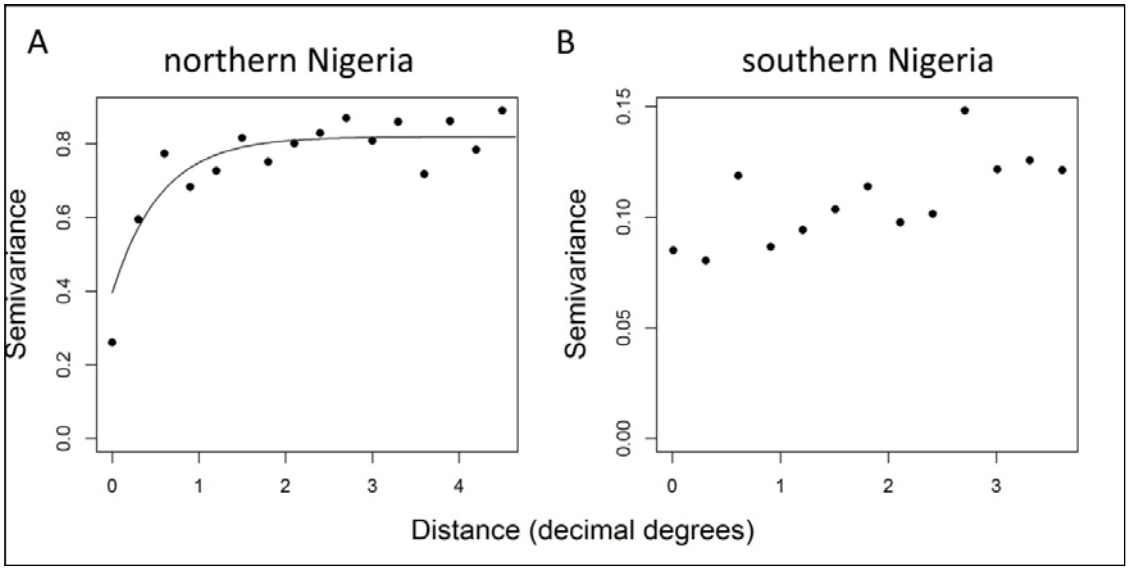



Figure 4.3 Semi-variograms model the spatial structure of TT/CO present in residuals from nonspatial Model 2 separately in northern (A) and southern (B) regions of Nigeria. While there is no evidence of spatial structure in the residuals from southern Nigeria, the semi-variance (or difference) in risk is observed to increase with distance in northern Nigeria. Of key interest is the ratio of sill to nugget variance, which provides information on how spatially structured the variance in prevalence is. In northern Nigeria the ratio is 2.07, suggesting that just over half of residual variance is spatially structured in the North. This structure may be due to dependency on unknown risk factors which are locally clustered in these areas or non-stationarity in relationships between observed risk factors and disease.

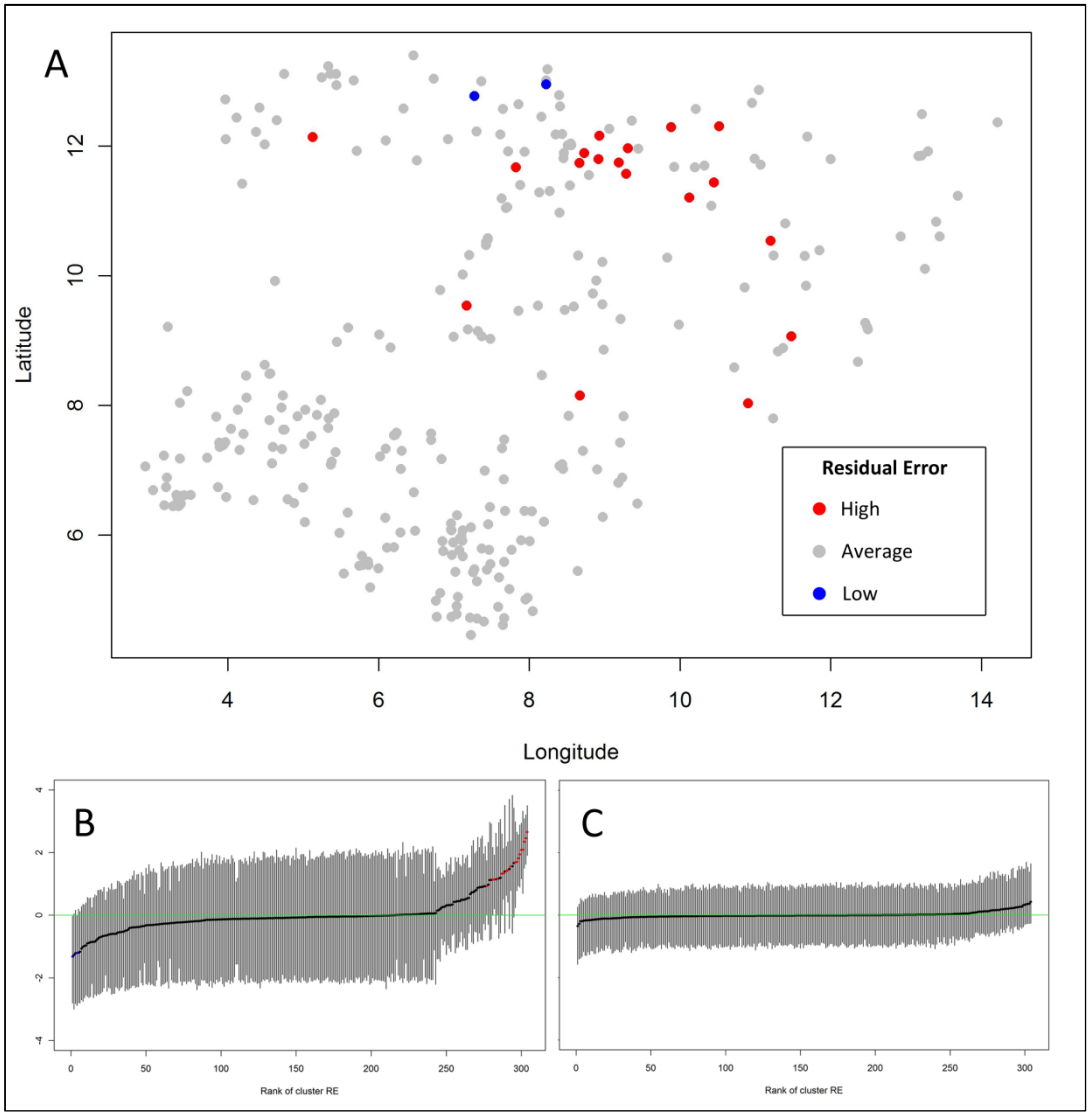



Figure 4.4 Residual error from non-spatial Model 2 (A, B) and spatial Model 3 (C). Mapped residuals with 95% Bayesian Credible Interval (BCI) above zero indicate that areas of high residual risk are localised in northern Nigeria, with a large cluster present encompassing southern Jigawa, eastern Kano and northern Bauchi states.

## 4.5 Discussion

The present study provides evidence that both individual-level risk factors and broader climatic conditions are associated with later stages of trachoma in adults over the age of 40 years in Nigeria, using uniquely detailed national survey data. The hierarchical approach used in this analysis has the advantage of incorporating risk factors at multiple levels and explicitly modelling residual spatial correlation in TT/CO that could affect the standard errors of estimates of association. A number of well-established individual risk factors for trichiasis were identified that included age, gender and occupation, as well as large-scale climatic and environmental factors (precipitation, residual LST, temperature and urban classification) that explained further variation in risk across the country. After adjusting for these factors, there is evidence that additional, unknown spatial risk factors in northern Nigeria underlie observed disease clustering.

Individual-level factors found to be associated with trichiasis are consistent with our general understanding of trachoma epidemiology. These associations replicate those previously found in a number of studies in Nigeria [269], other countries in sub-Saharan Africa [157,228,250] and trachoma endemic areas worldwide [210,270,271]. The risk of TT increases with age due to cumulative scarring caused by repeated infection over an individual's life, while the higher risk in females is commonly attributed to close contact with children and greater exposure to infection with the causative agent [5,113]. Occupation is a characteristic that captures various dimensions of an individual's SES and may be linked to underlying risk factors for infection, including hygienic behaviours, use of water, waste disposal, overcrowding or other conditions that encourage the proliferation of flies or increased transmission through contact and fomites. A lower score on various socioeconomic measures has also been associated with a greater risk of trachoma in previous studies, including occupation [59], illiteracy or lack of formal schooling [271,272] and living standards measures, but not uniformly across all settings [166]. Variation in relevant socioeconomic measures between settings may reflect differences in

underlying transmission dynamics, equity in access to treatment and surgical interventions, as well as reliability of the measures themselves. After accounting for these risk factors, living in urban areas remained associated with a lower risk of trichiasis. This finding supports anecdotal evidence underlying current trachoma survey strategies that preferentially select rural clusters, and may reflect reduced access to health services or increased contact with livestock and flies associated with rural lifestyles.

After controlling for individual-level risk factors, lower precipitation, higher land surface temperatures and lower mean annual temperatures were associated with a higher risk of TT in Nigeria. On this scale, climatic factors may influence transmission dynamics through hygienic behaviours related to perceived water availability, actual water availability or as determinants of fly physiology and behaviour [131,132,135,160,178,273]. Shared variation in precipitation and LST accounted for the most climatic variation (Appendix 4.2), and might be interpreted as variation common to different measures of climatic water availability. Higher residual LST, after removing collinear variation with precipitation, was independently associated with a further increased risk of TT/CO. These findings are consistent with previous analyses associating a higher risk of active trachoma with higher aridity and lower rainfall [154-158]. The higher risk of TT/CO associated with lower air temperatures (or higher altitudes) seems counter-intuitive, however this association has been reported in previous studies with limited control for potentially confounding variables [153,250]. Lower meterological temperatures are hypothesized to have an biological effect on the life span of the punitive fly vector, *M sorbens*, which has been shown to live from 12 days at 32° C to 35 days at 24°C [273]. It is expected that TT/CO in adults who are ≥40 years old at the time of the survey mainly reflects exposure to factors influencing transmission 30-40 years ago (assuming little population movement across clusters/climatic gradients). However, it is also possible that certain climatic factors may also influence the development of trachomatous scarring and hence TT/CO. Ongoing active disease and eye irritants like ocular dryness may be associated with drier climatic conditions and contribute to chronic conjunctival inflammation. This in turn has

been associated with a higher risk of TT and faster progression to later disease stages [18,19,154,247].

Despite strong links between water availability and transmission of trachoma, there are a number of potential reasons why household water source was not identified as a risk factor in this analysis. First, domestic water consumption and, importantly, its allocation for hygienic purposes will mediate any relationship between water availability and trachoma [132]. This is difficult to measure and while distance to water [116,126,178] and type of source [274] have been associated with trachoma in some studies, they are at best proxy measures of household and individual water use. It is likely that our classifications of water source were not able to capture these directly relevant measures. In addition, a study on water use patterns in Tanzania highlighted the importance of perceived water availability and its impact on water usage, rather than availability itself [131]. It is possible that perceptions around water availability are partly driven by climatic experiences and thus may influence subsequent behaviours, including allocation within the household. Second, this survey was done over 30 months and limited evidence suggests that the water source reported as "main" may vary seasonally in Nigeria [275]. Consequently, any observed relationship between distance and usage may be stronger in the dry season. Third, individual occupation as a socioeconomic measure may have captured any effect of water source, as those with higher incomes had improved water access. And finally, while water is likely to be associated with transmission, trichiasis prevalence may more strongly reflect historical transmission levels prior to any recent interventions or secular trends. In support of this hypothesis was our finding that an improved water source was associated with an increase in the unadjusted odds of disease in the driest areas, potentially reflecting targeting of water interventions to the driest areas in the last 20-30 years.

One of the strengths of this analysis lies in its explicit recognition of the hierarchical structure of the data and ability to incorporate residual spatial variation. After accounting for risk factors at the individual and cluster level, there was evidence that TT/CO was

spatially structured over a large (365 km) range in the north. This is likely to be due to a large cluster of residual risk, focused around southern Jigawa, eastern Kano and northern Bauchi states. Approximation of the VPC using a simulation approach suggested that 14% of the total variation was attributed to the cluster level. After accounting for risk factors at both levels, this was reduced to less than 1%. This suggests that risk of TT/CO is more variable within clusters than between clusters, and is consistent with the natural history of trachoma which requires repeated infections of *C trachomatous*, observed to cluster within households [106,250] and individuals [112]. In contrast, a recent study by Hagi et al. attributed nearly 40% of observed variation in active trachoma to the village level [152]. It is not clear what approach was used for this estimate, thus it may not be directly comparable to estimates from this study, but a higher proportion of variation between villages may reflect the importance of environmental factors on transmission dynamics via flies and water availability. The influence of these factors may give a relatively homogenous "spread" of active disease risk across a community, whereas clustering of TT may reflect the importance of individual-level risk factors which influence the predisposition to infection, duration of infection, or immunological response to infection over longer periods of time.

Despite the robust approach used to model these data, there are a number of limitations inherent in the data and methods. First, as anticipated, strong collinearity in both environmental and socioeconomic variables across the country placed limitations on our ability to disentangle their independent effects. Observed associations with climatic factors may reflect uncontrolled socioeconomic factors, as rural populations are likely to be dependent on agro-ecological conditions for crop and livestock productivity. Second, this survey was cross-sectional and TT is a condition that represents the cumulative effect of many infections over time. Thus potential decadal climate variability, migration during an individual's adult life or variation in other risk factors for TF between the period of exposure and time of survey limits any inferences of causality. Expected associations may be masked, or even reversed in some cases, where access to the SAFE strategy (including

surgery and environmental improvement) has been implemented in high transmission areas. No information was available on coverage of the SAFE strategy in the years leading up to the survey, but anecdotal evidence suggests that the SAFE strategy was not widely implemented within Nigeria at this time. And finally, not all points were able to be geolocated and any errors in geolocation could introduce misclassification at the cluster-level.

For the first time, we have quantified associations between environmental factors and risk of TT/CO in Nigeria while accounting for the effects of individual-level risk factors and residual spatial structure. While the results indicate that individual-level factors are an important source of variation, individuals living in drier and rural areas of Nigeria were at greater risk of chronic disease stages. This supports anecdotal evidence associating limited water availability with trachoma although other mechanisms may also be important, such as the effect of temperature on the abundance, breeding potential and activity of *M sorbens* [276]. Findings from this study may help to more reliably extrapolate trichiasis data within countries and regions and refine estimates of the burden of disease, although further work is required to investigate associations at larger scales.

As discussed in this chapter, one pathway through which climatic and environmental factors may influence the risk of TT is through their impact on transmission pathways of *C trachomatis* and active disease. A more proximate requirement for national programmes is to identify areas which should be prioritised for disease mapping or, even more importantly, can be reliably excluded from these activities. The next chapter extends the use of multilevel modelling and spatial analyses to investigate variation in the risk of active trachoma in Kenya and evaluate the potential for developing predictive risk maps.

# Chapter 5: Analysis of spatial patterns and risk factors for active trachoma in Kenya

## 5.1 Overview

The previous chapter provided evidence that the distribution of trichiasis within Nigeria is associated with well-established individual risk factors, as well as larger-scale factors that vary between communities. Geostatistical models such as these could potentially inform predictive models to refine estimates of the burden of disease at larger scales. Hierarchical models of active trachoma in Mali have found that cluster-level environmental factors explained a relatively high proportion (36.7%) of the overall variation in risk [152]. Quantification of associations at this level and characterisation of heterogeneity in risk will not only add to our epidemiological understanding of patterns of disease, but might provide a basis for targeting surveys or surveillance and optimising their design.

This chapter will use disaggregated, cluster-level data on the prevalence of active trachoma in Kenya to i) explore whether socioeconomic and environmental/climatic factors are associated with patterns of disease in endemic areas, ii) to investigate spatial heterogeneity of active trachoma between and within districts and iii) evaluate the potential for developing a spatial risk map of active trachoma in Kenya. This approach may offer insight into the spatial epidemiology of trachoma in Kenya, with implications for current mapping strategies and future surveillance of disease.

## 5.2 Introduction

Epidemiological investigations of trachoma have associated numerous socioeconomic and environmental factors with risk of active disease and infection, mainly linking exposures to hypothetical transmission pathways through water availability, crowding and fly density [54]. While the majority of studies have mainly focused on small scale (individual or household-level) risk factors, a few studies have characterised heterogeneity in the distribution of disease over larger scales [152,157,158]. As reviewed in Chapter 1, patterns of disease arise due to underlying exposures to *Chlamydia trachomatis*, which are in turn mediated by influences acting at different levels, including individual (e.g. genetic, behavioural) and household/familial factors (e.g. socio-economic and behavioural) through to large scale environmental factors.

Trachoma prevalence data are typically presented at the district level, which does not allow appreciation of the variation within districts and spatial patterns of risk that may cross administrative boundaries. While spatial heterogeneity in disease risk is of interest in informing survey design, a better understanding of risk factors associated with the distribution of disease may help to target future surveys or surveillance. In response to recent elimination targets requiring the prevalence of active disease to be below 5% in all subdistricts [3], national control programmes may be prompted to reconsider surveying areas which were initially classified as non-endemic. A more systematic approach making use of disaggregated data may help in identifying areas at high risk and target future surveys. Geographic data on a number of potential risk factors for trachoma are increasingly available at sub-national scales, including data on poverty and socioeconomic risk factors from census surveys, multiple-indicator cluster surveys (MICS), and demographic health surveys (DHS), as well as remotely sensed environmental and climatic data. As well as defining large-scale geographical trends in disease prevalence, higher resolution spatial data for both disease prevalence and underlying risk factors may also aid in the identification of hotspots or areas at higher risk of disease recrudescence. As

introduced in Chapter 1, predictive disease risk mapping is a well-established tool for a broad range of tropical and non-tropical infectious diseases [170,187,277] and has previously been used to predict the risk of TF in South Sudan based on strong associations with water availability and land cover [158]. Although both risk factors and the reliability of predictive maps are likely to vary by country, these methods could offer a more systematic approach to prioritise baseline surveys or potentially target future surveillance strategies.

The aims of this chapter are to identify potential risk factors for active trachoma within surveyed districts in Kenya, describe spatial heterogeneity in its distribution and evaluate the potential use of predictive risk maps in this context. The data, comprised of all of cluster level prevalence data from surveys conducted by the Kenya National Trachoma Control Programme, are representative of areas suspected to be endemic within Kenya.

## 5.3 Methods

### 5.3.1 Trachoma and its control in Kenya

The Kenya National Trachoma Control Programme (KNTCP) was officially launched in 2007 alongside the Kenya National Plan for Elimination of Trachoma by 2015 (KNPET 2008-2015). Surveys conducted by the KNTCP have been targeted to areas suspected to be at highest risk of trachoma, based on the evidence and assumptions described in the following paragraph (Figure 5.1; Table 5.1). The first surveys were carried out in 2004, in six districts suspected to be endemic for trachoma in the Rift Valley Province [278], followed by surveys in Turkana district and Eastern Province. As of 2012 baseline mapping using population-based prevalence surveys (PBPS) or trachoma rapid assessments (TRA) had been completed in all 17 districts classified as “endemic”.

Trachoma is commonly associated with rural pastoralists in Kenya, and targeted partly based on historical surveys which found a high prevalence of blinding trachoma in

Turkana [279] and Samburu [280], and others that informed the presumptive exclusion of areas in central Kenya and along the Eastern coast [156]. Surveyed districts lie primarily in the north-west lowlands and the south-central plains of Kenya, and comprise many (but not all) of those areas classified as arid and semi-arid lands (ASAL) which have significant pastoralist populations. ASAL areas account for 80% of the land in Kenya, tend to be rural and are generally characterised by low rainfall in arid (150-450 mm) and semi-arid (500-850 mm) districts. As might be expected, the poorest communities are found in these sparsely populated arid zones, mainly in the north, while areas of high agricultural potential in central and western Kenya are densely populated [281]. The exclusion of arid pastoralist districts within North Eastern Province and inner Coast Province from trachoma survey activities was based on the belief that it does not constitute a public health problem in these areas due to religious hygienic practices.

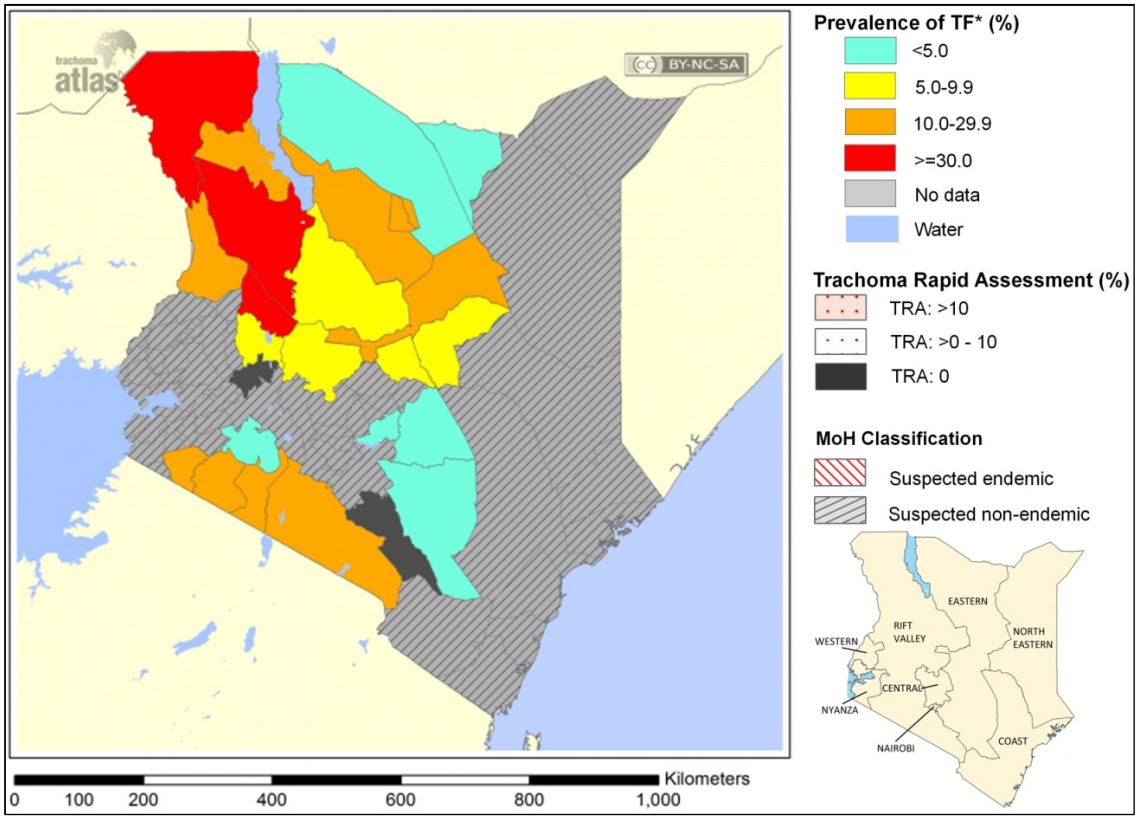



Figure 5.1 & Table 5.1 Districts classified as suspected endemic for trachoma and surveyed by the Kenyan National Trachoma Control Programme (KNTCP) using population-based prevalence surveys (PBPS) or trachoma rapid assessment (TRA) between 2004 and 2012.

| Province | District | Baseline survey[a] | ASAL Classification[b] |
|---|---|---|---|
| Rift Valley | Baringo | 2004; PBPS; district | 50-85% ASAL |
| Rift Valley | Kajiado | 2004; PBPS; district | 85-100% ASAL |
| Rift Valley | Laikipia | 2007; PBPS; district | 50-85% ASAL |
| Rift Valley | Narok | 2004; PBPS; district | 30-50% ASAL |
| Rift Valley | Samburu | 2004; PBPS; district | 85-100% ASAL |
| Rift Valley | Transmara | 2004; PBPS; district | 30-50% ASAL |
| Rift Valley | Turkana | 2010; PBPS; segmented | 100% ASAL |
| Rift Valley | West Pokot | 2004; PBPS; district | 50-85% ASAL |
| Eastern | Isiolo | 2011; PBPS; segmented | 100% ASAL |
| Eastern | Kitui | 2012; PBPS; segmented | 85-100% ASAL |
| Eastern | Koibatek[c] | 2011; TRA | 85-100% ASAL |
| Eastern | Mbeere | 2012; PBPS; segmented | 50-85% ASAL |
| Eastern | Makueni[c] | 2011; TRA | 50-85% ASAL |
| Eastern | Marsabit | 2011; PBPS; segmented | 100% ASAL |
| Eastern | Moyale | 2011; PBPS; segmented | 100% ASAL |
| Eastern | Mwingi | 2012; PBPS; segmented | 50-85% ASAL |
| Eastern | North Meru | 2004; PBPS; district | 50-85% ASAL |

[a] District surveys were based on administrative boundaries while segmented surveys divided districts into smaller geographical units based on risk assessments

[b] ASAL classification taken from the Draft National Policy for the sustainable development of arid and semi-arid lands of Kenya [282]

[c] Data are not included in this analysis

### 5.3.2 Survey Background

Surveys in Kenya have followed similar protocols, with total sample size calculated as recommended in the programme managers guide based on an expected prevalence of 10% and absolute precision of 0.04 [56]. Sublocations are used as the cluster unit and twenty are selected per evaluation unit using a systematic sampling method with probability proportional to size (PPS) [278]. All sublocations were surveyed by selecting two villages (or more) in each sublocation and, splitting the required sample size between these villages, using a compact-segment sampling method to select a sufficient number of households [278].

Since 2010, the evaluation unit has changed from the district (as determined by the 2007 administrative boundaries) to segmented "trachoma districts", of 100,000-200,000 people each. Trachoma districts correspond to divisions (sub-districts) aggregated according to scores from a pre-survey trachoma risk assessment, with the aim of creating more homogenous districts and increasing statistical precision of resulting prevalence estimates [87]. The level of risk was quantified by a summary score for each division, based on five key parameters (Table 5.2). Parameters were focused on socioeconomic and water indicators and were completed for each division by key informants, who were typically eye care or public health officers.

As of March 2013, Kenya's national re-administration has created 47 county governments based on the 1992 second level administrative divisions, and no longer uses former provinces or districts. The KNTCP classifies 12 of these 47 counties as trachoma endemic and will continue to base future evaluation units on segments defined in baseline surveys.

Table 5.2 Outline of parameters included on the risk assessment form used between 2010-2012 to aggregate divisions into "trachoma districts" according to similar scores

| Parameter | Categories | Score |
|---|---|---|
| **Previous evidence of trachoma** | Borders on an endemic district | 1 |
| | Trichiasis (TT) cases reported | 2 |
| | Trachoma Rapid Assessment | 3 |
| | Population-based Prevalence Survey | 4 |
| **Socioeconomic activity of the community** | Majority are settled urban | 1 |
| | Majority are settled agricultural | 2 |
| | Mixed nomadic herders and settled agricultural | 3 |
| | Majority are nomadic herders | 4 |
| **Water availability?** | Has piped water in most of the houses | 1 |
| | Has constant water supply: springs/rivers/dams/boreholes | 2 |
| | Dry less than 6 months in a year | 3 |
| | Dry most of the year | 4 |
| **Average number of hours most people take to fetch water (one round trip)?** | Does not fetch water; piped to every house | 1 |
| | Less than one hour | 2 |
| | One to two hours | 3 |
| | More than two hours | 4 |
| **Poverty level of communities** | Whole community is rich | 1 |
| | Majority rich; clusters of poor communities | 2 |
| | Majority poor | 3 |
| | Very poor and receiving famine relief | 4 |

Low risk: 5-10; Moderate risk: 11-15; High risk: 16-20

### 5.3.3 Conceptual framework

Selection of potential determinants of the risk of active trachoma in Kenya and subsequent model building strategies were based on a conceptual framework developed for this chapter (Figure 5.2). Inputs into this framework included risk factors on causal pathways reviewed in Chapter 1 and those considered in poverty risk frameworks [156,169,283-285]. Ultimately, the most proximal factors influencing transmission of *C trachomatous* are likely to be behavioural, as increased face and hand washing, latrine use and lower contact are expected to reduce its transmission. These measures are not available for this study; however, more distal factors will influence exposures and thus may be associated with the distribution of risk in Kenya at larger scales. Access to "Human Resources" (i.e. education, health services and water and sanitation), are hypothesised to have the most direct effect on these behaviours, but may be mediated by conditioning factors such as remoteness to these resources and social values. Access to these resources is determined

by income level and equality within a given area, which may also influence crowding and subsequent contact patterns.

Environmental and climatic conditions could potentially effect transmission by influencing hygienic behaviours proximally through water availability or, more distally, through constraining livelihood and economic opportunities and impacting socioeconomic factors. In Kenya, the rural economy depends mainly on smallholder subsistence agriculture, which produces 75% of total agriculture output [286]. As might be expected, studies have found close links between environmental factors and livelihoods in Kenya, as well as poverty and socioeconomic status, due to dependence on agro-ecological conditions for crop and livestock productivity [283,284]. As a consequence, the geographical distribution of natural resources is likely to be an important determinant of patterns of poverty and marginalised populations.

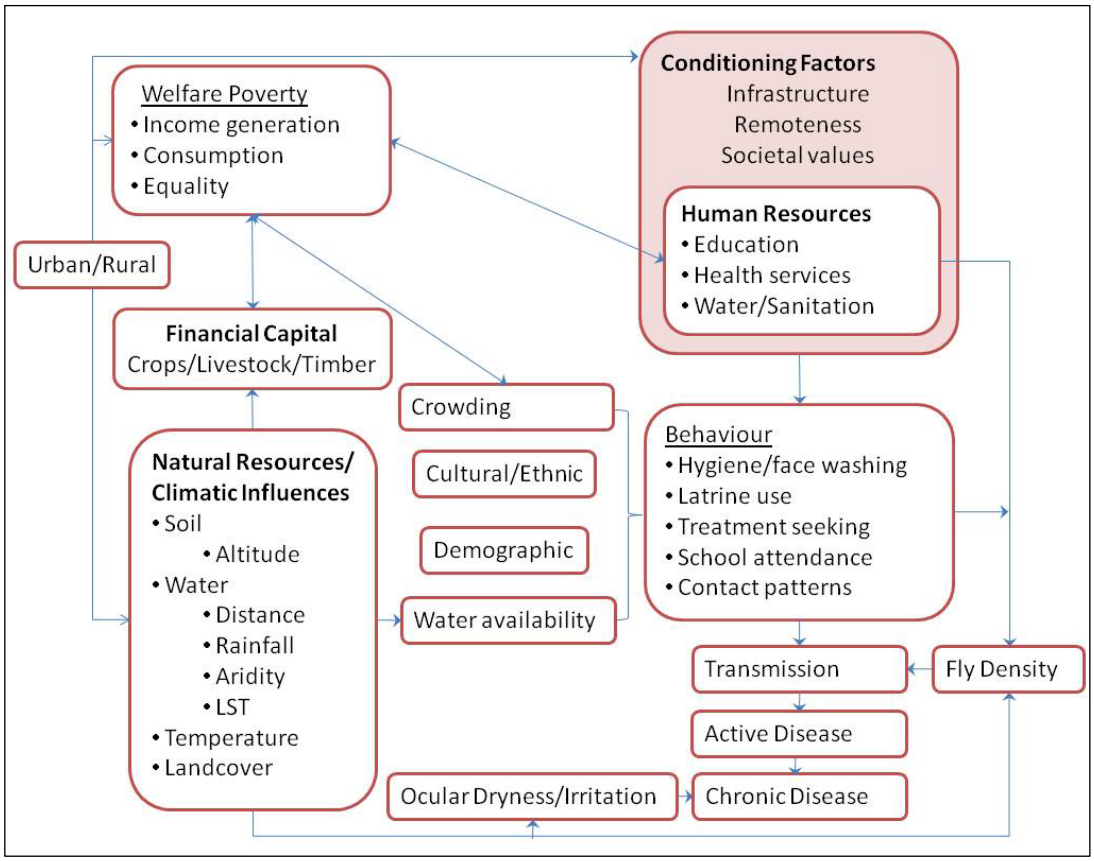



Figure 5.2 Conceptual framework of hypothesized pathways through which socioeconomic and environmental factors may determine the risk of trachoma, and mediating factors that may influence these pathways.

### 5.3.4 Data

*Field Collected data*

Data used in the this analysis are comprised of cluster-level (i.e. sublocation) baseline estimates of the number of children aged 1-9 years and numbers positive for TF, from PBPS in 17 districts conducted between 2004-2012 in Kenya. Although sublocations are technically small areas, freely available shapefiles dating from 2002 no longer provide a good match to the Kenyan Population and Housing Census boundaries. As a consequence, many clusters could not be matched to their corresponding sublocation by the KNTCP. Instead, 438 cluster were successfully assigned to a point location: either to the longitude and latitude of the centroid of the correct sublocation (84%), or, wherever possible, mapped to the corresponding town or village of the same name matched by district and division (16%). A total of 79 clusters were unable to be georeferenced.

*Covariate data*

Selection of covariate data was based on the conceptual framework described in section 5.3.3 and hypothesised relevance to causal pathways to active trachoma. Variables were available at different geographical levels, and so this approach assumes that certain aggregate socioeconomic characteristics of an area where a cluster is located (either district or constituency) could be contextual determinants of whether a cluster has a higher risk of trachoma. The following sections describe covariate data identified for this analysis, which include both socioeconomic data (including human resource assets and measures of poverty and inequality) as well as environmental and climatic data related to livelihoods, water availability and, potentially, fly density. All covariates are fully described with the source and resolution of the data in Appendix 5.1.

Socioeconomic Data

Indicators of welfare poverty included in the analyses were poverty incidence and the Gini Index. Data were available for each constituency in Kenya, from a small-area estimation mapping of household consumption expenditure [287] from the Welfare Monitoring Survey [288] conducted in 1997 using census data from the 1999 Population and Housing Census [289]. Poverty Incidence (also known as the headcount index) is the proportion of the total population in an area whose consumption is below the poverty line, defined as KShs 1239 per adult per month. The Gini Index is a widely used measure of inequality that refers to the dispersion of the distribution over the entire consumption aggregate, so that zero indicates perfect equality and one hundred indicates perfect inequality [290]. Although also a product of the poverty mapping, it is relatively independent of poverty headcount measure as it describes the distribution of welfare across the population and is not tied to the poverty line.

District-level indicators of the proportion of households with access to various water and sanitation indicators were obtained from predictive models using data from national cluster-sample surveys undertaken as part of MICs, DHS, national malaria and AIDs indicator surveys (MIS/AIS) and living standard measurement surveys [291]. Access to improved drinking water and sanitation were defined using the criteria outlined by the the Joint Monitoring Programme for Water Supply and Sanitation (JMP), and are measured by reported access and use. Access to improved sanitation facilities were defined as those that "hygienically separate human excreta from human contact," and improved ('safe') drinking-water sources defined as those that are "protected from outside contamination (especially faecal contamination)." In addition, a measure of the proportion of households with crowding (greater than 5 individuals per room) and the proportion of households reporting open defecation were also used as district-level covariates in this analysis.

While the above data were only available aggregated at the district or constituency level, other indicators directly related to socioeconomic status were available at higher

resolutions. The 2007 primary schools database was accessed from Kenya Open Data (www.opendata.go.ke) and used to measure access to education in terms of 1) the distance to the nearest primary school in kilometres and 2) the number of schools per 1000 population in each sublocation. In addition, attendance rate statistics collected in the Kenya Integrated Household Budget Survey were available at the district level [292]. Cattle density was available as a gridded surface from the Food and Agriculture Organization's (FAO) Gridded Livestock Database. However, it is noted that the last areal livestock census in Kenya was carried out more than three decades ago and these data are likely to be highly unreliable for nomadic populations and heavily extrapolated using environmental and climatic data that are also included in this analysis [293]. Finally, the distance to small scale irrigation projects, available through the World Resource Institute (http://www.wri.org/resources/data-sets/kenya-gis-data), was calculated as a measure of potential agricultural productivity and livelihood stability. Indicators on infrastructure were calculated including distance to nearest road and nearest major road using shapefiles from the Digital Chart of the World (www.diva-gis.org).

Environmental & Climatic Data

Studies have demonstrated strong geographic associations between environment and climatic conditions and livelihoods, as well as trachoma prevalence [116,157,158,294,295]. While socioeconomic data are typically only available at an aggregated level, these indices are often remotely sensed or available through routine collection of data through weather stations. Thus, the resolution of the data is between 1-10km and estimates may be extracted for each individual location, potentially explaining variation within districts. Data are fully described, including source, units and resolution, in Appendix 5.1, and included factors associated with water availability identified in the last analyses (mean annual rainfall, average and variance of the land surface temperature (LST), enhanced vegetation index (EVI; a measure of vegetation), and distance to water

bodies), temperature (mean annual temperature and altitude) and other environmental factors related to natural resources (altitude, urban extents and land cover). Mean annual rainfall, landcover and urban extents were available for multiple time points and so were matched as closely as possible with the year that each survey was conducted to account for potential changes over time.

*Data extraction*

Data described above were extracted for each survey point using ArcGIS version 10.0 (Redlands, CA) or R version 3.1.0, using the "maptools" and "raster" packages. The United Nations Second Administrative Level Boundaries data set project (SALB; [296]) were used to define district boundaries in this analysis, in order to coincide with district level water and sanitation indicators.

### 5.3.5 Categorisation and scaling

Variables were initially visualised using histograms to check for potential outliers that could indicate an error in the data or disproportionately influence the regression. Variables were then classified into quartiles and the relationship with TF observed using boxplots of the distribution of prevalence of TF within each quartiles and scatter plots. Where there was a clear pattern in the risk of TF across the factor values, variables were reclassified accordingly, otherwise categories were based off of quartiles. Access to improved source of drinking water and access to water within 1km were standardised for the analyses (by subtracting the mean and dividing by the standard deviation), as they were fit best as continuous variables and had a small scale with relatively limited range. In addition, six sites that were classified as peri-urban were reclassified as urban, due to insufficient power to look at their effects separately and based on a similar distribution of TF observed in urban areas.

### 5.3.6 Analysis

*Tools for the evaluation of spatial heterogeneity*

Spatial heterogeneity was observed and quantified throughout this analysis using three approaches: i) semi-variogram analyses, ii) local indicators of spatial association (LISA) and iii) hierarchical models allowing variance to be partitioned between levels. These represent both spatially explicit (i and ii) and non-spatial (iv) statistical approaches to quantifying spatial heterogeneity, defined here as the variation over space of the observed values from a spatially continuous process (Kolasa and Rollo 1991). Semi-variograms (previously introduced in Appendix 4.3) were used to observe and quantify spatial heterogeneity of variance in the data binned by distance lags, while LISAs provided a decomposed measure of correlation within a defined spatial range for each location (see Box 5.1). In contrast, hierarchical modelling attempts to account for gradient-type heterogeneity in the mean by inclusion of risk factors and partitioning residual variation between different levels. Thus, remaining variability between and within districts can be compared and assessed for evidence of residual spatial autocorrelation.

*Univariate models*

Univariate models were run using the "lmer" package in R version 3.1.0, which fit a generalised linear mixed effects logistic regression models with a random effect for year and district to allow for correlation within these groups. Univariate models were fitted with continuous and categorical variables in turn, and the variable with the lowest Akaike information criterion (AIC) retained for the modelling process. If included categorically, a model including the categorical variable was compared to one fitted with a quadratic term in addition to the continuous variable.

**Box 5.1 Moran's I and LISA statistics for evaluating global and local autocorrelation**

Moran's I coefficient of autocorrelation is a commonly used measure of global spatial association and quantifies the similarity of an outcome based on a spatial relationship within a distance class (i.e. first neighbours or distance based). The statistic provides a weighted measure based on the sum of the cross- products of centred pairs of points, and varies from -1 to 1; where negative values indicate negative autocorrelation (i.e. neighboring areas have dissimilar values) and positive values indicate positive autocorrelation (i.e. neighbouring areas have similar values). Moran's I is used to evaluate whether an outcome is clustered in space, compared to what would be expected under a simulated random distribution, and can provide a formal test as to whether spatial autocorrelation needs to be taken into account in an analysis.

The statistic takes the form:

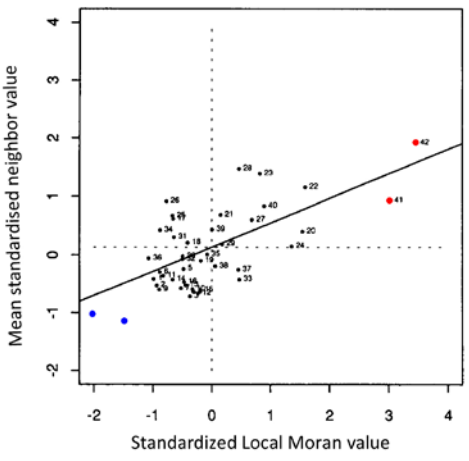
$$I = \frac{n\sum_i\sum_j W_{ij}(Z_i-\bar{Z})(Z_j-\bar{Z})}{(\sum_i\sum_j W_{ij})\sum_k(Z_k-\bar{Z})^2}$$

Where the cross product of the difference between the values ($Z$) and the mean ($\bar{Z}$), calculated for locations *i* and *j* according to a weights matrix ($W_{ij}$), are summed over all *n* locations.

Moran's I is a global statistic; i.e. an overall indicator of the degree of autocorrelation and assumes underlying spatial processes do not vary in space. This assumption can be explored by: 1) splitting the area into smaller segments assumed to have constant spatial process (as in Chapter 4), 2) decomposing global indicators into local components and 3) using geographically-weighted regression models to evaluate variation in associations.

Local indicators of spatial association (LISAs) are statistics that i) give an indication of the extent of spatial clustering of similar values around an observation and ii) sums to be proportional to the global indicator. A general formula can be expressed as $L_i = f(y_i, y_{Ji})$ where f is a function, and the $y_{Ji}$ are the values observed in the neighbourhood $J_i$ of *i*.

One LISA is the local Moran's I test [1], which decomposes the overall Moran's, by calculating a measure of similar or dissimilar disease frequency values around each point. These indicators can be graphed and mapped, in order to provide evidence of instability in the spatial association throughout the study area. In the below plot, all locations near the centre conform to the global mean. Points in the top right quadrant are areas where a high prevalence point is surrounded by other high prevalence points, and points in the bottom left correspond to areas where a low prevalence point is surrounded by other low prevalence points. Outliers in the various quadrants identify points where local autocorrelation is stronger and may unduly influence the global statistic (red and blue).



*Multivariate models*

Bayesian multivariate logistic regression models were developed in WinBUGs version 14.1 (MRC Biostatistics Unit, Cambridge and Imperial College London, UK) and run from R version 3.1.0, using the package R2WinBUGs. Addition of variables used a stepwise variable selection that was informed by the conceptual framework outlined in section 5.3.3, whereby variables more proximal to disease risk were added first. All variables were first tested for collinearity, as this can lead to model instability. Where continuous variables provided the best fit and were both highly correlated (i.e. improved water source and access within 1km), the variable deemed more proximal or relevant was selected. During multivariate modelling, some variables were found to be collinear as categorical variables. In this situation, the variable with the stronger association was retained and reduced categorisation of the secondary variable explored, based on box plots and scatter plots of its relationship with TF to identify alternative thresholds. All continuous variables were centred in the final models, to improve convergence within WinBUGS, and all variables re-included in the final model in turn and model fit was assessed using the deviance information criteria (DIC).

A binomial model for the number of children diagnosed with TF was fitted as follows:

$$Y_{ijk} \sim \text{Binomial}(n_{ijk}, p_{ijk})$$

$$\text{logit}(p_{\text{ijk}}) = (\alpha + u_k) + \sum_{\text{l}=1}^{\text{N}} \beta_{\text{l}} X_{\text{l},ijk} + \sum_{\text{m}=1}^{\text{N}} \beta_{\text{m}} X_{\text{m},jk} + \sum_{\text{n}=1}^{\text{N}} \beta_{\text{n}} X_{\text{n},k} + v_{ijk} + w_{ijk}$$

where $Y_{ijk}$ is the number graded positive for TF from a sample of *n* individuals at each of *i* sites (conducted in constituency *j* and district *k*). The resulting proportion of children with disease, $p_{ijk}$ , was modelled using a hierarchical regression model which estimated coefficients (β) for predictors at three levels: site level ($X_{l,ijk}$), constituency level ($X_{m,jk}$) and district level ($X_{n,k}$). As shown above, the model included unstructured random effect

($u_k$) which allowed the district level intercept to vary. Then around the district-level intercept, the intercept for each site was allowed to vary by inclusion of unstructured ($w_{ijk}$) and spatially correlated ($v_{ijk}$) geostatistical random effects. In addition, a district-level temporal random effect was used to assess whether allowing for variation in the prevalence of TF between different survey time points improved the fit of the model based on DIC.

The model was fitted using Markov chain Monte Carlo (MCMC), and after allowing a burn in of 60,000 iterations, the values for the intercept and coefficients were stored for 1,000 iterations and thinned one in ten. Model convergence was based on visual inspection of multiple MCMC chains on time-series plots, checking that the Gelman & Rubin diagnostic is between 0.999 and 1.2 (based on an analysis of within and between chain variances for each variable) and plots were assessed for autocorrelation, ie a pattern of serial correlation in the chain where sequential draws of a parameter were correlated [297,298]. Non-informative priors were used for α, unstructured random effects and the coefficients (normal prior with mean 0 and precision $1x10^{-6}$), the prior distribution of $\phi$ was uniform, with an upper and lower bound of 0.05 and 100, and all prior distributions of the random effects variances were given uninformative gamma distributions (dgamma(0.001,0.001)). Any district level coefficients included as continuous variables from the water and sanitation mapping were allowed a degree of uncertainly around the mean, by fitting a beta distribution based on the predicted values.

The residuals from the final non-spatial model were assessed for the presence of residual spatial autocorrelation, which can affect standard error estimates. Moran's I index was used to formally test for residual spatial autocorrelation. This statistic takes the form:

$$I = \frac{n \sum_i \sum_j W_{ij}(Z_i - \bar{Z})(Z_j - \bar{Z})}{(\sum_i \sum_j W_{ij}) \sum_k (Z_k - \bar{Z})^2}$$

Where the cross product of the difference between the values ($Z$) and the mean ($\bar{Z}$), calculated for locations *i* and *j* according to a weights matrix ($W_{ij}$), are summed over all *n* locations. In this case, a binary weights marix was used based on the minimum distance required to allow all points at least one neighbour (56 km). The significance of the resulting statistic is assessed using Monte Carlo randomisation within R. Where residual spatial correlation was found, the geostatistical random effect was modelled using an isotropic, stationary exponential decay function: $f(d_{ab};\phi) = \exp[-(\phi d_{ab})]$ where $d_{ab}$ is the straight-line distance between pairs of points $a$ and $b$, and $\phi$ is the rate of decline of spatial correlation.

### 5.3.7 Model validation

To evaluate the performance of the final model, the dataset was split into four subsets. Each quarter of the data was predicted from the remaining data, and recombined to form the validation dataset. This resulted in an observed and predicted value for all 438 survey locations, as well as calculation of the predicted probability of the prevalence of TF in a given site being greater than 5 and 10%. Prevalence thresholds used were selected based on those used to guide start of MDA at the district level (10%) and inform certification of trachoma elimination, which are based on a 5% threshold of TF within subdistricts.

The performance of the predictive model was assessed using the following statistics: sensitivity, specificity, and area under the curve (AUC) of the receiver operator characteristics (ROC). The ROC plots the sensitivity (true positive fraction) against 1-specificity (false positive fraction) to illustrate the compromises in discriminatory performance by varying probability cut offs for assigning a point as above or below a given threshold. The corresponding AUC provides a summary measure of the predictive accuracy over all probability cut-off points, commonly using values of <0.7 to indicate poor discriminatory performance, 0.7-0.8 acceptable, 0.8-0.9 excellent and >0.9 outstanding

performance [299]. Mean error and mean absolute error were used to assess bias and accuracy of predictions.

### 5.3.8 Prediction

Polygon shapefiles of 6520 sublocations in Kenya were obtained from the Kenya Central Bureau of Statistics (CBS) to use for prediction. The mean, minimum and maximum values for all gridded data were summarised for each polygon. Sublocations were classified by urban and land cover class when more than 50% of the area fell into the relevant category. Any sublocations missing covariate data (because of errors in covariate data) were assigned an average of those sublocations located within 15km. Nairobi province and sublocations that represented other major cities (Kisumu, Mombasa, Nakuru and Eldoret) were excluded from the predictive data, as both covariate and trachoma survey data are generally representative of rural areas.

Model parameter estimates were then used to predict the prevalence of TF based on the mean values for each variable within each sublocation. Prediction was performed in WinBUGs, by calculating the sum of the products of the coefficients and their corresponding covariate values for each location. Non-spatial random effects were added on to incorporate extra variability and, where residual variation showed spatial autocorrelation, an interpolated geostatistical random effect was added.

Within a Bayesian framework, predictions are in the form of a posterior distribution, which consists of the last 1000 samples from a converged model. These represent the distribution of the possible values a sublocation can take conditional on the model data, and allow estimation of the probability that the prevalence of TF will be greater than a specified threshold.

## 5.4 Results

*Distribution of active trachoma and surveyed sites*

Data on the numbers examined and numbers positive for TF were available for 438 geolocated sites from 19 districts, representing clinical data from 21,003 individuals aged between 1 and 9 years. The overall prevalence of active trachoma was 18.8%, however the distribution of prevalence values was skewed and overdispersed, with a median of 2.7% (range 0-100%) (Figure 5.3A).

The majority of clusters were identified as rural (91%), were located in arid or semi-arid climates (75%) and received less than 800mm of rainfall annually (78%). The sites were relatively underserved and remote, with approximately half of the sites in sublocations with fewer than one school per 1000 population (48%) and most located more than 30 km away from a major road (84%). The median and range values for each covariate are included in Appendix 5.2.

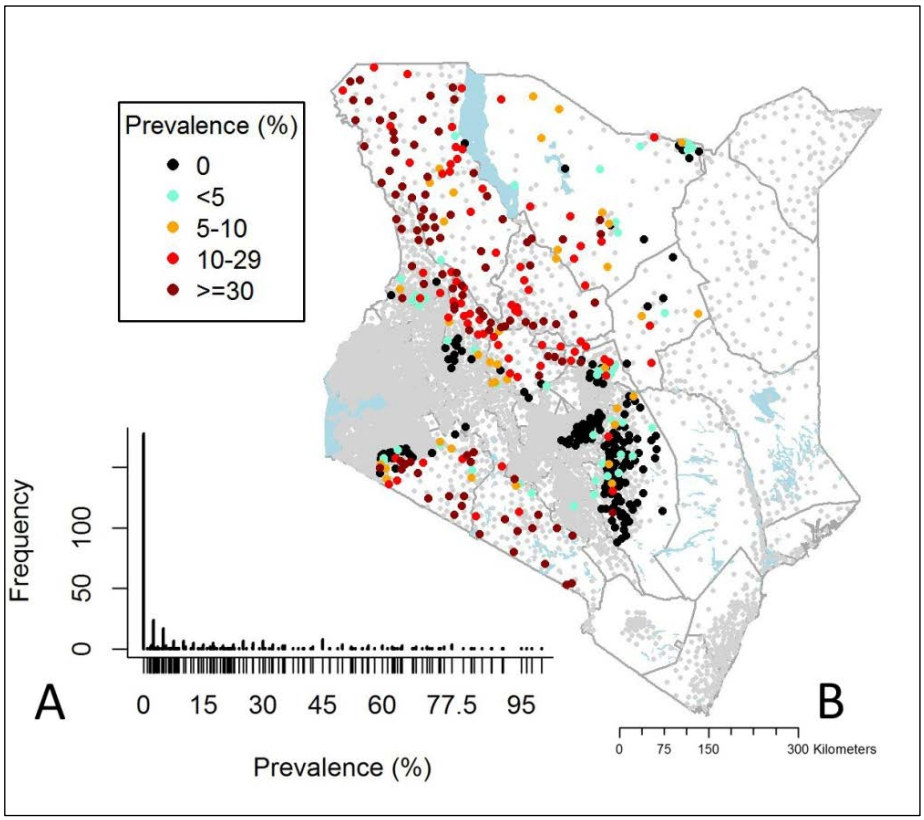



Figure 5. 3 The distribution of active trachoma prevalence in 438 sites in Kenya (A) and distribution of surveyed sites (B). Grey dots represent un-surveyed sublocations.

As might be expected, the uneven geographical distribution of data and relatively high endemicity levels reported are a reflection of how surveys have been targeted within Kenya and provide an incomplete picture of the distribution of trachoma (Figure 5.3B). This introduces certain challenges into the analysis; as estimated parameters (particularly the random effects) cannot be assumed to arise from the full distribution within Kenya and thus may not be generalisable outside of the study area.

As a consequence of using PPS to select sublocations, the distribution of surveyed sites is noted to follow the population density within districts and thus, provide relatively more information on the prevalence of TF in these areas. Areas of higher endemicity are observed throughout the northern and central districts of Rift Valley Province and bordering areas in Eastern Province. However, the disaggregated data highlight the heterogeneity in disease prevalence within districts, some which appears to be spatially structured: with a tendency for lower prevalence sites, and particularly absence of disease, to cluster along the borders of populous districts classified non-endemic.

*Socioeconomic and environmental associations*

Univariate associations

Correlation was observed between various socioeconomic, environmental and climatic variables. District-level estimates of the proportion of households with improved drinking water and proportion with improved sanitation were highly correlated (0.73-0.83) throughout Kenya and in surveyed districts. As might be expected, various subsets of environmental indices were also highly correlated, including aridity, altitude, average LST, rainfall and EVI. On the whole, however, patterns of environmental and climatic factors in relation to trachoma appear to be more variable in Kenya than in Nigeria, and thus reclassifying variables according to box plots was sufficient to avoid multicollinearity.

The majority of variables were found to be associated with the risk of TF in univariate models, as presented in Table 5.3. For the most part, odds ratios reflected expected associations between socioeconomic and environmental variables. A lower risk of disease was associated with a higher population density, greater access to education (higher number of schools, higher attendance and shorter distance to schools), improved water and sanitation, greater infrastructure (closer to major and minor roads) and more fertile, wetter environmental conditions (higher rainfall, vegetation and lower LST). This step of the analysis highlighted the importance of allowing intercepts to vary at the district level, suggesting a high degree of variation between districts and relatively less variation within districts associated with certain factors. Unexpectedly, neither poverty incidence nor inequality at the constituency level was clearly associated with TF prevalence. This may be because these income-based measures are poor predictors of behavioural determinants of disease or access to requisite education, water and sanitation in this context.

Multivariate associations

After including cluster level covariates into the model, there was no evidence of spatial structure in the residuals using semi-variogram analyses or Moran's I as a formal test ($p=0.41$); thus a spatial random effect was not included. The final multivariate model was non-spatial and retained covariates at two levels: district and sublocation. Table 5.4 presents the results from the final Bayesian model, which suggested that eight covariates should be retained. District-level mean open defecation was positively associated with the prevalence of TF, as was a greater distance to the nearest school, water bodies, nearest road, and major roads. The risk of TF was lower in sublocations with more than two schools per 1000 population, and was also negatively associated with EVI (greenness) and annual mean temperature. Upon inclusion of all covariates and the district level random effect, the addition of a temporal random effect did not significantly improve the model and so was excluded.

Table 5.3 Univariate associations between socioeconomic, environmental and climatic variables and active trachoma in 438 clusters in Kenya between 2004 and 2012

| Variable | OR (95% CI) |
|---|---|
| **Socioeconomic** | |
| District level variables | |
| Access to an improved water source [a] | |
| ≥ 55% | - |
| < 55% | 7.06 (5.44, 9.18) |
| Access to improved sanitation [a,b] | |
| ≥ 43% | - |
| < 43% (lowest quartile) | 6.91 (5.37, 8.91) |
| Open defecation [a,c] | 1.22 (1.10, 1.36) |
| Crowded living quarters (>5 per room) | 1.39 (1.13, 1.69) |
| Average school attendance | 0.92 (0.85 – 0.99) |
| Constituency level variables | |
| Poverty Incidence | 0.97 (0.92, 1.01) |
| Lower Inequality (<= 30%) | 0.50 (0.13, 1.90) |
| Cluster level variables | |
| Distance to nearest school (km) | |
| <1 | - |
| 1 – 1.9 | 1.28 (1.09, 1.50) |
| 2.0 – 4.9 | 2.50 (2.17, 2.81) |
| ≥ 5 | 2.23 (1.95, 2.55) |
| Schools per 1000 population | |
| <2.0 | - |
| ≥2.0 | 0.42 (0.33, 0.52) |
| Cattle density (animals per 5km cell) | |
| 0 | - |
| 1 – 9 | 2.62 (2.12, 3.23) |
| 10 – 49 | 1.92 (1.51, 2.46) |
| 50 - 650 | 1.21 (0.94, 1.56) |
| Distance to small scale irrigation | |
| 0-19 | - |
| 20-39 | 1.49 (1.34, 1.67) |
| 40-305 | 2.02 (1.80, 2.27) |
| **Environmental** | |
| Urban classification | 0.33 (0.26, 0.43) |
| Land cover | |
| Savannah/Grasslands | 1.23 (1.1, 1.4) |
| Barren/Sparsely vegetated | 0.54 (0.4, 0.7) |
| Distance to water bodies (km) | |
| 0 – 29 | - |
| 30 – 170 | 2.15 (1.97, 2.36) |
| Distance to road (km) | |
| 0-3.0 | - |
| ≥3.0 | 1.52 (1.39, 1.67) |
| Distance to primary road | |
| <40 | - |
| 40 – 77 | 1.87 (1.61, 2.17) |
| ≥ 77 | 3.60 (3.05, 4.25) |
| Population density | |
| 0 – 12 | - |
| 13-34 | 0.76 (0.69, 0.83) |
| 35-110 | 0.59 (0.52, 0.67) |
| 111-2550 | 0.29 (0.25, 0.34) |

Table continued on next page

Table 5.3 continued

| | |
|---|---|
| Altitude (m) | |
| 214 – 721 | - |
| 722 – 1051 | 2.84 (2.5, 3.3) |
| 1052 – 1567 | 3.83 (3.2, 4.5) |
| 1568 - 2816 | 2.31 (1.9, 2.8) |
| Enhanced Vegetation Index (EVI) | |
| ≥ 0.35 | - |
| 0.25-0.34 | 3.75 (3.11, 4.52) |
| 0.15-0.24 | 5.69 (4.64, 6.99) |
| < 0.15 | 3.68 (2.89, 4.69) |
| **Climate** | |
| Variance Land Surface Temperature (LST) (°C) | |
| 1-8 | - |
| 8-12 | 1.86 (1.66, 2.09) |
| 13-16 | 1.68 (1.44, 1.96) |
| 16-39 | 0.69 (0.60, 0.80) |
| Average Land Surface Temperature (LST) (°C) | |
| 19.0 – 32.5 | - |
| 32.6 – 36.3 | 1.45 (1.27, 1.66) |
| 36.4 – 40.7 | 1.39 (1.21, 1.60) |
| 40.8 – 52.6 | 1.13 (0.97, 1.31) |
| Mean Annual Temperature (°C) | |
| 12-18 | - |
| 19-21 | 2.21 (1.93, 2.53) |
| 22-23 | 1.47 (1.25, 1.73) |
| 24-29 | 0.80 (0.65, 0.98) |
| Mean Annual Precipitation (mm) | |
| < 800 | - |
| ≥ 800 | 0.39 (0.33, 0.46) |
| Annual aridity index | |
| Humid (>0.65) | - |
| Dry Sub-Humid (0.5-0.65) | 2.93 (2.33, 3.68) |
| Semi Arid (0.2-0.5) | 4.53 (3.63, 5.65) |
| Arid (<0.2) | 1.76 (1.34, 2.30) |

°C: Celsius; mm: millimetres; km: kilometers

[a] Proportion of households reporting; [b] Lowest quartile; [c] Standardised

Table 5.4 Multivariate Bayesian non-spatial logistic regression models for active trachoma in 438 clusters in Kenya between 2004 and 2012

| | | Mean OR (95% BCI) | |
|---|---|---|---|
| **Variable** | **Null Model** | **District Model** | **Full Covariates** |
| Open defecation | | 1.19 (1.11, 1.29) | 1.16 (1.06, 1.26) |
| Distance to schools ≥2 km | | | 1.79 (1.06, 2.80) |
| >2 schools per 1000 population | | | 0.40 (0.19, 0.79) |
| Distance to water body ≥30 km | | | 1.77 (1.08, 2.66) |
| Distance to road ≥3 km | | | 1.76 (1.14, 2.6) |
| Distance to primary road (km) | | | |
| < 40 | | | - |
| 40-77 | | | 4.11 (1.91, 7.97) |
| ≥ 77 | | | 6.17 (2.53, 12.82) |
| EVI (measure of greenness) | | | |
| ≥ 0.35 | | | - |
| 0.25-0.34 | | | 2.70 (1.16, 5.44) |
| 0.15-0.24 | | | 4.42 (1.58, 10.03) |
| < 0.15 | | | 4.39 (1.15, 11.96) |
| Annual Mean Temperature C° | | | 0.86 (0.76, 0.97) |
| Squared term | | | 0.97 (0.95, 0.99) |
| | | | |
| Alpha | -3.21 | -9.15 | -10.14 |
| | (-5.11, -0.55) | (-13.26, -5.97) | (-14.41, -6.42) |
| **Cluster $\sigma^2$** | 3.24 (2.62, 4.00) | 3.26 (2.62, 3.98) | 2.43 (1.95, 3.02) |
| **District $\sigma^2$** | 10.07 (4.12, 22.80) | 5.22 (2.02, 12.21) | 3.69 (1.29, 9.01) |
| **DIC** | **1544** | **1544** | **1538** |

OR – Odds Ratio; BCI – Bayesian Credible Interval; DIC – Deviance Information Criterion

*Heterogeneity between and within districts*

Inclusion of the district level measure of open defecation removed the large-scale trend observed in semi-variogram analysis; suggesting that second order spatial dependency was present over a range of approximately 300 km, which was reduced further by inclusion of the remaining covariates (Figure 5.4).

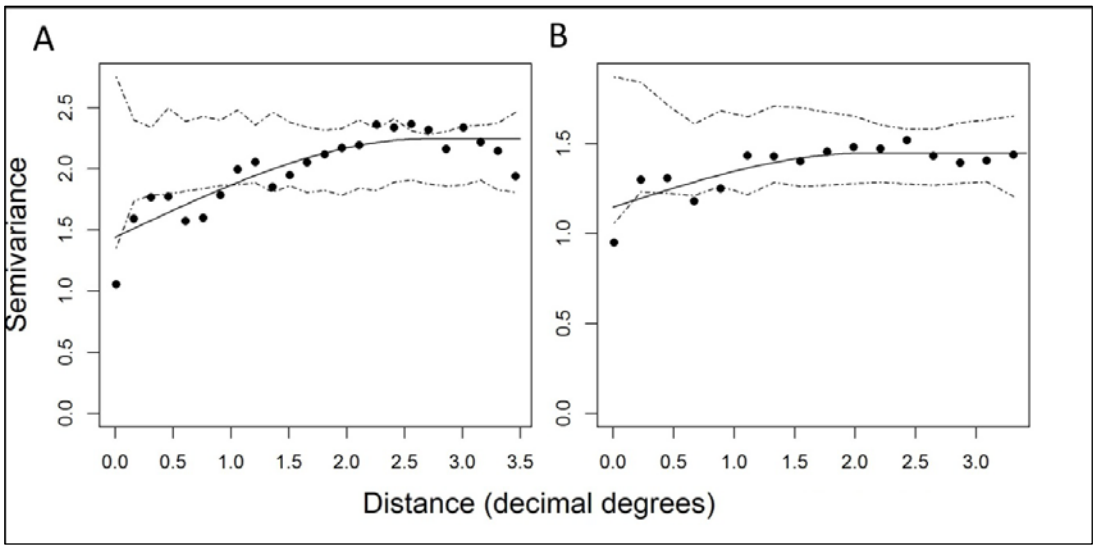



Figure 5.4 Semi-variogram plots of residual variation in TF from regression models in Table 5.4 including district (A) and cluster-level (B) covariates. Omnidirectional semi-variogram and best-fitted line of exponential spatial model is presented, with Monte Carlo simulation envelope (dotted lines). Note: at the equator, one decimal degree equates to approximately 110 kilometers.

Although formal tests for spatial autocorrelation did not support inclusion of a spatial random effect in the final model, local spatial dependency in the residuals was examined using a scatterplot of local Moran's I statistics (Figure 5.5A). This was used to identify points which had extreme values of positive (consistent with Low-Low and High-High) or negative autocorrelation (High-Low and Low-High) in relation to neighbouring points (defined by the minimum distance for at least one nearest neighbour). This plot highlights the substantially different levels of association in space compared to the global mean (which is close to zero), reinforcing the notion that trachoma has highly variable

associations in space. While there were no clear outliers likely to bias global estimates of Moran's I, the results suggest that underlying processes are not stationary, precluding a standard geostatistical model.

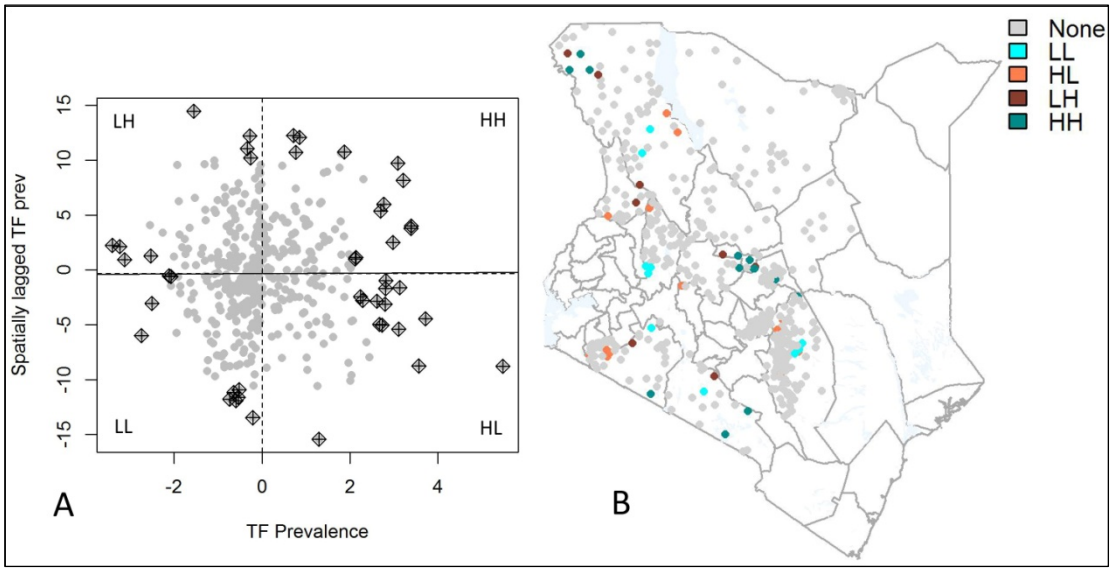



Figure 5.5 Local Moran's I scatterplot with the global Moran's I regression line fitted (A) and map of clusters with residual values above or below 2 standard deviations away from the origin (B). Positive spatial autocorrelation, or clustering of similar values, is indicated by Low-Low (LL) and High-High (HH) associations between neighbouring points, while negative spatial autocorrelation indicates association of dissimilar values: high values surrounded by low neighbouring values (HL) and low values surrounded by high values (LH) associations.

Patterns of local autocorrelation varied widely, with many "hotspots" and "coldspots" generated by dispersal patterns as well as areas where positive autocorrelation is present (Figure 5.6B). However, there does seem to be visible clustering present, with high prevalence clusters (HH) in north-west Turkana and areas bordering Tanzania radiating out towards lower prevalence areas moving closer to central and eastern Kenya.

A substantial amount of non-spatial residual variation remained in the final multivariate model (Figure 5.7). Inspection of the cluster-level residuals suggested that the model generally under-predicted the risk of TF, but over-predicted at very low prevalence values. As might be expected, high and low site-specific residuals corresponded to "hotspots" of disease where the prevalence of TF was substantially higher or lower than the district

mean. Overall, these findings highlight the spatial heterogeneity of trachoma in Kenya and emphasise the presence of non-stationary processes underlying the observed distribution.

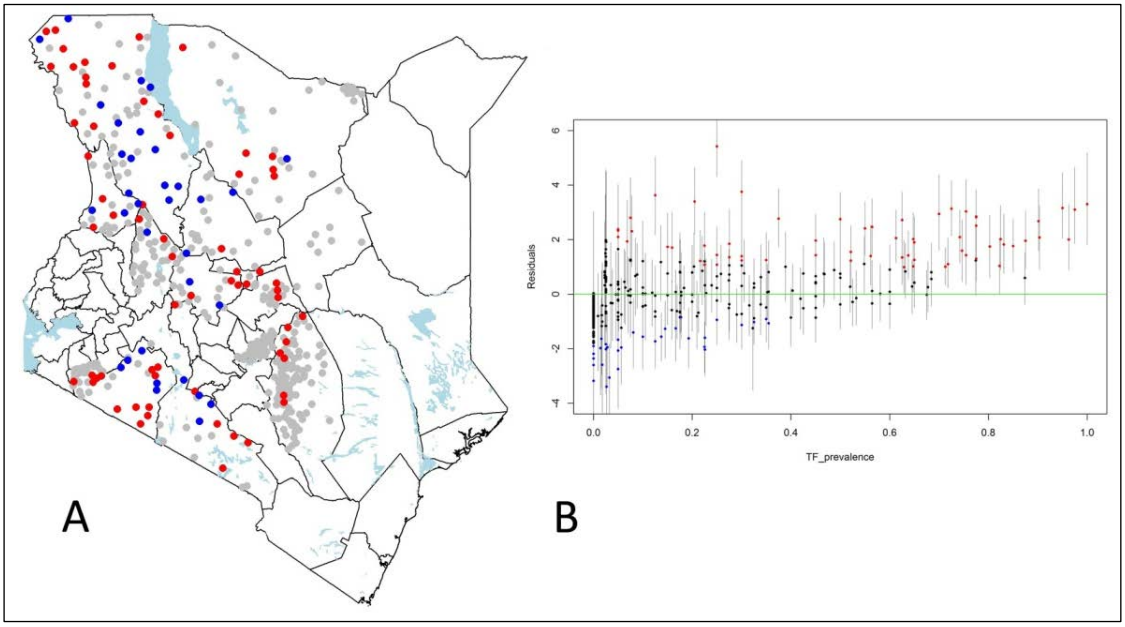



Figure 5.6 Residual variation from the final multivariate regression model in Table 5.4, captured by the unstructured random effect. Points where the 95% credible interval of the random effect is below (blue) or above (red) zero are presented.

*Model validation*

The non-spatial binomial model was validated in terms of its ability to predict the prevalence of TF for the 438 clusters for which there were survey data. The final model had a similarly poor sensitivity around both the 5% and 10% thresholds, but the specificity was quite high: indicating the model was able to correctly identify the majority of clusters where the prevalence was below these thresholds (Table 5.5). The mean error suggested a tendency to underestimate prevalence (- 7%) and based on the mean absolute error, predictions were out by ± 16%. Based on these poor statistics, a predictive map would be unable to distinguish areas of higher risk, which is of key interest to trachoma control programmes, and therefore was not developed.

Table 5.5 Validation statistics for the final multivariate model presented in Table 5.4

| | **Threshold** | | **Sensitivity** | **Specificity** | **AUC** |
|---|---|---|---|---|---|
| Nonspatial Binomial | 5% | | 0.34 | 0.86 | 0.69 |
| | | | (0.24, 0.42) | (0.78, 0.91) | (0.64, 0.74) |
| | 10% | | 0.26 | 0.91 | 0.69 |
| | | | (0.18, 0.34) | (0.84, 0.96) | (0.65, 0.74) |
| | Mean error | -0.07 | | | |
| | Mean absolute error | 0.16 | | | |
| | Correlation | 0.28 | | | |

AUC – Area under the (ROC) curve. Values in parentheses indicate 95% confidence intervals using a 50% probability cutoff

## 5.5 Discussion

Although socioeconomic factors commonly associated with trachoma are assessed subjectively and used to inform survey design in Kenya, large scale risk factor studies have not been used to support these strategies. Using cluster level data available from all baseline surveys, the present analysis used a Bayesian, multi-level modelling approach to investigate underlying risk factors and spatial heterogeneity in the distribution of active trachoma in Kenya. The results provide strong evidence of spatial variation in disease risk in relation to socioeconomic and environmental risk factors; generally supporting the notion those populations with lower access to education and located in sparsely vegetated and rural areas have a higher prevalence of disease in Kenya. However, the analysis identified substantial non-spatial residual variation both between and within districts. This suggests that, in this context and with these data, predictions of the risk of disease at unsurveyed locations will be unreliable. Generalisation to areas outside of the study area is not recommended, due to the high variability of risk between districts and perceived differences between surveyed and non-surveyed districts.

The observation that clusters with a higher prevalence of active trachoma tended to have lower measures on socio-economic indicators is broadly consistent with findings from previous studies. This analysis found that both district-level (open defecation) and cluster-level (distance to roads and access to education) socioeconomic factors helped to explain the distribution of trachoma within Kenya. The proportion of households reporting open defecation within districts helped explain large scale trends observed in semi-variogram analysis of the raw data. A higher mean risk of TF in districts where open defecation is common may reflect a higher density of flies and potential transmission of *C trachomatis* within clusters, but may also act as a marker for general trends in socioeconomic status or education. Cluster-level socioeconomic factors associated with an increased risk of active trachoma included a greater distance to roads and primary roads, as well as lower access to education (in terms of number of schools and greater distance to schools). Proximity to

a road may allow more livelihood diversity and access to human resources, and has been associated with sublocation-level poverty incidence in a meso-scale study in Kajiado district as well as over larger scales in Kenya [283,284].

The strong associations observed between access to education and active trachoma are consistent with the findings from the majority of previous studies conducted at the individual level, which observed a negative relationship between various measures of household education and trachoma [116,300-302]. One cluster-level study in Niger did report a positive association between mean number of years of education completed by the head of household and active trachoma in children aged 0-5 years; suggesting that this association may vary in different contexts according to cultural and educational differences [124]. Recent work by King et al. (2013) in four African countries has highlighted the higher risk of trachoma in non-school attending children; who logically may cluster in communities that are further away from schools or have few schools in their locality [181]. Access to education has also been associated with sublocation-level poverty incidence at meso- and macro- scales in Kenya [283,284]. As well as the potential impact of education on hygienic behaviours such as hand/face washing and use of sanitation facilities, populations with many schools may also have greater access to other resources such as better water supplies, sanitation facilities and livelihood opportunities.

Risk of active trachoma was associated with EVI, annual mean temperature, and distance to water bodies in this study. As EVI is a measure of greenness, it is higher in areas with more vegetation and reflects differences in rainfall and land-surface temperature, as well as land cover. Thus a negative association with TF is consistent with previous studies from South Sudan and Mali, which found that active trachoma was more prevalent in areas with higher aridity and lower rainfall [157,158]. In addition to its relevance to water availability and subsequent use for hygienic purposes, greater aridity and dust in areas with low vegetation may contribute to disease by i) drying the conjunctiva and increasing susceptibility to infection and/or ii) irritating the eye further and causing more chronic

disease conditions [15]. Perhaps more importantly, EVI is closely correlated with the population density, irrigation and cattle density in Kenya. This reflects its potential relevance to socioeconomic mapping in terms of agricultural productivity and pasture potential, which is consistent with a study relating vegetation indices to sublocation-level poverty incidence in a district-level study in Kenya [283,284]. The increased risk of TF in clusters located at large (over 30 km) distances from water may also reflect lower access to water for irrigation and have similar links to agricultural and livestock productivity. Finally, after controlling for the above factors, the prevalence of TF was negatively associated with the mean annual temperature. As discussed in the previous chapter, associations between temperature and active trachoma have been reported in other studies [152,153] and may possibly influence the density of flies [135].

The high levels of spatial heterogeneity of trachoma prevalence between and within endemic districts in Kenya supports widely held beliefs around its distribution within countries. A more detailed exploratory spatial analysis of this variation suggested that, prior to including cluster-level risk factors, levels of spatial autocorrelation varied across Kenya. Based on local Moran's I, spatial autocorrelation was particularly strong in Turkana and southern areas of Kenya near the Tanzanian border. Although inclusion of cluster-level covariates accounted for observed spatial dependency, high levels of residual, aspatial variation persisted throughout the study area. These findings support the current strategy of the KNTCP, which administers a pre-survey risk questionnaire at the sub-district level in endemic districts to create smaller, more homogenous trachoma districts, in order to generate more precise data at finer resolutions. Higher resolution census data for more proximate risk factors (including distance to water and crowding at the sublocation level) may improve the models and provide a more reliable basis for risk mapping within endemic districts. This analysis also highlights the potential use of satellite derived data, particularly EVI and distance to roads, in conjunction with data on water and sanitation interventions to identify areas with high environmental risk.

Much of the residual variation was not observed to be spatially structured, and thus may be attributed either to risk factors that are truly aspatial (such as behavioural differences that do not also aggregate in space) or variation that is spatially structured at smaller scales (i.e. within cluster). This small scale, nonspatial heterogeneity in trachoma prevalence presents a constraint on the use of models to predict its distribution in unsurveyed areas. The poor predictive ability of the model developed in this chapter contrasts with work published by Clements et al. (2010), who used a geostatistical model to predict the risk of trachoma in South Sudan with only two covariates: rainfall and landcover [158]. Validation statistics of this model found a very high predictive ability to discriminate prevalence of active trachoma at the location level relative to thresholds of 0%, 10%, 40% and 70% (ROC 0.96). This model also under predicted prevalence, but to a much lesser degree (mean prediction error: -0.012). There are a number of potential reasons why these models may perform so differently. First, South Sudan has a strong environmental/climatic gradient across the country, much like Nigeria. In contrast, the exploratory PCA using Kenyan environmental data highlighted the more variable climatic combinations that exist across the country, which may lead to seasonal variation in water availability and different behaviours around water usage. Second, South Sudan is much less developed than Kenya and is likely to have less heterogeneity in terms of access to education, water and sanitation. Thus, environmental determinants of water availability may be the driving factor behind hygienic behaviours. Third, with the exception of Western Equatoria, the data used to model trachoma in South Sudan were clustered within districts and may not have provided a full distribution of the true variation within districts. However, it is also likely that patterns of risk of active trachoma differ between countries and its geographic distribution is less likely to follow broad environmental trends in countries with more variable socioeconomic conditions. It would be interesting to compare the performance of a cluster-level model using the TT data from the previous chapter, as one might expect the distribution of trichiasis (and historical TF) in Nigeria to reflect environmental drivers of water availability rather than more recent interventions.

The most striking discrepancy between districts currently identified as endemic by the KNTCP (Figure 5.1) and areas that would be predicted to be at high risk based on identified risk factors, are areas in North Eastern Province. This province is highly remote, has limited access to education or health facilities, is arid and has little vegetation and is populated by nomadic pastoralists; yet reportedly has no trachoma. Anecdotally, this is attributed to face washing practices believed to be practiced by the pastoralist Muslim Somali inhabitants, however no empirical surveys exist to support this. Regardless of whether this is uniformly true throughout the province or is correctly attributed to a specific religious/ethnic characteristic, hygienic behavioural factors are likely to be the most important determinants of active disease and moderate observed associations in any risk factor analysis. As discussed in the previous chapter, allocation of water for hygienic purposes depends on having a reliable water source, the ability to transport large enough quantities to exceed basic needs, the knowledge of how much water is required and the perceived importance of washing [54,125,132,275].

The assumption that the relationships between risk factors and active trachoma are constant in space (stationary) is a strong assumption that probably does not hold true in Kenya. A recent study suggested that geographic associations with poverty in Kenya varied by province and similar findings might be expected with diseases that have complex relationships with behavioural, socioeconomic and environmental factors [283]. One extension to this analysis might be the use of geographically weighted regression to explore non-stationarity in observed relationships. This technique calculates locally weighted regression models using a set of more proximal points, in order to estimate a coefficient for each survey site [303], which may be useful for hypothesis formation about where behavioural factors strongly moderate observed associations.

The results from this chapter should be considered in the context of a number of limitations of the data used in the models and assumptions behind the modelling process. First, as discussed in the methods section, the distribution of the data is heavily biased

towards the most trachoma endemic areas in Kenya. This represents the most significant limitation of the model, as there is relatively little data in districts where the overall prevalence is less than the 10% threshold for mass treatment. This has the effect of restricting the range of covariate data in which predictions could be based and potentially invalidating any predictions made outside of the surveyed districts, which would be of most programmatic use. Second, cluster level surveys are conducted so that they are representative of a sublocation (small area) rather than an exact point location. Although sublocations do tend to be small (between 5 and 40 km across), this assumption is not ideal, as the appropriate measure at this level would be "proportion of the subdistrict classified as urban" or "mean distance to a school". Third, not all data were able to be geolocated, particularly those areas in northern Eastern province surveyed most recently. As these surveys found disease to be absent in many sublocations, this is likely to have contributed to the overprediction observed in these areas. Finally, there will invariably be measurement errors both in covariate data available and diagnosis of TF. These misclassifications could influence observed associations and introduce error into subsequent predictions. Clinically active follicular trachoma is known to have low specificity as a marker of *C trachomatis* infection, particularly in areas with low levels of endemic trachoma and clinical signs may be due to non-chlamydial bacterial pathogens [15]. This may contribute to observed "hotspots" in districts which have a very low prevalence of disease overall.

Identified risk factors and small scale heterogeneity in Kenya have a number of implications for the design of surveys. First, the prevalence of TF was negatively correlated with EVI, which in turn is correlated with population density and potential socioeconomic factors. As a consequence, use of PPS to select clusters within districts is likely to select more clusters from low-risk areas. Where the aim is a district-level population representative prevalence estimate to initiate SAFE implementation, perhaps this is not a concern. However, the lower geographic coverage in areas that are more remote and rural, may provide unreliable estimates for subdistricts at higher risk and limit

inferences around elimination targets. A previous study by Schemann et al. (2002) supports this hypothesis, as it found that smaller villages (which using PPS would be less likely to be surveyed) were more likely to have a higher prevalence of trachoma in Mali [116]. Second, despite the presence of spatial autocorrelation in active trachoma prevalence estimates over approximately 300 km, the high levels of non-spatial heterogeneity reinforce the need for high resolution and robust survey strategies. Finally, the clear negative association between the prevalence of trachoma at the cluster-level and access to schools suggests that use of a school-based survey platform may introduce bias in certain areas.

In summary, this chapter illustrates how existing data might be used to identify risk factors underlying the distribution of active trachoma and investigate spatial heterogeneity between and within districts. Although the high levels of heterogeneity in risk preclude use of this model to predict the risk of trachoma, the findings highlight a number of implications for survey design. As the need for reliable epidemiological data increases with the rapid scale up of trachoma mapping activities in many countries to meet elimination targets, alternative survey methodologies have been proposed to fill these gaps. The next chapter will compare two trachoma survey methodologies using a computerised simulation approach.

# Chapter 6: Comparing the performance of cluster random sampling and Integrated Threshold Mapping for targeting trachoma control, using computer simulation

## 6.1 Overview

Chapter 5 highlighted the heterogeneity in risk of active trachoma and variable spatial structure observed throughout areas classified as endemic in Kenya. These results emphasise the importance of reliable data on the prevalence of active trachoma to guide programmatic action. Cluster randomized surveys (CRS) are currently recommended for baseline and post-intervention surveys and are used to provide district-level prevalence estimates of a specified precision. However, this methodology is relatively costly and time-consuming, particularly when implemented at the required scale to achieve global coverage of survey data. While trachoma control programmes currently use a community-based strategies to implement CRS, other diseases often use school-based survey platforms, including soil-transmitted heminths (STH) and schistosomiasis [304,305] and malaria [306,307]. The risk of trachoma is widely believed to vary by attendance (and enrolment) in trachoma endemic contexts [97,208], which is supported by results from the previous chapter. Recently, however, Integrated Threshold Mapping (ITM) has been proposed as an integrated and cost-effective means of rapidly surveying trachoma in order to classify districts according to treatment thresholds.

In this chapter, a computerised sampling approach is used to evaluate the equivalence of ITM and CRS, and explore the impact of varying key parameters on the performance of these sampling methodologies. This chapter has been published in *PLOS NTDs*: *Smith JL, Sturrock HJ, Olives C, Solomon AW, Brooker SJ (2013) Comparing the performance of cluster random sampling and integrated threshold mapping for targeting trachoma control, using computer simulation. PLoS Negl Trop Dis 7: e2389*. I conceived the study design, was

responsible for the analysis and drafted the manuscript with high level input from the other authors.

## 6.2 Introduction

Since the establishment in 1998 of the Global Elimination of Trachoma by 2020 (GET2020) Alliance, an increasing number of endemic countries have implemented national programmes in an effort to meet elimination targets. These targets are less than one case of trachomatous trichiasis (TT) per 1000 total population unknown to the health system, and <5% trachomatous inflammation–follicular (TF) in children aged 1-9 years, at the sub-district level [3]. In response to these targets and a need to finalise global mapping in time to allow programmatic impact, there has been a renewed interest in developing cost-effective mapping strategies and integrating survey and control activities with other neglected tropical diseases (NTDs) [153,308-310]. Population-based prevalence surveys (PBPS) remain the accepted "gold standard" for estimating the prevalence of trachoma within target populations and usually use cluster random sampling (CRS) to select non-overlapping subpopulations (clusters)[85]. This methodology is relatively expensive, however, and there is interest in developing cheaper and more rapid methods as well as integrating with other disease surveys [311]. Integrated Threshold Mapping (ITM) is a sampling methodology currently being put forward as a cost-effective means of rapidly surveying trachoma in remaining unmapped districts and to allow treatment decisions to be made and timely scale up of interventions to be achieved [95].

Both CRS and ITM diagnose trachoma based on the presence of key clinical signs using the 1987 WHO simplified grading system: TF in children aged 1-9 and TT in adults aged over 14 [26]. These measures are easily collected in the field and routinely used to inform intervention strategies. For example, in districts where the prevalence of TF is greater

than 10%, annual mass drug administration (MDA) of azithromycin should be implemented (Table 6.1).

Table 6. 1 Azithromycin treatment strategies and classification at designated TF prevalence thresholds [312]

| TF Prevalence (district level) | Classification | Treatment strategy |
|---|---|---|
| <5% | Active trachoma not a public health problem | No MDA |
| 5-9.9% | Hypo-endemic | Determine need for MDA at sub-district level |
| 10-29.9% | Meso-endemic | MDA at district level (≥3 years[a]) |
| >30% | Hyper-endemic | MDA at district level (≥5 years[a]) |

[a] before reassessment to determine whether to stop or continue

However, ITM differs from the accepted "gold standard" survey methodology in a number of important ways (briefly outlined in Table 6.2), including the use of a school-based sampling platform for children aged 1-9 years and a different age distribution of participants. Differences in selection of participants can have a varying impact on resulting prevalence estimates and treatment decisions, depending on how disease is distributed in the population. Age patterns of active trachoma indicate a higher burden in children under 10 years, with the highest prevalences found in preschool-aged children in hyper-endemic areas [220,313]. A recent meta-analysis has reported the risk of TF to be lower in children attending school in four African countries [181], supporting widely-held beliefs that the risk of trachoma is likely to vary by attendance (and enrolment) in trachoma endemic contexts. While CRS takes a community-based sample, that theoretically is representative of the true age distribution and prevalence of disease in this population, ITM may over- or under-sample certain age groups and introduce a bias if the risk differs between enrolled and non-enrolled children. In addition, clustering of active trachoma by household has been observed in a number of studies [106,108,179], and the precision of estimates from both sampling methodologies are expected to be influenced by this factor. A careful evaluation of how participant selection and variation in epidemiological parameters impact prevalence estimates and treatment decisions using the two methodologies is warranted.

Table 6.2 Methodological differences between cluster random sampling (CRS) and Integrated Threshold Mapping (ITM)

| | CRS | ITM |
|---|---|---|
| Platform | Community-based | School-based with younger children brought from the community |
| Cluster selection | Probability proportional to size or random selection | Random selection: minimum 2 per subdistrict |
| Participant selection | Household | Children aged 6-9 at school & 1-5 year old children from communities |
| Sample size and age groups | 100 aged 1-9 years | 25 aged 1-5 years and 25 aged 6-9 years |

Although ITM was internally validated against CRS during the pilot phase of the methodology's development in Mali and Senegal [95], and used in a nationwide mapping of Togo [96], these evaluations were limited by several issues. In Mali and Senegal, only a single district was surveyed providing limited evidence in trachoma meso- and hyper-endemic settings. Furthermore, the CRS sample in these settings was partially comprised of existing ITM clusters, which could potentially have biased the CRS estimates and resulted in an overly-optimistic assessment of ITM. Finally, although this methodology was used to map trachoma in all districts in Togo, it is a trachoma hypo-endemic country and so results could not be generalised to other trachoma endemic contexts.

Computerised sampling simulations have provided a convenient platform recently to evaluate alternative survey designs for tropical diseases including soil-transmitted helminthes, trachoma and schistosomiasis [99-101,228]. This approach entails generating realistic "gold standard" data for a population that maintains observed disease clustering, using epidemiological parameters derived from existing datasets. A survey methodology can then be evaluated using these data by selecting participants according to the specified sampling protocol and deriving a prevalence estimate. There are a number of advantages to using computerized sampling simulations to compare survey designs, including the ability to i) simulate fully enumerated data (allowing estimation of "true" prevalence of

disease), ii) incorporate sampling error by repeating simulations a large number of times, iii) evaluate performance across a range of endemicity settings and iv) explore how variation in factors underlying clustering of disease in communities will influence the performance of sampling methodologies. A similar comparison performed empirically might be prohibitively expensive to carry out, as it would require at minimum a full census survey of a large number of districts across different endemicity settings and implementation of each sampling protocol in the field.

This analysis used computerised sampling simulations to compare the precision and accuracy of district level prevalence estimates based on ITM versus CRS. Furthermore, we compared the performance of both survey methodologies, in terms of their ability to correctly classify districts according to established TF prevalence thresholds and the factors that affect the degree of equivalence. Equivalence between the two survey methods, under different scenarios, was formally evaluated by testing the null hypothesis that ITM yields the same programmatic results compared to CRS.

## 6.3 Methods

### 6.3.1 Overview

Simulating sampling designs require gold standard data from which to draw samples and compare sample estimates. There are no perfect datasets available to conduct this analysis, which would necessitate standarised, full census datasets of demographic and epidemiological information for multiple districts. An alternative is to simulate these data, using parameter estimates from empirical data to generate realistic pseudo gold standard data on active trachoma [314,315]. In this study, full census data from a single community are used to parameterize disease clustering and, incorporating information on between-district variation, to 'expand' the available dataset and generate data for a large number of simulated communities within many districts.

### 6.3.2 Empirical datasets

*Community level dataset*

One dataset used to parameterize this analysis comes from Kahe Village, Rombo District, northern Tanzania, which is a single community that consists of 90 local administrative units called balozis. A fully enumerated census and survey of trachoma was conducted in April to June 2000 by means of a house-to-house survey, using the WHO simplified grading system, prior to the initiation of any interventions against trachoma. A single examiner collected these data and clinical grading was validated through a live-patient inter-grader agreement exercise using an international expert reference grader with an agreement of 100% for TF. The dataset in total consists of 5748 individuals in 1103 households, with between 41-126 individuals and 8-23 households per balozi. The dataset included information on the presence or absence of TF in 1831 children aged 1-9 years, where the prevalence was 33.4%. Data on school enrolment were also available for a subset (23%) of children aged 6-9 years.

The demographic (age and gender) and household structure present in Kahe was used for all simulated communities in the expanded dataset. This dataset was also used to provide initial values used to parameterize the models, including the relative risk of TF between children aged 1-5 years and 6-9 years and the intra-cluster correlation (ICC) measuring the degree of disease clustering within households. The subset of data with information on enrolment provided an initial value for the relative risk of TF in children aged 6-9 who were enrolled in school to those who did not. In addition, this dataset was used to assess whether there was an additional household level risk associated with having a schoolgoing/non-schoolgoing sibling and inform the simulation model (results presented in Appendix 6).

*District- level dataset*

Data on the prevalence of active trachoma were available for 305 clusters (non-overlapping sampling populations) from 29 districts in Kenya, surveyed as part of the National Trachoma Control Programme between 2004-2012 and included within the Global Atlas of Trachoma [278,316]. These data represent available disaggregated data in a broadly similar context, and importantly include nearly all endemic districts. These data were used to model variation between and within districts (Figure 6.1) in order to inform simulation of realistic district and cluster-level prevalence values.

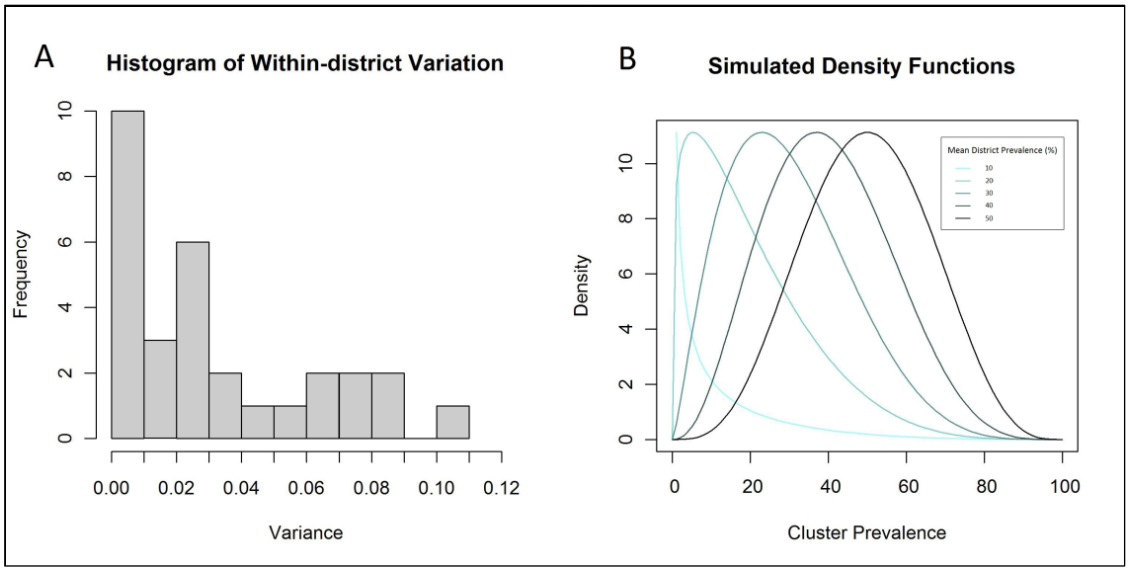



Figure 6.1 Histogram of the district variance of TF in Kenya (A) and density functions used to simulate data (B). Variance in the prevalence of active trachoma was quantified within 29 districts in Kenya. The mean within-district variance was then used to inform beta density functions for simulating cluster-level prevalence values for varying district level prevalence values.

### 6.3.3 Dataset expansion

The process of expanding the community dataset to simulate realistic data for 100 communities within each of 100 districts is fully described in Appendix 6. In brief, district

level prevalence estimates were generated covering all endemicity classes and used to simulate community level estimates of TF in children aged 1-9 years. The burden of TF within each simulated community was distributed among the population according to parameters initially defined by the above datasets (Table 6.3) in order to maintain disease clustering within households and subpopulations. Enrolment is defined as being "officially registered in a given educational programme, or stage or module thereof, regardless of age" [317], while attendance refers to an individual's presence at school at a given time. In these simulations we have assumed that all enrolled children attend on the day of the survey, however recognize that enrolment statistics are typically much higher than attendance. Enrolment was varied to assess the impact it has on sampling performance, and children identified as "school-going" were allowed to vary during the simulation process.

To avoid basing simulations on data parameterised by single village-level and district level datasets, additional pseudo-gold standard datasets were simulated varying each of the epidemiological parameters identified in Table 6.3 while holding other factors constant. This allowed an exploration of the impact of those parameters on the performance of the different sampling methodologies and the robustness of the different sampling approaches over other epidemiological settings. This included varying the level of household clustering quantified by the ICC, the relative risk of TF observed between enrolled and non-enrolled children, and the relative risk of TF between age group using parameters shown in Table 6.3.

Table 6.3 Description of key epidemiological parameters used in the simulation model and sensitivity analysis[a]

| Key Parameter | Rationale | Estimation method & Initial Value | Sensitivity Analysis |
|---|---|---|---|
| 1. Age-specific prevalence of TF: TF in 1-5 years versus 6-9 years | In order to expand a cluster level prevalence estimate in children aged 1-9 years to the two age groups, need to know RR between groups. This will likely vary with endemicity. | Estimated from gold standard datasets Initial value: 2.0 | Varied parameter: 1.3, 1.5, 1.8, 1.0, 2.0 |
| 2. Risk of TF in enrolled children vs non-enrolled children | Likely that enrolled children will have lower TF prevalence | Estimated from gold standard datasets Initial value: 0.5 | Varied parameter: 0.25, 0.33, 0.5, 0.75, 1.0 |
| 3. School attendance | This will affect the sample size in schools of 6-15 year olds and affect the impact of parameter 2. | Ministry of Education data Initial value: 0.7 | Varied parameter: 0.4 and 0.7 |
| 4. Clustering within households: risk of TF in children aged 1-5 years with a TF positive/negative sibling | Clustering at the household level will mean that children with TF positive siblings are more likely to have TF | Estimated from gold standard datasets Initial value: 0.2 | Varied parameter: 0.1, 0.2, 0.3, 0.4, 0.5 |

TF: trachomatous inflammation–follicular; RR: relative risk

[a]Random selection of 20 clusters were used in simulations for both methodologies

### 6.3.4 Sampling simulations

*Survey methodologies*

CRS for trachoma uses a standard two-stage or multi-stage design, often comprising a random selection of approximately 20 villages (clusters) at the first stage and selection of households at the second [56]. Selection of households may be carried out using simple random sampling, systematic sampling, the random walk or compact segment sampling. The sample size for CRS is calculated by defining parameters which include: expected prevalence estimates, acceptable error margin or precision, required confidence level, and design effect. In contrast, ITM employs convenience sampling of school children, pre-school children and women of child-bearing age to estimate the prevalence of trachoma [94]. At least two villages are selected per sub-district, with a minimum of 20 villages selected per district. In each village, a single school is randomly selected as the testing site. Children enrolled at that school are asked to come to the location, and adults from the community are also asked to assemble here and bring children aged 1-5 years. Systematic

sampling is then used to select 25 children aged 1-5, 25 children aged 6-9 and 50 adult women (or 100 adults) aged ≥15 years.

*Sampling process*

A computerized simulation approach, using Monte Carlo methods, was used to randomly select 20 clusters from each district and sample individuals within each cluster according to the protocol for ITM and CRS (Table 6.2). For this analysis, a sample size of 100 individuals was assumed for CRS and participants selected from a random selection of households until the sample size met. It was assumed that children aged 1-5 years that would be brought to schools by their mother (or other adult household member) and sampled by ITM would be those with school-going siblings aged 6-9 years. We explored the impact of this assumption by also sampling a random selection of children in this age group. Sampling simulations were repeated 1000 times on each dataset using both methodologies.

### 6.3.5 Analysis

District-level prevalence estimates generated by the two sampling methodologies were used to classify districts according to endemicity class for each simulation, using categories corresponding to established treatment thresholds: hypo-endemic (<10%), meso-endemic (10-30%) and hyper-endemic (>30%) (Table 6.1). The performance of each method was then quantified in terms of the proportion of times each district was correctly classified over 1000 simulations according to TF treatment thresholds.

*Operating Characteristic (OC) curve*

Due to the complicated sampling distributions of these methodologies, it is not possible to calculate the full theoretical OC curves. However, we can visualize the empirical OC curves resulting from these simulation studies, which are generated from the proportion of times a district is correctly classified in each endemicity class using the two methodologies, over a "range" of district prevalence values. For each survey method, this allowed us to establish the range of district prevalence values in which the probability of correctly classifying a district is less than or equal to 0.80.

*Equivalency*

Overall agreement in district endemicity classifications by the two methodologies was assessed using a weighted kappa-statistic. This statistic provides a measure of agreement between the two methods adjusted for chance, where a value of zero indicates agreement no better than chance. Weighting is useful when there are more than two ordered categories, so that the magnitude of disagreement between categories is allowed to vary (i.e., difference between <10% and 10-30% is not as great as that between <10% and >30%). Increasing kappa values correspond to better agreement between the two methods, where agreement is often interpreted as slight (<0.2), fair (0.2-0.4), moderate (0.4-0.6), substantial (0.6-0.8) and almost perfect (≥0.8) [318].

Equivalence between the two survey methods was formally evaluated by testing the null hypothesis that ITM yields the same programmatic results compared to CRS. The distribution of the difference in the proportion of correctly classified districts by ITM and CRS was generated and the mean and 95% CIs plotted in relation to delta, $\Delta$, a threshold corresponding to a predefined level of difference deemed programmatically important. In these analyses, delta was initially assumed to be 20%, based on the rationale that this is equal to 80% of the simulations being classified the same by ITM and CRS and roughly

corresponding to a standard level of acceptable error. Where the CI fell within this range, the survey methods were classified as equivalent for that district, while those that fell outside were classified as not equivalent and those that overlapped with the thresholds as inconclusive. Districts were stratified by the relative risk of TF and endemicity class to evaluate whether the equivalence of the two methodologies varied with these parameters.

## 6.4 Results

### 6.4.1 Estimated prevalence

Overall, the results indicate that ITM under-estimates the true prevalence of TF compared to CRS and that the magnitude of difference between estimates from these methodologies increases with endemicity. This is illustrated in Figures 6.2 and 6.3, which compare the two sampling strategies where all parameters are set to the initial values described in Table 6.3. Figure 6.2 presents filled density plots in example hypo-, meso-, and hyper-endemic districts, where the red line represents the true prevalence value for that district, the curves represent the distribution of prevalence estimates from the 1000 simulations using the CRS method (grey) and ITM (red). The results suggest that the systematic error resulting from school-based sampling is proportional to the prevalence, so that the absolute bias increases linearly as the prevalence increases.

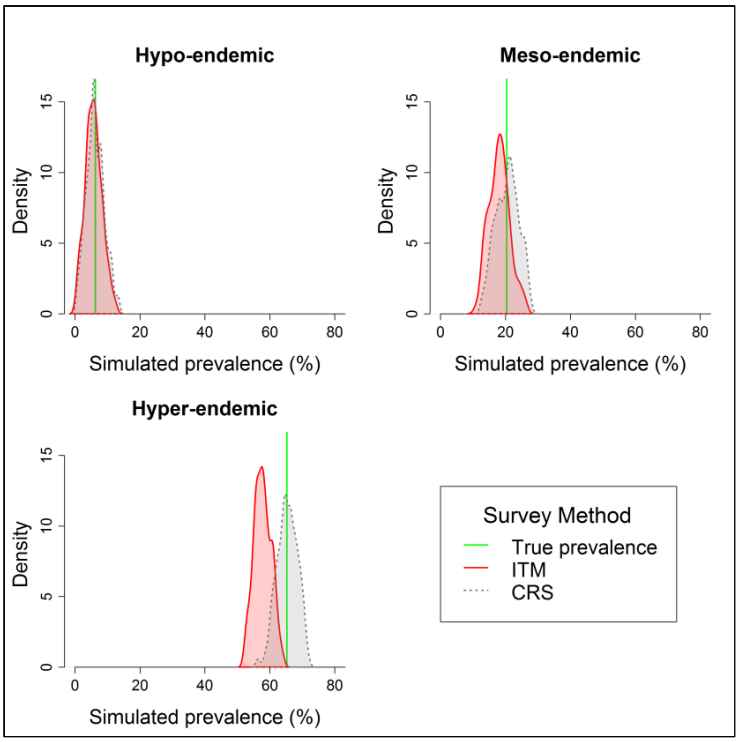



Figure 6.2 Density plots of prevalence estimates generated by CRS and ITM sampling methodologies. Plots are generated using simulated data and present results from a single district within each endemicity class. The red line represents the true district-level prevalence, the curves are histograms of values from 1000 simulations using the CRS method (grey) and ITM method (red).

### 6.4.2 District-level classification

Figure 6.3 plots the proportion of times each of 100 districts were correctly classified (of 1000 simulations) against the district-level true prevalence for each sampling methodology, where the relative risk of TF in enrolled and non-enrolled children is equal to 0.5 and enrolment rate is 0.7. The green lines correspond to the treatment thresholds while the areas shaded red and grey around these thresholds have a "higher" risk of misclassification by the corresponding sampling methodology. Within these prevalence ranges, districts will be correctly classified less than 80% of the time. Performance of both CRS and ITM was lower closer to treatment thresholds. Compared to CRS, where misclassification error was fairly symmetrical around treatment thresholds, ITM tended to underestimate the prevalence of TF, resulting in a corresponding shift and widening of the region where potential error is known to be high.

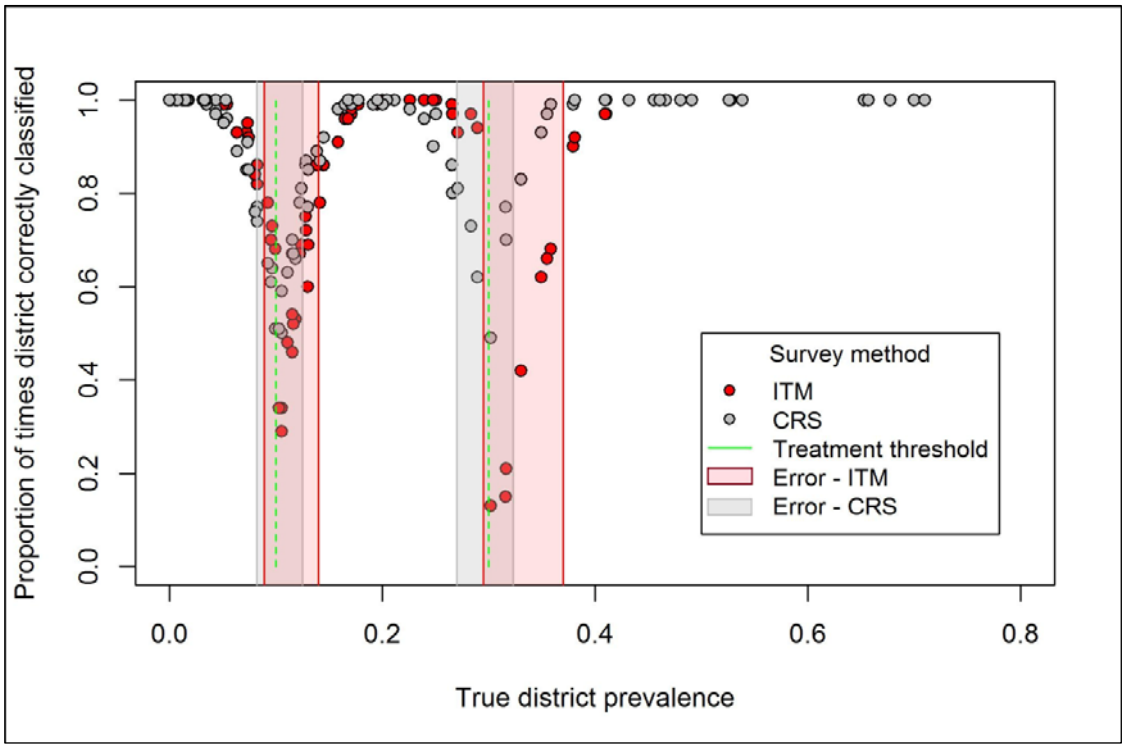



Figure 6.3 Performance of ITM and CRS compared to true prevalence. The proportion of times each of 100 districts were correctly classified by ITM and CRS were compared to true prevalence, where the relative risk of TF in enrolled and non-enrolled children is equal to 0.5 and enrolment rate is 0.7. The green lines correspond to the treatment thresholds and the boxes in red and grey around these thresholds to areas of "higher" misclassification, where the districts will be correctly classified less than 80% of the time.

Using a relative risk of TF in enrolled versus non-enrolled children equal to 0.5, there was "almost perfect" agreement (Kappa=0.86) in district-level endemicity classification between ITM and CRS overall in the 1000 simulated samples. However, agreement between ITM and CRS decreased with increasing endemicity category, with substantial agreement found in hypo-endemic districts (Kappa=0.71) and only moderate agreement in meso-endemic (Kappa=0.47) and hyper-endemic districts (Kappa=0.41).

The equivalence analysis in Figure 6.4 illustrates changes in the distribution of the difference in the proportion of correctly classified districts by ITM and CRS by endemicity class. The results suggest that the two sampling methodologies are equivalent in hypo-endemic areas but the wider confidence intervals in meso- or hyper-endemic areas indicate that they less likely to be equivalent in these settings due to a greater degree of bias.

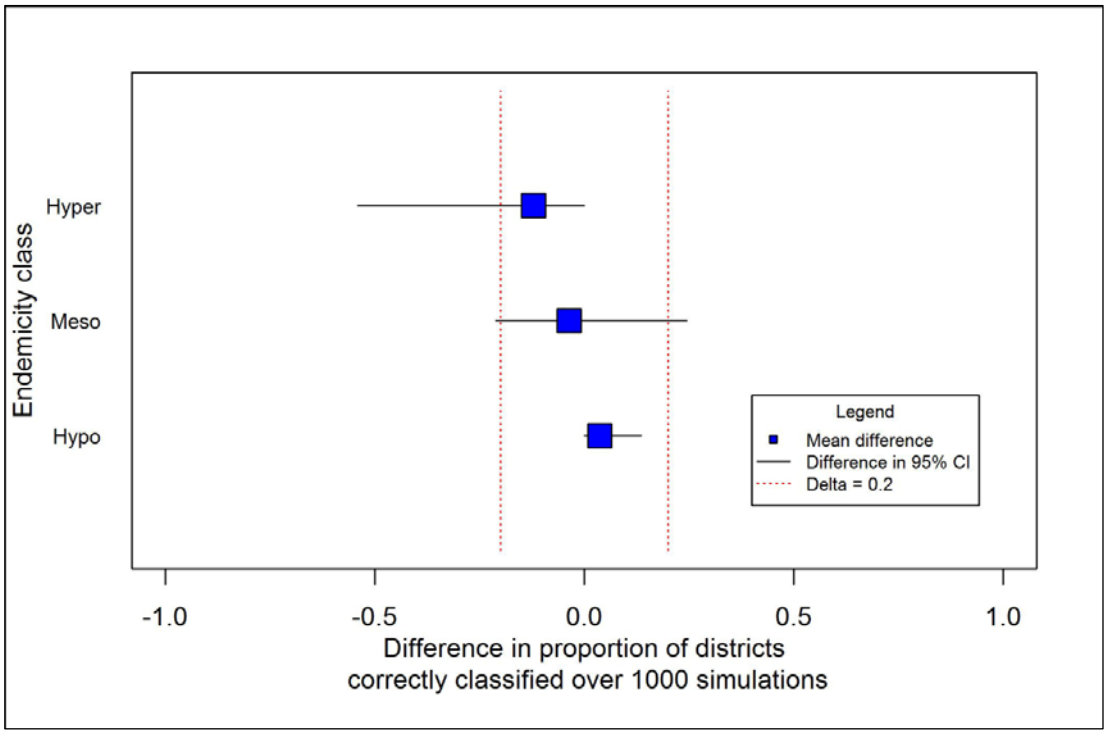



Figure 6.4 Equivalence of ITM compared to CRS by endemicity class. The figure presents the difference in the proportion of times ITM correctly classified districts compared to CRS (over 1000 simulations) by endemicity class in relation to an assumed value (20%) representing an important programmatic difference. The blue square is the mean difference in proportions and the lines correspond to the difference in the 95% CI. The two methods are deemed equivalent when ITM correctly classifies districts differently to CRS no more than 20% of the time.

### 6.4.3 Sensitivity analysis

Sensitivity analysis of the impact of varying key parameters as shown in Table 6.3 suggested that the relative risk of TF between enrolled and non-enrolled children and the enrolment rate will define the performance of ITM. This is illustrated in Figure 6.5 which plots the probability that ITM and CRS will give equivalent results in a district (i.e. the probabilities of correctly classifying a district using ITM and CRS differ no more than 0.20) given endemicity class and varying these parameters. Where enrollment is set as 0.7 and the relative risk is 0.75 or above, there is a high (≥ 80%) probability that ITM and CRS will be equivalent across all endemicity classes.

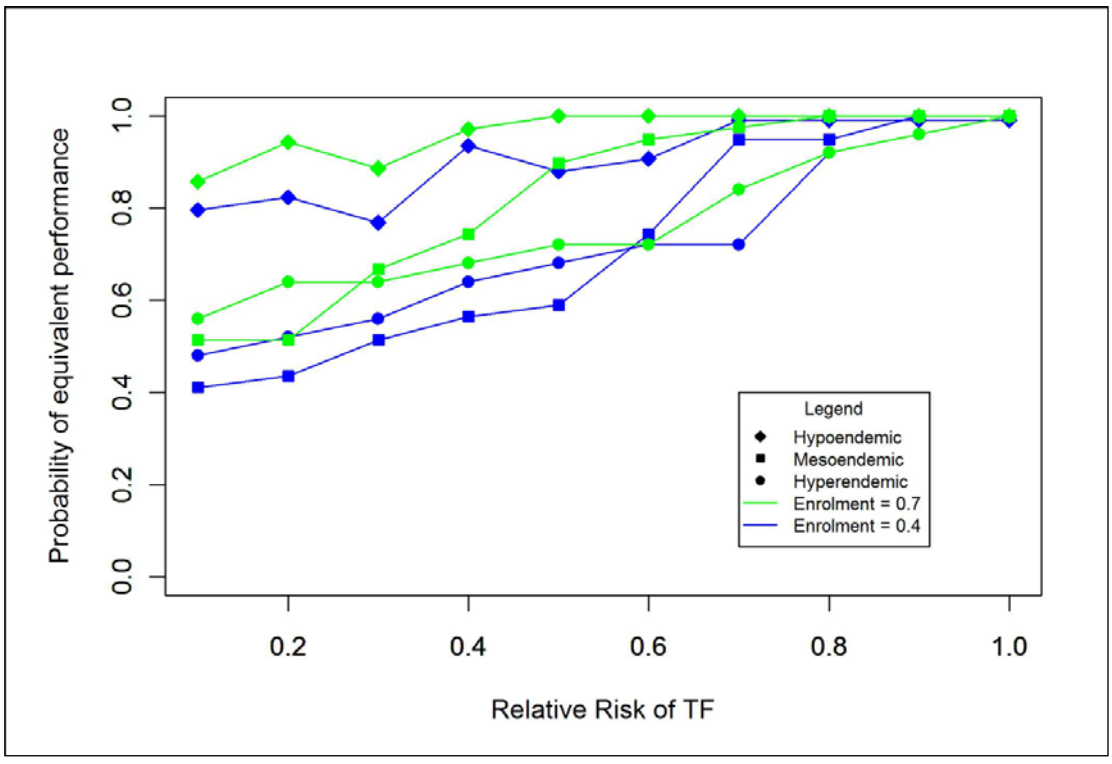



Figure 6.5 Equivalence in district classification by ITM and CRS. Equivalence is determined by calculating the difference in the probabilities that CRS and ITM will correctly classify a given district over 1000 simulations, and estimating whether this difference exceeds a delta equal to 0.2, signifying that two methods classify districts differently no more than 20% of the time. The figure presents equivalence by endemicity class and relative risk of TF in enrolled and non-enrolled children, where enrolment is equal to 0.4 (blue) or 0.7 (green).

As enrolment decreases and the difference in risk between enrolled and non-enrolled children increases, ITM increasingly misclassifies districts compared to CRS. This effect is likely to be greater in meso- and hyper- endemic districts, due to a greater magnitude of bias and resulting in misclassification over a wider range of prevalence values around the 10% and 30% thresholds. The impact on misclassification is also illustrated by Figure 6.6, which plots the range of prevalence values where the risk of misclassification using the two survey methodologies is greater or equal to 0.20. Classification error associated with CRS is symmetrically distributed approximately ±2 percent around each threshold and does not vary with these parameters. In contrast, the range of misclassification associated with ITM not only increases with a greater difference between enrolled and non-enrolled children, but also shifts to include more prevalence values above the threshold. Within this range of misclassification, the performance of ITM also decreases as a response to the degree of underestimation, so that in certain contexts ITM is unable to correctly classify any districts at or slightly above 30% prevalence. Variation in the relative risk of TF between age groups and the degree of household clustering defined by the ICC did not have an impact on performance. Evaluation of our assumption that children aged 1-5 years sampled by ITM were siblings of enrolled children also had no observable impact on the performance of ITM.

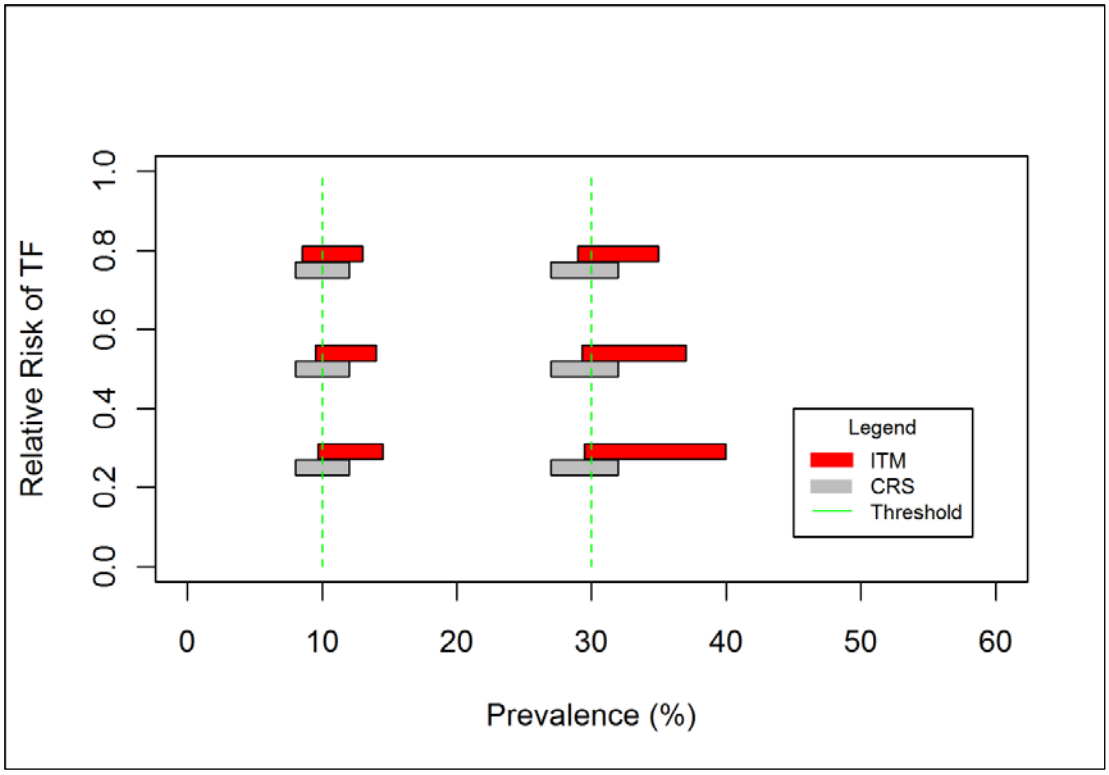



Figure 6.6 Range of true prevalence values with high risk of misclassification by CRS and ITM. Range of values in which the risk of misclassifying a district using CRS and ITM sampling methodologies is greater or equal to 0.20 around the 10% and 30% thresholds, with the enrolment rate equal to 0.7.

## 6.5 Discussion

Our simulations show that over a range of epidemiological settings, ITM will under-estimate the true prevalence of TF. The error introduced by ITM also means that districts are more prone to misclassification according to treatment thresholds than by CRS. The extent of underestimation and misclassification of districts introduced by ITM is dependent on three main factors: (i) the district prevalence of TF; (ii) the relative risk of TF between enrolled and non-enrolled children within clusters; and (iii) the enrollment rate in schools. In general, the overall agreement between the two methods is high, but as the difference in risk of TF between enrolled and non-enrolled children becomes more pronounced, there is a shift in prevalence estimates corresponding to the magnitude of the bias. In these situations, the null hypothesis of programmatic equivalence between the two methodologies is not supported.

Use of a school-based platform is a key methodological difference between CRS and ITM and, while the potential pitfalls of this approach are well recognised, the impact of this strategy on treatment decisions has not been systematically evaluated until now [85,319]. Our simulations highlighted the key influence of the relative risk of TF between enrolled and non-enrolled children and the enrollment rate on the performance of ITM. Furthermore, we were able to quantify the impact of these parameters on district classification over a range of endemicity settings. In areas where the risk of TF is similar between enrolled and non-enrolled children, there is evidence that CRS and ITM will be equivalent and classify districts correctly within an acceptable range of difference. Where risk is lower in enrolled children, a negative bias is introduced that is proportional to the magnitude of the difference in risk and reflected in greater absolute discrepancies between the two sampling methodologies as prevalence of TF increases. A lower enrolment rate effectively constrains the "sample" of the total population of children aged 6-9 attending schools and has the effect of increasing uncertainty around the prevalence estimate due to the greater effect of a positive child in the sample [320]. Compared to CRS,

where misclassification error is fairly symmetrical around treatment thresholds across all scenarios, ITM can introduce a bias-dependent right shift and widen the range of prevalence values where misclassification error is high. In contrast, varying the relative risk of TF between age groups and the average ICC did not have a noticeable impact on performance of ITM and CRS at the district level, either in magnitude or shift.

As a consequence of this potential bias, ITM may be less likely than CRS to misclassify areas as greater than 10% or 30% when the true prevalence is below this threshold, but more likely to misclassify areas as lower when the true prevalence is higher. Misclassification is more comparable between the two methodologies at the 10% threshold, particularly when the relative risk between enrolled and non-enrolled children is closer to one. At this threshold, the misclassification by ITM would result in resources being allocated for further surveys at the subdistrict level instead of implementing MDA for the entire district. In practice, the difference in performance is most likely to impact interventions around the 30% threshold, where areas misclassified by ITM would be treated for three years before an impact survey instead of being treated for five years. Districts that fall within areas of high misclassification are of operational interest and the optimal choice of survey design is likely to be a function of the cost of the surveys, the costs of treatment associated with misclassification around both thresholds and the likely impact of treatment decisions on long term transmission dynamics. For example, while a particular survey design may be a cost-effective method to classify districts at a given round, a more accurate but more expensive survey design may allow quicker elimination of the disease leading to cost-savings in the future. Incorporating costs and the impact of treatment decisions on transmission was beyond the scope of this paper, but is the focus of future study.

Our use of computerised simulation has a number of advantages over field evaluations of trachoma sampling approaches [95,96]. First, whereas inadequate evidence was available for meso- and hyperendemic settings, our approach allowed evaluation of ITM and CRS

over a range of epidemiological settings. Second, simulations allowed the two sampling methodologies to be carried out independently of one another and repeated 1000 times for each district, thus accounting for sampling error in our estimates of performance. Finally, this approach allows key parameters to be explicitly defined and varied in a sensitivity analysis in order to explore their impact on performance in different contexts. This aspect of the study is important, as these parameters are likely to vary widely in settings where ITM might be used to generate TF prevalence estimates.

Although our study explored the performance of ITM and CRS in varying contexts, there are a number of potential limitations that may limit its generalisability. First, although key factors were varied in order to test sampling strategies in different epidemiological scenarios, exploring datasets similar to the data from Kahe in Tanzania and from Kenya would allow a more realistic range of parameters to be incorporated. In addition, parameterisation of the model assumed constant relationships which may be more complex in reality. Certain factors, like household clustering of trachoma, may vary markedly based on local transmission intensity, however no clear and consistent relationship was supported by available data. This may partly be due to random error introduced by the clinical sign TF, which is known to be an unreliable marker of *C. trachomatis* infection [7,34]. A better estimation of these parameters, such as the relative risk of TF between enrolled and non-enrolled children, based on their relationship with endemicity may require collection of new data in the field. Second, these simulations sampled participants from a single demographic and household structure based on a community from Tanzania. Although the children selected as "enrolled" varied in the simulated datasets, it is possible that disease clustering within households might have a greater effect in other community structures. Furthermore, these simulations represent a general sampling scenario, and in the field there is more variation in the way that ITM and CRS are implemented. (For example, ITM randomly samples two clusters per subdistrict with a minimum of 20 per district, so the number of clusters sampled varies indirectly with district size [94]. In contrast, the number of clusters sampled by CRS is dependent on

population size and is often selected using probability proportional to size in order to estimate a reliable district-level prevalence [85].) While use of a school-based survey platform offers a number of operational advantages, it is difficult to justify this approach in many contexts. In actual practice, one might expect trachoma "hotspots" to have poorer socioeconomic conditions and lower school enrollment, thus limiting the potential use of ITM to identify disease foci. More widespread collection of indicators of enrollment and attendance as part of trachoma surveys is encouraged in order to inform survey design. In addition, there is a lack of guidance on how ITM sampling methods would be operationalised in the event of non-response from family members bringing young children to the school. If the older children were oversampled, or a smaller sample of older children accepted, then ITM would underestimate the prevalence of TF to a greater degree. Finally, both the threshold of "acceptable difference" to be used in the equivalence analysis and the thresholds themselves deserve more discussion. To some degree, treatment thresholds are imprecise as they are based on unreliable clinical indicators and the impact on transmission of misclassifying a district that has a prevalence of 9% versus 12% is not well defined. As the elimination target for active trachoma is to reduce its prevalence to less than 5% in every sub-district, the transmission dynamics around these lower thresholds is of crucial interest. The degree of acceptable difference in performance between survey designs will depend on these transmission dynamics over the course of a control programme, as well as costs associated with misclassification.

In summary, the results from this chapter strengthen the evidence base around trachoma sampling methodologies and demonstrate the advantages of using a simulated approach to evaluate different sampling scenarios. To a large extent, the results from these simulations reflect a known limitation of school-based sampling: that resulting prevalence estimates are unreliable when the enrollment is low and/or the risk of disease in schools differs from communities. However, quantification of the performance of ITM at the district level in different contexts provides important information for national control programmes. In areas where enrolment is known to be very high, and it can be reliably

inferred that the bias is minimized, then ITM may provide a rapid, cost-effective alternative to CRS [95,321]. In addition to strengthening the evidence base around trachoma sampling methodologies, the results from this analysis demonstrate the advantages of using a simulated approach to evaluate different sampling scenarios. Future work could incorporate costing of different survey approaches and extension to include mathematical modeling to simulate the impact of different combinations of control interventions on transmission [322].

# Chapter 7: Summary and discussion of findings

This thesis has used existing geographical data to define the current distribution of trachoma and spatial epidemiology of disease at large scales, with the overall aim of informing current disease mapping strategies. This broader aim was motivated by three key requirements in global efforts to eliminate blinding trachoma. First, in order to garner support amongst partners, target surveys and achieve maximum impact by 2020, there was a need to define the current distribution and burden of trachoma. Second, a better understanding of the spatial heterogeneity of trachoma and determinants underlying its distribution will add to our epidemiological understanding of the disease and may help strengthen the evidence around future survey methodologies. Finally, in order to scale up mapping in all endemic districts, the performance of survey designs proposed as an alternative to PBPS must undergo robust evaluation. The following chapters aimed to address these issues and to provide a basis for these methods in trachoma research.

## 7.1 Summary of findings

As a first step, Chapter 2 described the collection of available data used to generate the Global Atlas of Trachoma and presented summary statistics of this database along with maps of the global distribution. The results from this chapter highlighted the regional differences in data availability and endemicity patterns. In Africa alone, an estimated 129.4 million individuals are in areas known to be trachoma-endemic and a further 155 million in areas suspected to be endemic. In addition to highlighting the widespread endemicity and high burden in Africa, compared to more focal distribution in Asian and Latin American countries, the results emphasised the substantial uncertainty in estimates introduced by a lack of data in India and China. This heterogeneous global distribution of

both data and disease presents challenges in filling in the gaps using appropriate survey methodologies, and will also have an important impact on projected estimates of the burden of trachoma which are typically based on limited data and heavily extrapolated within geographic regions.

Chapter 3 built on the data resource presented in Chapter 2, and presented a framework for using collated data on the prevalence of trichiasis to model trachomatous vision loss where causal blindness data are lacking. DALYs are increasingly used to provide a comparable measure to other diseases, track changes in burden over time and justify requests for and allocate of resources against competing priorities. The results from this chapter emphasised the added burden associated with trichiasis; which added a further 155,500 DALYs to the 2010 estimates. While there are a number of key advantages associated with the methodology used in this chapter, the sensitivity analysis highlighted a need to explore alternative extrapolation methods to address gaps in geographical coverage of existing data.

This need was explicitly addressed in Chapter 4, which explored individual and cluster-level risk factors underlying the distribution of trichiasis and/or corneal opacity in Nigeria. This analysis demonstrated associations between a number of well-established individual risk factors (age, gender, occupation) and TT/CO, as well as large scale environmental factors. Although these associations may indicate that socioeconomic and environmental factors could be used to extrapolate existing data, the variance partition coefficient suggested that a relatively low proportion of the total variation was at higher hierarchical levels (although this is likely to vary locally). After adjusting for these factors, there was evidence of a large, local cluster of risk in the north but no residual autocorrelation in the south; highlighting variation in spatial dependency across Nigeria.

A similar risk analyses was carried out with cluster data on active trachoma in Chapter 5, representing survey data from all 17 districts currently classified as endemic in Kenya. Results showed that large scale deterministic trends were explained by risk factors at the

district and cluster level. Based on the higher risk observed in less populated areas and local clustering within districts, the use of PPS for trachoma surveys might be questioned in the context of elimination targets. While PPS does provide prevalence estimates representative of the district population, there is a risk of missing subdistricts with a higher prevalence. Although many survey designs create "trachoma districts" based on populations of 100,000-200,000, this is not always the case. Districts with a high population and resulting poor geographical coverage of rural areas may miss subdistricts of higher prevalence; possibly suggesting use of a stratified approach to ensure adequate geographical coverage at this level.

In this chapter, a group of spatially-varying covariates explained large-scale spatial patterns in active trachoma in Kenya, leaving independent errors. However, if these model covariates were not included, there might still be larger scale autocorrelation that can bias regression parameter estimates and cause standard errors to be underestimates, potentially leading to incorrect inferences regarding exposure-disease associations [323]. In larger scale analyses of trachoma risk, the potential introduction of geographical trends introduced by autocorrelation from environmental factors is usually ignored as regression methods commonly used for these analyses cannot account for these spatial relationships. The results from this thesis suggest that larger scale autocorrelation should be evaluated and corrected for where necessary in standard risk factor analyses.

Collectively, the results in Chapters 4 and 5 support the notion that the risk of trachoma exhibits marked spatial heterogeneity and observed patterns of spatial dependency (likely generated by underlying associations with risk factors at larger scales) may vary between endemic areas. These results emphasised the need for robust and well designed survey methodologies to identify areas of high risk. This is particularly important in the current context of rapid scale up in mapping activities. Therefore, Chapter 6 systematically evaluated the use of ITM compared to CRS using computerised simulations, allowing quantification of the impact of varying key parameters. Results suggested that ITM was

likely to underestimate the prevalence of trachoma in a range of epidemiological settings, reflecting the unreliability of school-based sampling when the attendance is low and/or the risk of disease in schools differs from communities. While ITM may provide a cost-effective means of surveying target populations in settings where school attendance is high, generalisation of school attendance rates within large areas may be unreliable. Consistent with results from Chapter 3 that suggest that variation in access to education was a key determinant of large-scale trends, a school-based platform may be particularly problematic for trachoma given that non-attendance is likely to cluster in those communities that have fewer schools or are located further from schools.

## 7.2 Future directions

The notion of trachoma as a highly heterogeneous disease is supported by the findings from this thesis, which has highlighted variability in risk and underlying determinants at different scales. In addition to the requirement for reliable survey methods to guide local intervention strategies, the results have demonstrated how an understanding of the epidemiology and spatial structure of disease in different contexts can inform evaluations of alternative survey methodologies. There are numerous ways in which these methods can be refined and potential directions that future analyses might take to further inform survey strategies over the course of a control programme.

District level estimates of trichiasis provide a wider evidence base to estimate the global burden of disease; however, results presented in Chapter 3 continue to be limited by scarcity of data in time and space. Although the Global Trachoma Mapping Project will fill gaps in baseline data, the (variable) success of control programmes currently in place may introduce further uncertainty into relationships between TT and trachomatous blindness. Future research might i) better quantify how the relationship between TT and trachomatous blindness varies in different settings, ii) investigate the impact of control on

the age-prevalence distributions of trichiasis and trachomatous blindness, iii) explore different methods of extrapolation in space, and iv) investigate whether mathematical models might be used to better estimate the changing burden of trachomatous blindness, using data on TF and TT as well as country-specific mortality rates. Over these large scales, stochastic variation in disease prevalence is expected to be less and thus, district level socioeconomic or environmental data may be useful in capturing large-scale trends.

While country and regional estimates of the burden of disease are the key objectives for large-scale priority setting, sub-national estimates of trachoma are directly relevant to control programmes. Thus, although the Global Atlas of Trachoma is a useful tool for partners, this thesis illustrates the extent to which aggregated data belie the small scale heterogeneity of disease and mask important variation within districts. From an operational research perspective, disaggregated data are far more valuable for investigation of spatial heterogeneity and patterns of risk. Spatial data at multiple scales are increasingly available through the use of smartphones for data collection. These data may help to define important risk factors underlying the distribution of disease and have a number of implications for routine risk factor analyses and design of surveys.

While variation in risk of trachoma was observed to be spatially structured over a 200-300 km range, current survey methodologies sample at a fine enough resolution (30 cluster per district) to capture variation within this range. However, the results from this thesis indicate that the epidemiology and spatial structure of trachoma vary in space; suggesting that both survey methods and potential targeting of surveillance should be context specific. Future operational research should focus on defining a framework for targeted surveillance that may incorporate some of the techniques introduced in this thesis. As endemicity decreases, trachoma may become more focal and further study should be targeted to local clusters of risk within districts in order to better understand i) important local risk factors (i.e. is there more travel between these areas that might pose a future

risk of reintroduction; what are the shared risk factors driving high risk in this locality) and ii) better target specific interventions (i.e. water availability and behavioural change).

Evaluation of survey designs in different contexts will become increasingly important in the context of elimination, where there is more pressure to integrate surveillance of trachoma with other diseases and verify the absence of disease. Smartphone data collection has allowed the rapid proliferation of spatial data on disease presence and risk factors at different scales, which may offer an opportunity to better parameterise models and conduct "context specific" sampling simulations. Included parameters may include observed spatial heterogeneity between and within districts, clustering by household and various risk factors, and varying endemicity; all of which will influence the optimal spatial resolution of surveys and the sampling effort required. In addition, the likely survey and treatment costs can be incorporated into sampling simulations, to compare the cost-effectiveness of different survey designs [100,101].

Furthermore, survey simulations could be integrated with mathematical intervention models for two key extensions. First, mathematical models could provide a post-intervention gold standard upon which to conduct sampling simulations where these data are lacking. Second, models could be used to explore the implications of the performance of different survey designs on long term control outcomes [324]. For example an initial survey step could be incorporated into intervention models in order to allow an exploration of the implications of the performance of different survey designs on control decisions and ultimately on disease transmission.

In conclusion, this thesis provides the first systematic and detailed investigation into the spatial epidemiology of trachoma at different scales. The findings have shown that while the distribution of trachoma is associated with a number of spatially-varying risk factors, it is highly heterogeneous over multiple scales. Thus, there remains a substantial need to ensure that survey designs generate reliable data to allow targeting of interventions at appropriate scales over the course of a control programme. The optimal design of

trachoma surveys and targeting of interventions will be increasingly important as programmes are scaled up to meet the challenge of elimination.

# References


1. Anselin L (1995) Local Indicators of Spatial Association—LISA. Geographical Analysis 27: 93-115.

2. Resnikoff S, Pascolini D, Etya'ale D, Kocur I, Pararajasegaram R, et al. (2004) Global data on visual impairment in the year 2002. Bull World Health Organ 82: 844-851.

3. World Health Organization. Report of the fifteenth meeting of the WHO alliance for the Global Elimination of Blinding Trachoma by 2020; 2011; Geneva. WHO.

4. Treharne JD (1985) The community epidemiology of trachoma. Rev Infect Dis 7: 760-764.

5. Hu VH, Harding-Esch EM, Burton MJ, Bailey RL, Kadimpeul J, et al. (2010) Epidemiology and control of trachoma: systematic review. Trop Med Int Health 15: 673-691.

6. Cook J, S M (2011) Trachoma. In: Selendy J, editor. Water and Sanitation-Related Diseases and the Environment: Challenges, Interventions and Preventive Measures: Wiley-Blackwell.

7. Solomon AW, Peeling RW, Foster A, Mabey DC (2004) Diagnosis and assessment of trachoma. Clin Microbiol Rev 17: 982-1011.

8. Bailey R, Duong T, Carpenter R, Whittle H, Mabey D (1999) The duration of human ocular Chlamydia trachomatis infection is age dependent. Epidemiol Infect 123: 479-486.

9. Hu VH, Holland MJ, Burton MJ (2013) Trachoma: protective and pathogenic ocular immune responses to Chlamydia trachomatis. PLoS Negl Trop Dis 7: e2020.

10. Wolle MA, Munoz BE, Mkocha H, West SK (2009) Constant ocular infection with Chlamydia trachomatis predicts risk of scarring in children in Tanzania. Ophthalmology 116: 243-247.

11. Burton MJ, Holland MJ, Faal N, Aryee EA, Alexander ND, et al. (2003) Which members of a community need antibiotics to control trachoma? Conjunctival Chlamydia trachomatis infection load in Gambian villages. Invest Ophthalmol Vis Sci 44: 4215-4222.

12. Bobo LD, Novak N, Munoz B, Hsieh YH, Quinn TC, et al. (1997) Severe disease in children with trachoma is associated with persistent Chlamydia trachomatis infection. J Infect Dis 176: 1524-1530.

13. Solomon AW, Holland MJ, Burton MJ, West SK, Alexander ND, et al. (2003) Strategies for control of trachoma: observational study with quantitative PCR. Lancet 362: 198-204.

14. Burr SE, Hart JD, Edwards T, Baldeh I, Bojang E, et al. (2013) Association between ocular bacterial carriage and follicular trachoma following mass azithromycin distribution in The Gambia. PLoS Negl Trop Dis 7: e2347.

15. Burton MJ, Hu VH, Massae P, Burr SE, Chevallier C, et al. (2011) What is causing active trachoma? The role of nonchlamydial bacterial pathogens in a low prevalence setting. Invest Ophthalmol Vis Sci 52: 6012-6017.

16. Wolle MA, Munoz B, Mkocha H, West SK (2009) Age, sex, and cohort effects in a longitudinal study of trachomatous scarring. Invest Ophthalmol Vis Sci 50: 592-596.

17. Munoz B, Bobo L, Mkocha H, Lynch M, Hsieh YH, et al. (1999) Incidence of trichiasis in a cohort of women with and without scarring. Int J Epidemiol 28: 1167-1171.

18. Bowman RJ, Faal H, Myatt M, Adegbola R, Foster A, et al. (2002) Longitudinal study of trachomatous trichiasis in the Gambia. Br J Ophthalmol 86: 339-343.

19. Burton MJ, Bowman RJ, Faal H, Aryee EA, Ikumapayi UN, et al. (2006) The long-term natural history of trachomatous trichiasis in the Gambia. Invest Ophthalmol Vis Sci 47: 847-852.

20. West SK, Munoz B, Mkocha H, Hsieh YH, Lynch MC (2001) Progression of active trachoma to scarring in a cohort of Tanzanian children. Ophthalmic Epidemiol 8: 137-144.

21. Beatty WL, Byrne GI, Morrison RP (1994) Repeated and persistent infection with Chlamydia and the development of chronic inflammation and disease. Trends Microbiol 2: 94-98.

22. Wright HR, Turner A, Taylor HR (2008) Trachoma. Lancet 371: 1945-1954.

23. Hu VH, Weiss HA, Ramadhani AM, Tolbert SB, Massae P, et al. (2012) Innate immune responses and modified extracellular matrix regulation characterize bacterial infection and cellular/connective tissue changes in scarring trachoma. Infect Immun 80: 121-130.

24. Burton MJ, Adegbola RA, Kinteh F, Ikumapayi UN, Foster A, et al. (2007) Bacterial infection and trachoma in the gambia: a case control study. Invest Ophthalmol Vis Sci 48: 4440-4444.

25. Hu VH, Massae P, Weiss HA, Chevallier C, Onyango JJ, et al. (2011) Bacterial infection in scarring trachoma. Invest Ophthalmol Vis Sci 52: 2181-2186.

26. Thylefors B, Dawson CR, Jones BR, West SK, Taylor HR (1987) A simple system for the assessment of trachoma and its complications. Bull World Health Organ 65: 477-483.

27. MacCallan AF (1931) The Epidemiology of Trachoma. Br J Ophthalmol 15: 369-411.

28. Dawson CR, Hoshiwara I, Daghfous T, Messadi M, Vastine DW, et al. (1975) Topical tetracycline and rifampicin therapy of endemic trachoma in Tunisia. Am J Ophthalmol 79: 803-811.

29. Miller K, Schmidt G, Melese M, Alemayehu W, Yi E, et al. (2004) How reliable is the clinical exam in detecting ocular chlamydial infection? Ophthalmic Epidemiol 11: 255-262.

30. Bailey RL, Hampton TJ, Hayes LJ, Ward ME, Whittle HC, et al. (1994) Polymerase chain reaction for the detection of ocular chlamydial infection in trachoma-endemic communities. J Infect Dis 170: 709-712.

31. Keenan JD, See CW, Moncada J, Ayele B, Gebre T, et al. (2012) Diagnostic characteristics of tests for ocular Chlamydia after mass azithromycin distributions. Invest Ophthalmol Vis Sci 53: 235-240.

32. Wright HR, Taylor HR (2005) Clinical examination and laboratory tests for estimation of trachoma prevalence in a remote setting: what are they really telling us? Lancet Infect Dis 5: 313-320.

33. Baral K, Osaki S, Shreshta B, Panta CR, Boulter A, et al. (1999) Reliability of clinical diagnosis in identifying infectious trachoma in a low-prevalence area of Nepal. Bull World Health Organ 77: 461-466.

34. See CW, Alemayehu W, Melese M, Zhou Z, Porco TC, et al. (2011) How reliable are tests for trachoma?--a latent class approach. Invest Ophthalmol Vis Sci 52: 6133-6137.

35. Bhosai SJ, Bailey RL, Gaynor BD, Lietman TM (2012) Trachoma: an update on prevention, diagnosis, and treatment. Curr Opin Ophthalmol 23: 288-295.

36. Jenson A, Dize L, Mkocha H, Munoz B, Lee J, et al. (2013) Field evaluation of the Cepheid GeneXpert Chlamydia trachomatis assay for detection of infection in a trachoma endemic community in Tanzania. PLoS Negl Trop Dis 7: e2265.

37. Kuper H, Solomon AW, Buchan J, Zondervan M, Foster A, et al. (2003) A critical review of the SAFE strategy for the prevention of blinding trachoma. The Lancet Infectious Diseases 3: 372-381.

38. Courtright P (1994) Acceptance of surgery for trichiasis among rural Malawian women. East Afr Med J 71: 803-804.

39. West S, Lynch M, Munoz B, Katala S, Tobin S, et al. (1994) Predicting surgical compliance in a cohort of women with trichiasis. Int Ophthalmol 18: 105-109.

40. West S (2012) Trachoma and antibiotic use: the 'A' in SAFE. Expert Review of Anti-Infective Therapy 10: 75-83.

41. House JI, Ayele B, Porco TC, Zhou Z, Hong KC, et al. (2009) Assessment of herd protection against trachoma due to repeated mass antibiotic distributions: a cluster-randomised trial. Lancet 373: 1111-1118.

42. Evans JR, Solomon AW (2011) Antibiotics for trachoma. Cochrane Database Syst Rev: CD001860.

43. Fraser-Hurt N, Bailey RL, Cousens S, Mabey D, Faal H, et al. (2001) Efficacy of oral azithromycin versus topical tetracycline in mass treatment of endemic trachoma. Bull World Health Organ 79: 632-640.

44. Bowman RJ, Sillah A, Van Dehn C, Goode VM, Muqit MM, et al. (2000) Operational comparison of single-dose azithromycin and topical tetracycline for trachoma. Invest Ophthalmol Vis Sci 41: 4074-4079.

45. Malaty R, Zaki S, Said ME, Vastine DW, Dawson DW, et al. (1981) Extraocular infections in children in areas with endemic trachoma. J Infect Dis 143: 853.

46. Porco TC, Gebre T, Ayele B, House J, Keenan J, et al. (2009) Effect of mass distribution of azithromycin for trachoma control on overall mortality in Ethiopian children: a randomized trial. JAMA 302: 962-968.

47. Biebesheimer JB, House J, Hong KC, Lakew T, Alemayehu W, et al. (2009) Complete local elimination of infectious trachoma from severely affected communities after six biannual mass azithromycin distributions. Ophthalmology 116: 2047-2050.

48. Gaynor BD, Miao Y, Cevallos V, Jha H, Chaudary JS, et al. (2003) Eliminating trachoma in areas with limited disease. Emerg Infect Dis 9: 596-598.

49. Burton MJ, Holland MJ, Makalo P, Aryee EA, Alexander ND, et al. (2005) Re-emergence of Chlamydia trachomatis infection after mass antibiotic treatment of a trachoma-endemic Gambian community: a longitudinal study. Lancet 365: 1321-1328.

50. Shah NA, House J, Lakew T, Alemayehu W, Halfpenny C, et al. (2010) Travel and implications for the elimination of trachoma in Ethiopia. Ophthalmic Epidemiol 17: 113-117.

51. Ayele B, Gebre T, Moncada J, House JI, Stoller NE, et al. (2011) Risk factors for ocular chlamydia after three mass azithromycin distributions. PLoS Negl Trop Dis 5: e1441.

52. West SK, Munoz B, Mkocha H, Holland MJ, Aguirre A, et al. (2005) Infection with Chlamydia trachomatis after mass treatment of a trachoma hyperendemic community in Tanzania: a longitudinal study. Lancet 366: 1296-1300.

53. Edwards T, Cumberland P, Hailu G, Todd J (2006) Impact of health education on active trachoma in hyperendemic rural communities in Ethiopia. Ophthalmology 113: 548-555.

54. Emerson PM, Cairncross S, Bailey RL, Mabey DC (2000) Review of the evidence base for the 'F' and 'E' components of the SAFE strategy for trachoma control. Trop Med Int Health 5: 515-527.

55. Khandekar R, Ton TK, Do Thi P (2006) Impact of face washing and environmental improvement on reduction of active trachoma in Vietnam-a public health intervention study. Ophthalmic Epidemiol 13: 43-52.

56. Solomon AW, Zondervan M, Kuper H, Buchan J, Mabey DC, et al. (2006) Trachoma control - a guide for programme managers. Switzerland: World Health Organization.

57. West SK, Bailey R, Munoz B, Edwards T, Mkocha H, et al. (2013) A randomized trial of two coverage targets for mass treatment with azithromycin for trachoma. PLoS Negl Trop Dis 7: e2415.

58. Gebre T, Ayele B, Zerihun M, Genet A, Stoller NE, et al. (2012) Comparison of annual versus twice-yearly mass azithromycin treatment for hyperendemic trachoma in Ethiopia: a cluster-randomised trial. Lancet 379: 143-151.

59. Harding-Esch EM, Edwards T, Sillah A, Sarr-Sissoho I, Aryee EA, et al. (2008) Risk factors for active trachoma in The Gambia. Trans R Soc Trop Med Hyg 102: 1255-1262.

60. Ray KJ, Lietman TM, Porco TC, Keenan JD, Bailey RL, et al. (2009) When can antibiotic treatments for trachoma be discontinued? Graduating communities in three African countries. PLoS Negl Trop Dis 3: e458.

61. Trompoukis C, Kourkoutas D (2007) Trachoma in late Greek antiquity and the early Byzantine periods. Can J Ophthalmol 42: 870-874.

62. Tower P (1963) The history of trachoma: its military and sociological implications. Arch Ophthalmol 69: 123-130.

63. Meyerhof M (1932) A short history of ophthalmia during the Egyptian campaigns of 1798-1807. Br J Ophthalmol 16: 129-152.

64. Jones BR (1980) Changing concepts of trachoma and its control. Trans Ophthalmol Soc UK 100: 25-29.

65. Allen SK, Semba RD (2002) The trachoma menace in the United States, 1897-1960. Surv Ophthalmol 47: 500-509.

66. World Health Organization (2012) Global WHO Alliance for the Elimination of Blinding Trachoma by 2020. WHO Weekly epidemiological record 17: 161–168.

67. Wiabut F (1929) Historical map of worldwide trachoma distribution. International Council of Ophthalmology, Library of the Nuffield Laboratory of Ophthalmology, University of Oxford.

68. Sidky M, Freyce M (1949) World distribution and prevalence of trachoma in recent years. Epidem Vital Stat Rep II: 230-277.

69. Siebeck R (1961) The Global Distribution of Trachoma 1930–1955. In: Rodenwaldt E, Justatz H, editors. World atlas of Epidemic Diseases Part III. Hamburg: Falk-Verlag.

70. Polack S, Brooker S, Kuper H, Mariotti S, Mabey D, et al. (2005) Mapping the global distribution of trachoma. Bull World Health Organ 83: 913-919.

71. Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, et al. (2013) Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380: 2197-2223.

72. Salomon JA, Vos T, Murray CJ (2013) Disability weights for vision disorders in Global Burden of Disease study - Authors' reply. Lancet 381: 23-24.

73. Murray CJ, Acharya AK (1997) Understanding DALYs (disability-adjusted life years). J Health Econ 16: 703-730.

74. Kuper H, Polack S, Limburg H (2006) Rapid assessment of avoidable blindness. Community Eye Health 19: 68-69.

75. World Health Organization (2006) International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10), Chapter VII. H54. Blindness and low vision

76. Dandona L, Dandona R (2006) Revision of visual impairment definitions in the International Statistical Classification of Diseases. BMC Med 4: 7.

77. World Health Organization (1988) Coding instructions for the WHO/PBL eye examination record (Version III).

78. Foster A (1987) Cataract blindness in Africa. Ophthalmic Surg 18: 384-388.

79. Ranson MK, Evans TG (1995) The global burden of trachomatous visual impairment: I. Assessing prevalence. Int Ophthalmol 19: 261-270.

80. Frick KD, Hanson CL, Jacobson GA (2003) Global burden of trachoma and economics of the disease. Am J Trop Med Hyg 69: 1-10.

81. Thylefors B, Ranson K, Negrel AD, Pararajasegaram R (2004) Trachoma-related visual loss. In: Murray CJL, Lopez A, editors. The Global Epidemiology of Infectious Diseases: Harvard University Press. pp. pp. 301-324.

82. Shibuya K, Mariotti S, Mathers CD (2006) Global burden of blinding trachoma in the year 2000: Summary of methods and data sources. WHO.

83. Frick KD, Melia BM, Buhrmann RR, West SK (2001) Trichiasis and disability in a trachoma-endemic area of Tanzania. Arch Ophthalmol 119: 1839-1844.

84. Woreta TA, Munoz BE, Gower EW, Alemayehu W, West SK (2009) Effect of trichiasis surgery on visual acuity outcomes in Ethiopia. Arch Ophthalmol 127: 1505-1510.

85. Ngondi J, Reacher M, Matthews F, Brayne C, Emerson P (2009) Trachoma survey methods: a literature review. Bull World Health Organ 87: 143-151.

86. Karimurio J (2012) The “segment knockout” survey method for large trachoma-endemic districts. Melbourne: University of Melbourne.

87. Karimurio J, Rono H, Le Mesurier R, Mwanthi M, Keeffe J (2011) What is the appropriate age range of individuals to be included in a survey to estimate the prevalence of trachomatous trichiasis? Br J Ophthalmol 95: 1058-1060.

88. Negrel AD, Mariotti S (1999) Trachoma rapid assessment: rationale and basic principles Community Eye Health 12: 51-53.

89. Limburg H, Bah M, Johnson GJ (2001) Trial of the Trachoma Rapid Assessment methodology in The Gambia. Ophthalmic Epidemiol 8: 73-85.

90. Rabiu MM, Alhassan MB, Abiose A (2001) Trial of Trachoma Rapid Assessment in a subdistrict of northern Nigeria. Ophthalmic Epidemiol 8: 263-272.

91. Paxton A (2001) Rapid assessment of trachoma prevalence--Singida, Tanzania. A study to compare assessment methods. Ophthalmic Epidemiol 8: 87-96.

92. Liu H, Ou B, Paxton A, Zhao P, Xu J, et al. (2002) Rapid assessment of trachoma in Hainan Province, China: Validation of the new World Health Organization methodology. Ophthalmic Epidemiol 9: 97-104.

93. Myatt M, Mai NP, Quynh NQ, Nga NH, Tai HH, et al. (2005) Using lot quality-assurance sampling and area sampling to identify priority areas for trachoma control: Viet Nam. Bull World Health Organ 83: 756-763.

94. Mathieu E (2011) Field Manual for Integrated NTD Surveys. Atlanta, GA: Centers for Disease Control and Prevention,.

95. Pelletreau S, Nyaku M, Dembele M, Sarr B, Budge P, et al. (2011) The field-testing of a novel integrated mapping protocol for neglected tropical diseases. PLoS Negl Trop Dis 5: e1380.

96. Dorkenoo AM, Bronzan RN, Ayena KD, Anthony G, Agbo YM, et al. (2012) Nationwide integrated mapping of three neglected tropical diseases in Togo: countrywide implementation of a novel approach. Trop Med Int Health 17: 896-903.

97. Courtright P, Sheppard J, Lane S, Sadek A, Schachter J, et al. (1991) Latrine ownership as a protective factor in inflammatory trachoma in Egypt. Br J Ophthalmol 75: 322-325.

98. Tavornpanich S, Gardner IA, Carpenter TE, Johnson WO, Anderson RJ (2006) Evaluation of cost-effectiveness of targeted sampling methods for detection of Mycobacterium avium subsp paratuberculosis infection in dairy herds. Am J Vet Res 67: 821-828.

99. Olives C, Valadez JJ, Brooker SJ, Pagano M (2012) Multiple category-lot quality assurance sampling: a new classification system with application to schistosomiasis control. PLoS Negl Trop Dis 6: e1806.

100. Sturrock HJ, Gething PW, Clements AC, Brooker S (2010) Optimal survey designs for targeting chemotherapy against soil-transmitted helminths: effect of spatial heterogeneity and cost-efficiency of sampling. Am J Trop Med Hyg 82: 1079-1087.

101. Sturrock HJW, Gething PW, Ashton RA, Kolaczinski JH, Kabatereine NB, et al. (2011) Planning schistosomiasis control: investigation of alternative sampling strategies for Schistosoma mansoni to target mass drug administration of praziquantel in East Africa. International Health 3: 165-175.

102. Levin SA (1992) The Problem of Pattern and Scale in Ecology: The Robert H. MacArthur Award Lecture. Ecology 73: 1943-1967.

103. Pfeiffer DU, Robinson TP, Stevenson M, Stevens KB, Rogers DJ, et al. (2008) Spatial Analysis in Epidemiology. New York: Oxford University Press.

104. Harding-Esch EM, Sillah A, Edwards T, Burr SE, Hart JD, et al. (2013) Mass Treatment with Azithromycin for Trachoma: When Is One Round Enough? Results from the PRET Trial in The Gambia. Plos Neglected Tropical Diseases 7.

105. Barenfanger J (1975) Studies on the role of the famiy unit in the transmission of trachoma. Am J Trop Med Hyg 24: 509-515.

106. Bailey R, Osmond C, Mabey DC, Whittle HC, Ward ME (1989) Analysis of the household distribution of trachoma in a Gambian village using a Monte Carlo simulation procedure. Int J Epidemiol 18: 944-951.

107. Mabey DC, Bailey RL, Ward ME, Whittle HC (1992) A longitudinal study of trachoma in a Gambian village: implications concerning the pathogenesis of chlamydial infection. Epidemiol Infect 108: 343-351.

108. Katz J, Zeger SL, Tielsch JM (1988) Village and household clustering of xerophthalmia and trachoma. Int J Epidemiol 17: 865-869.

109. Blake IM, Burton MJ, Bailey RL, Solomon AW, West S, et al. (2009) Estimating household and community transmission of ocular Chlamydia trachomatis. PLoS Negl Trop Dis 3: e401.

110. Grassly NC, Ward ME, Ferris S, Mabey DC, Bailey RL (2008) The natural history of trachoma infection and disease in a Gambian cohort with frequent follow-up. PLoS Negl Trop Dis 2: e341.

111. Bowman RJ, Jatta B, Cham B, Bailey RL, Faal H, et al. (2001) Natural history of trachomatous scarring in The Gambia: results of a 12-year longitudinal follow-up. Ophthalmology 108: 2219-2224.

112. Taylor HR, Siler JA, Mkocha HA, Munoz B, West S (1992) The natural history of endemic trachoma: a longitudinal study. Am J Trop Med Hyg 46: 552-559.

113. Cromwell EA, Courtright P, King JD, Rotondo LA, Ngondi J, et al. (2009) The excess burden of trachomatous trichiasis in women: a systematic review and meta-analysis. Trans R Soc Trop Med Hyg 103: 985-992.

114. Kupka K, Nizetic B, Reinhards J (1968) Sampling studies on the epidemiology and control of trachoma in southern Morocco. Bull World Health Organ 39: 547-566.

115. Golovaty I, Jones L, Gelaye B, Tilahun M, Belete H, et al. (2009) Access to water source, latrine facilities and other risk factors of active trachoma in Ankober, Ethiopia. PLoS One 4: e6702.

116. Schemann JF, Sacko D, Malvy D, Momo G, Traore L, et al. (2002) Risk factors for trachoma in Mali. Int J Epidemiol 31: 194-201.

117. Abdou A, Nassirou B, Kadri B, Moussa F, Munoz BE, et al. (2007) Prevalence and risk factors for trachoma and ocular Chlamydia trachomatis infection in Niger. Br J Ophthalmol 91: 13-17.

118. Ngondi J, Matthews F, Reacher M, Onsarigo A, Matende I, et al. (2007) Prevalence of risk factors and severity of active trachoma in southern Sudan: an ordinal analysis. Am J Trop Med Hyg 77: 126-132.

119. Taylor HR, West SK, Mmbaga BB, Katala SJ, Turner V, et al. (1989) Hygiene factors and increased risk of trachoma in central Tanzania. Arch Ophthalmol 107: 1821-1825.

120. Taylor HR, Velasco FM, Sommer A (1985) The ecology of trachoma: an epidemiological study in southern Mexico. Bull World Health Organ 63: 559-567.

121. Ketema K, Tiruneh M, Woldeyohannes D, Muluye D (2012) Active trachoma and associated risk factors among children in Baso Liben District of East Gojjam, Ethiopia. BMC Public Health 12: 1105.

122. Marshall CL (1968) The relationship between trachoma and piped water in a developing area. Arch Environ Health 17: 215-220.

123. West S, Lynch M, Turner V, Munoz B, Rapoza P, et al. (1989) Water availability and trachoma. Bull World Health Organ 67: 71-75.

124. Amza A, Kadri B, Nassirou B, Stoller NE, Yu SN, et al. (2012) Community risk factors for ocular Chlamydia infection in Niger: pre-treatment results from a cluster-randomized trachoma trial. PLoS Negl Trop Dis 6: e1586.

125. Bailey R, Downes B, Downes R, Mabey D (1991) Trachoma and water use; a case control study in a Gambian village. Trans R Soc Trop Med Hyg 85: 824-828.

126. Polack S, Kuper H, Solomon AW, Massae PA, Abuelo C, et al. (2006) The relationship between prevalence of active trachoma, water availability and its use in a Tanzanian village. Trans R Soc Trop Med Hyg 100: 1075-1083.

127. Edwards T, Harding-Esch EM, Hailu G, Andreason A, Mabey DC, et al. (2008) Risk factors for active trachoma and Chlamydia trachomatis infection in rural Ethiopia after mass treatment with azithromycin. Trop Med Int Health 13: 556-565.

128. Cumberland P, Hailu G, Todd J (2005) Active trachoma in children aged three to nine years in rural communities in Ethiopia: prevalence, indicators and risk factors. Trans R Soc Trop Med Hyg 99: 120-127.

129. White GF, Bradley DJ, White AU (2002) Drawers of water: domestic water use in East Africa. Bull World Health Organ 80: 63-73.

130. Rog M, Swenor B, Cajas-Monson LC, McHiwe W, Kiboko S, et al. (2011) A cross-sectional survey of water and clean faces in trachoma endemic communities in Tanzania. BMC Public Health 11: 495.

131. McCauley AP, Lynch M, Pounds MB, West S (1990) Changing water-use patterns in a water-poor area: lessons for a trachoma intervention project. Soc Sci Med 31: 1233-1238.

132. Cairncross S (1999) Trachoma and water. Community Eye Health 12: 58-59.

133. Lynch M, West SK, Munoz B, Kayongoya A, Taylor HR, et al. (1994) Testing a participatory strategy to change hygiene behaviour: face washing in central Tanzania. Trans R Soc Trop Med Hyg 88: 513-517.

134. West S, Munoz B, Lynch M, Kayongoya A, Chilangwa Z, et al. (1995) Impact of face-washing on trachoma in Kongwa, Tanzania. Lancet 345: 155-158.

135. Hafez M (1958) Studies on the ecology of *Musca Sorbens* wied. in Egypt. Bulletin of the Entomological Society of Egyp XLII: 83-121.

136. Taye A, Alemayehu W, Melese M, Geyid A, Mekonnen Y, et al. (2007) Seasonal and altitudinal variations in fly density and their association with the occurrence of trachoma, in the Gurage zone of central Ethiopia. Ann Trop Med Parasitol 101: 441-448.

137. Reinhards J, Weber A, Nizetic B, Kupka K, Maxwell-Lyons F (1968) Studies in the epidemiology and control of seasonal conjunctivitis and trachoma in southern Morocco. Bull World Health Organ 39: 497-545.

138. West SK, Rapoza P, Munoz B, Katala S, Taylor HR (1991) Epidemiology of ocular chlamydial infection in a trachoma-hyperendemic area. J Infect Dis 163: 752-756.

139. Hafez M (1958) The relation of *Musca Sorbens* wied. to eye diseases, in Egypt. Bulletin of the Entomological Society of Egyp XLII: 275-283.

140. Brechner RJ, West S, Lynch M (1992) Trachoma and flies. Individual vs environmental risk factors. Arch Ophthalmol 110: 687-689.

141. Schemann JF, Guinot C, Ilboudo L, Momo G, Ko B, et al. (2003) Trachoma, flies and environmental factors in Burkina Faso. Trans R Soc Trop Med Hyg 97: 63-68.

142. Emerson PM, Bailey RL, Mahdi OS, Walraven GE, Lindsay SW (2000) Transmission ecology of the fly Musca sorbens, a putative vector of trachoma. Trans R Soc Trop Med Hyg 94: 28-32.

143. Emerson PM, Bailey RL, Walraven GE, Lindsay SW (2001) Human and other faeces as breeding media of the trachoma vector Musca sorbens. Med Vet Entomol 15: 314-320.

144. Emerson PM, Simms VM, Makalo P, Bailey RL (2005) Household pit latrines as a potential source of the fly Musca sorbens--a one year longitudinal study from The Gambia. Trop Med Int Health 10: 706-709.

145. Curtis CF, Hawkins PM (1982) Entomological studies of on-site sanitation systems in Botswana and Tanzania. Trans R Soc Trop Med Hyg 76: 99-108.

146. Emerson PM, Lindsay SW, Walraven GE, Dibba SM, Lowe KO, et al. (2002) The Flies and Eyes project: design and methods of a cluster-randomised intervention study to confirm the importance of flies as trachoma vectors in The Gambia and to test a sustainable method of fly control using pit latrines. Ophthalmic Epidemiol 9: 105-117.

147. Emerson PM, Lindsay SW, Alexander N, Bah M, Dibba SM, et al. (2004) Role of flies and provision of latrines in trachoma control: cluster-randomised controlled trial. Lancet 363: 1093-1098.

148. Stoller NE, Gebre T, Ayele B, Zerihun M, Assefa Y, et al. (2011) Efficacy of latrine promotion on emergence of infection with ocular Chlamydia trachomatis after mass antibiotic treatment: a cluster-randomized trial. Int Health 3: 75-84.

149. Haile M, Tadesse Z, Gebreselassie S, Ayele B, Gebre T, et al. (2013) The association between latrine use and trachoma: a secondary cohort analysis from a randomized clinical trial. Am J Trop Med Hyg 89: 717-720.

150. Montgomery MA, Desai MM, Elimelech M (2010) Comparing the effectiveness of shared versus private latrines in preventing trachoma in rural Tanzania. Am J Trop Med Hyg 82: 693-695.

151. Simms VM, Makalo P, Bailey RL, Emerson PM (2005) Sustainability and acceptability of latrine provision in The Gambia. Trans R Soc Trop Med Hyg 99: 631-637.

152. Hagi M, Schemann JF, Mauny F, Momo G, Sacko D, et al. (2010) Active trachoma among children in Mali: Clustering and environmental risk factors. PLoS Negl Trop Dis 4: e583.

153. Koukounari A, Toure S, Donnelly CA, Ouedraogo A, Yoda B, et al. (2011) Integrated monitoring and evaluation and environmental risk factors for urogenital schistosomiasis

and active trachoma in Burkina Faso before preventative chemotherapy using sentinel sites. BMC Infect Dis 11: 191.

154. Prost A, Negrel AD (1989) Water, trachoma and conjunctivitis. Bull World Health Organ 67: 9-18.

155. Sarkies JW (1967) Dust and the incidence of severe trachoma. Br J Ophthalmol 51: 97-100.

156. Schwab L, Whitfield R, Jr., Ross-Degnan D, Steinkuller P, Swartwood J (1995) The epidemiology of trachoma in rural Kenya. Variation in prevalence with lifestyle and environment. Study Survey Group. Ophthalmology 102: 475-482.

157. Schemann JF, Laffly D, Sacko D, Zephak G, Malvy D (2007) Trichiasis and geoclimatic factors in Mali. Trans R Soc Trop Med Hyg 101: 996-1003.

158. Clements AC, Kur LW, Gatpan G, Ngondi JM, Emerson PM, et al. (2010) Targeting Trachoma Control through Risk Mapping: The Example of Southern Sudan. PLoS Negl Trop Dis 4.

159. Alemayehu W, Melese M, Fredlander E, Worku A, Courtright P (2005) Active trachoma in children in central Ethiopia: association with altitude. Trans R Soc Trop Med Hyg 99: 840-843.

160. West SK (2003) Blinding trachoma: prevention with the safe strategy. Am J Trop Med Hyg 69: 18-23.

161. Ramesh A, Kovats S, Haslam D, Schmidt E, Gilbert CE (2013) The impact of climatic risk factors on the prevalence, distribution, and severity of acute and chronic trachoma. PLoS Negl Trop Dis 7: e2513.

162. Holm SO, Jha HC, Bhatta RC, Chaudhary JSP, Thapa BB, et al. (2001) Comparison of two azithromycin distribution strategies for controlling trachoma in Nepal. Bulletin of the World Health Organization 79: 194-200.

163. Lee DC, Chidambaram JD, Porco TC, Lietman TM (2005) Seasonal effects in the elimination of trachoma. Am J Trop Med Hyg 72: 468-470.

164. da Cruz L, Dadour IR, McAllister IL, Jackson A, Isaacs T (2002) Seasonal variation in trachoma and bush flies in north-western Australian Aboriginal communities. Clin Experiment Ophthalmol 30: 80-83.

165. Assaad FA, Maxwell-Lyons F, Sundaresan T (1969) Use of local variations in trachoma endemicity in depicting interplay between socio-economic conditions and disease. Bull World Health Organ 41: 181-194.

166. Jansen E, Baltussen RM, van Doorslaer E, Ngirwamungu E, Nguyen MP, et al. (2007) An eye for inequality: how trachoma relates to poverty in Tanzania and Vietnam. Ophthalmic Epidemiol 14: 278-287.

167. Tielsch JM, West KP, Jr., Katz J, Keyvan-Larijani E, Tizazu T, et al. (1988) The epidemiology of trachoma in southern Malawi. Am J Trop Med Hyg 38: 393-399.

168. Bartley M (2004) Health inequality: an introduction to theories, concepts and methods. Cambridge: Polity Press.

169. Reardon T, Vosti SA (1995) Links between rural poverty and the environment in developing countries: Asset categories and investment poverty. World Development 23: 1495-1506.

170. Pullan RL, Sturrock HJ, Soares Magalhaes RJ, Clements AC, Brooker SJ (2012) Spatial parasite ecology and epidemiology: a review of methods and applications. Parasitology 139: 1870-1887.

171. Gatrell AC, Bailey TC, Diggle PJ, Rowlingson B (1996) Spatial Point Pattern Analysis and Its Application in Geographical Epidemiology. Transactions of the Institute of British Geographers 21: 256-274.

172. Diggle PJ (1993) Point process modelling in environmental epidemiology. In: Barnett V, Turkman KF, editors. Statistics for the environment. Chichester: John Wiley.

173. Legendre P, Fortin MJ (1989) Spatial pattern and ecological analysis. Vegetation 80: 107-138.

174. Legendre P, Dale M, Fortin MJ, Gurevitch J, Hohn M, et al. (2002) The consequences of spatial structure for the design and analysis of ecological field studies. Ecography 25: 601-615.

175. Fan J, Wenyang Z (2008) Statistical methods with varying coefficient models. Statistics and its interface 1: 179-195.

176. Fotheringham AS (1997) Trends in quantitiative methods I: stressing the local. Progress in Human Geography 21: 88-96.

177. Broman AT, Shum K, Munoz B, Duncan DD, West SK (2006) Spatial clustering of ocular chlamydial infection over time following treatment, among households in a village in Tanzania. Invest Ophthalmol Vis Sci 47: 99-104.

178. Baggaley RF, Solomon AW, Kuper H, Polack S, Massae PA, et al. (2006) Distance to water source and altitude in relation to active trachoma in Rombo district, Tanzania. Trop Med Int Health 11: 220-227.

179. Polack SR, Solomon AW, Alexander ND, Massae PA, Safari S, et al. (2005) The household distribution of trachoma in a Tanzanian village: an application of GIS to the study of trachoma. Trans R Soc Trop Med Hyg 99: 218-225.

180. Lansingh VC, Carter MJ (2007) Trachoma surveys 2000-2005: results, recent advances in methodology, and factors affecting the determination of prevalence. Surv Ophthalmol 52: 535-546.

181. King JD, Odermatt P, Utzinger J, Ngondi J, Bamani S, et al. (2013) Trachoma among children in community surveys from four African countries and implications of using school surveys for evaluating prevalence. Int Health 5: 280-287.

182. Dunson DB (2001) Commentary: practical advantages of Bayesian analysis of epidemiologic data. Am J Epidemiol 153: 1222-1226.

183. Gilks WR, Richardson S, Spiegelhalter DJ (1996) Markov Chain Monte Carlo in Practice: Chapman & Hall. null p.

184. Richardson S, Thomson A, Best N, Elliott P (2004) Interpreting posterior relative risk estimates in disease-mapping studies. Environ Health Perspect 112: 1016-1025.

185. Magalhaes RJ, Clements AC, Patil AP, Gething PW, Brooker S (2011) The applications of model-based geostatistics in helminth epidemiology and control. Adv Parasitol 74: 267-296.

186. Patil AP, Gething PW, Piel FB, Hay SI (2011) Bayesian geostatistics in health cartography: the perspective of malaria. Trends Parasitol 27: 246-253.

187. Pullan RL, Gething PW, Smith JL, Mwandawiro CS, Sturrock HJ, et al. (2011) Spatial modelling of soil-transmitted helminth infections in Kenya: a disease control planning tool. PLoS Negl Trop Dis 5: e958.

188. Stern DI, Gething PW, Kabaria CW, Temperley WH, Noor AM, et al. (2011) Temperature and malaria trends in highland East Africa. PLoS One 6: e24524.

189. Brooker S (2010) Estimating the global distribution and disease burden of intestinal nematode infections: adding up the numbers--a review. Int J Parasitol 40: 1137-1144.

190. Pullan RL, Bethony JM, Geiger SM, Cundill B, Correa-Oliveira R, et al. (2008) Human helminth co-infection: analysis of spatial patterns and risk factors in a Brazilian community. PLoS Negl Trop Dis 2: e352.

191. Clements AC, Firth S, Dembele R, Garba A, Toure S, et al. (2009) Use of Bayesian geostatistical prediction to estimate local variations in Schistosoma haematobium infection in western Africa. Bull World Health Organ 87: 921-929.

192. Summary of the Seventeenth Meeting of the International Task Force for Disease Eradication; 2000 October 12, 2010; Atlanta, GA. The Carter Center,.

193. International Coalition for Trachoma Control (2011) 2020 INSight: A Global Strategy to Eliminate Blinding Trachoma.

194. Smith JL, Flueckiger RM, Hooper PJ, Polack S, Cromwell EA, et al. (2013) The Geographical Distribution and Burden of Trachoma in Africa. PLoS Negl Trop Dis 7: e2359.

195. Hay SI, Snow RW (2006) The Malaria Atlas Project: developing global maps of malaria risk. PLoS Med 3: e473.

196. Brooker S, Hotez PJ, Bundy DA (2010) The global atlas of helminth infection: mapping the way forward in neglected tropical disease control. PLoS Negl Trop Dis 4: e779.

197. Brooker S, Kabatereine NB, Smith JL, Mupfasoni D, Mwanje MT, et al. (2009) An updated atlas of human helminth infections: the example of East Africa. Int J Health Geogr 8: 42.

198. Hurlimann E, Schur N, Boutsika K, Stensgaard AS, Laserna de Himpsl M, et al. (2011) Toward an open-access global database for mapping, control, and surveillance of neglected tropical diseases. PLoS Negl Trop Dis 5: e1404.

199. Simarro PP, Cecchi G, Paone M, Franco JR, Diarra A, et al. (2010) The Atlas of human African trypanosomiasis: a contribution to global mapping of neglected tropical diseases. Int J Health Geogr 9: 57.

200. PubMed: National Center for Biotechnology Information. Available: http://www.ncbi.nlm.nih.gov/pubmed.

201. Embase: Elsevier B.V. Available: http://www.embase.com.

202. World Health Organization. Report of the twelfth meeting of the WHO alliance for the Global Elimination of Blinding Trachoma by 2020; 2008 April 28-30, 2008; Geneva. WHO.

203. WHO (2013) Global Alliance for the Elimination of Blinding Trachoma by 2020. Wkly Epidemiol Rec 88: 242-251.

204. Taylor HR, Fox SS, Xie J, Dunn RA, Arnold AL, et al. (2010) The prevalence of trachoma in Australia: the National Indigenous Eye Health Survey. Med J Aust 192: 248-253.

205. Australian Bureau of Statistics (2011) Australian Statistical Geography Standard (ASGS): Volume 5 - Remoteness Structure, July 2011.

206. Linard C, Gilbert M, Snow RW, Noor AM, Tatem AJ (2012) Population Distribution, Settlement Patterns and Accessibility across Africa in 2010. PLoS ONE 7: e31743.

207. United Nations DoEaSA, Population Division, (2013) World Population Prospects: The 2012 Revision. United Nations.

208. World Health Organization. Report of the sixteenth meeting of the WHO alliance for the Global Elimination of Blinding Trachoma by 2020; 2012; Geneva. WHO.

209. World Health Organization. Report of the ninth meeting of the WHO alliance for the Global Elimination of Blinding Trachoma by 2020; 2005; Geneva. WHO.

210. Khandekar R, Mohammed AJ (2007) The prevalence of trachomatous trichiasis in Oman (Oman eye study 2005). Ophthalmic Epidemiol 14: 267-272.

211. World Health Organization. Report of the tenth meeting of the WHO alliance for the Global Elimination of Blinding Trachoma by 2020; 2006; Geneva. WHO.

212. Mathew AA, Keeffe JE, Le Mesurier RT, Taylor HR (2009) Trachoma in the Pacific Islands: evidence from Trachoma Rapid Assessment. Br J Ophthalmol 93: 866-870.

213. Jie Y, Xu L, Ma K, Zhang S, Zhu J, et al. (2008) Prevalence of trachoma in the adult Chinese population. The Beijing Eye Study. Eye (Lond) 22: 790-791.

214. Boost M, Cho P (2005) High incidence of trachoma in rural areas of Guangxi, China. Lancet Infect Dis 5: 735-736.

215. Wang LH, Wang B, Wang HY, Jiao WZ, Zhou CC, et al. (2010) [Prevalence of trachoma in rural primary school children in Tengzhou City of Shandong Province in China]. Zhonghua Yan Ke Za Zhi 46: 395-399.

216. Goldschmidt P, Vanzzini Zago V, Diaz Vargas L, Espinoza Garcia L, Morales Montoya C, et al. (2007) Chlamydia trachomatis in the conjunctiva of children living in three rural areas in Mexico. Rev Panam Salud Publica 22: 29-34.

217. Furtado JM, Lansingh VC, Carter MJ, Milanese MF, Pena BN, et al. (2012) Causes of blindness and visual impairment in Latin America. Surv Ophthalmol 57: 149-177.

218. Miller H, Gallego G, Rodriguez G (2010) [Clinical evidence of trachoma in Colombian Amerindians of the Vaupes Province]. Biomedica 30: 432-439.

219. Salim AR, Sheikh HA (1975) Trachoma in the Sudan. An epidemiological study. Br J Ophthalmol 59: 600-604.

220. Harding-Esch EM, Edwards T, Sillah A, Sarr I, Roberts CH, et al. (2009) Active trachoma and ocular Chlamydia trachomatis infection in two Gambian regions: on course for elimination by 2020? PLoS Negl Trop Dis 3: e573.

221. Yayemain D, King JD, Debrah O, Emerson PM, Aboe A, et al. (2009) Achieving trachoma control in Ghana after implementing the SAFE strategy. Trans R Soc Trop Med Hyg 103: 993-1000.

222. Hagan M, Yayemain D, Ahorsu F, Aboe A (2009) Prevalence of active trachoma two years after control activities. Ghana Med J 43: 54-60.

223. Bamani S, King JD, Dembele M, Coulibaly F, Sankara D, et al. (2010) Where do we go from here? Prevalence of trachoma three years after stopping mass distribution of antibiotics in the regions of Kayes and Koulikoro, Mali. PLoS Negl Trop Dis 4: e734.

224. Dembele M, Bamani S, Dembele R, Traore MO, Goita S, et al. (2012) Implementing preventive chemotherapy through an integrated National Neglected Tropical Disease Control Program in Mali. PLoS Negl Trop Dis 6: e1574.

225. The Carter Center. Summary Proceedings: Twelfth Annual Trachoma Control Program Review; 2011 Feb. 22-24, 2011; Atlanta, GA, USA. The Carter Center.

226. (2006) Morocco Achieves Intervention Goals in Trachoma Control Campaign. Trachoma Matters,: International Trachoma Initiative.

227. Khandekar R (2009) Elimination of blinding trachoma in oman. Oman Med J 24: 67-69.

228. Edwards T, Smith J, Sturrock HJ, Kur LW, Sabasio A, et al. (2012) Prevalence of trachoma in unity state, South Sudan: results from a large-scale population-based survey and potential implications for further surveys. PLoS Negl Trop Dis 6: e1585.

229. Emerson PM, Ngondi J, Biru E, Graves PM, Ejigsemahu Y, et al. (2008) Integrating an NTD with one of "The big three": combined malaria and trachoma survey in Amhara Region of Ethiopia. PLoS Negl Trop Dis 2: e197.

230. Gomez Rubio V, Best N, Richardson S, Li G, Clarke P Bayesian Statistics Small Area Estimation. Working paper: The BIAS project, Imperial College, London.

231. Harding-Esch EM, Edwards T, Mkocha H, Munoz B, Holland MJ, et al. (2010) Trachoma prevalence and associated risk factors in the gambia and Tanzania: baseline results of a cluster randomised controlled trial. PLoS Negl Trop Dis 4: e861.

232. West S, Munoz B (2010) Trachoma in Latin America: an opportunity for elimination. Biomedica 30: 315-316.

233. Longfield K, Smith B, Gray R, Ngamkitpaiboon L, Vielot N (2013) Putting health metrics into practice: using the disability-adjusted life year for strategic decision making. BMC Public Health 13: S2.

234. Murray CJ, Salomon JA, Mathers C (2000) A critical examination of summary measures of population health. Bull World Health Organ 78: 981-994.

235. Evans TG, Ranson MK (1995) The global burden of trachomatous visual impairment: II. Assessing burden. Int Ophthalmol 19: 271-280.

236. Ezz al Arab G, Tawfik N, El Gendy R, Anwar W, Courtright P (2001) The burden of trachoma in the rural Nile Delta of Egypt: a survey of Menofiya governorate. Br J Ophthalmol 85: 1406-1410.

237. Salomon JA, Vos T, Hogan DR, Gagnon M, Naghavi M, et al. (2012) Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010. Lancet 380: 2129-2143.

238. Gouda H, Powles J, Barendregt J, Emerson P, Ngondi J (2012) The burden of trachoma in South Sudan: assessing the health losses from a condition of graded severity. PLoS Negl Trop Dis 6: e1538.

239. (IAPB) TIAftPoB (2013) IAPB to Host Global Vision Database. IAPB Newsletter.

240. Solomon AW (2013) Incident cases of trichiasis after the elimination of *Chlamydia trachomatis* infection.

241. Ngondi J, Ole-Sempele F, Onsarigo A, Matende I, Baba S, et al. (2006) Prevalence and causes of blindness and low vision in southern Sudan. PLoS Med 3: e477.

242. Taylor HR, Jonas JB, Keeffe J, Leasher J, Naidoo K, et al. Disability weights for vision disorders in Global Burden of Disease study. The Lancet 381: 23.

243. Kirkwood B, Smith P, Marshall T, Prost A (1983) Relationships between mortality, visual acuity and microfilarial load in the area of the Onchocerciasis Control Programme. Trans R Soc Trop Med Hyg 77: 862-868.

244. The World Bank (2013) Africa's Pulse. Washington DC: World Bank.

245. Rajak SN, Makalo P, Sillah A, Holland MJ, Mabey DCW, et al. (2010) Trichiasis Surgery in The Gambia: A 4-Year Prospective Study. Investigative Ophthalmology & Visual Science 51: 4996-5001.

246. Mozzato-Chamay N, Mahdi OS, Jallow O, Mabey DC, Bailey RL, et al. (2000) Polymorphisms in candidate genes and risk of scarring trachoma in a Chlamydia trachomatis--endemic population. J Infect Dis 182: 1545-1548.

247. Burton MJ, Kinteh F, Jallow O, Sillah A, Bah M, et al. (2005) A randomised controlled trial of azithromycin following surgery for trachomatous trichiasis in the Gambia. Br J Ophthalmol 89: 1282-1288.

248. Haileselassie T, Bayu S (2007) Altitude-a risk factor for active trachoma in southern Ethiopia. Ethiop Med J 45: 181-186.

249. Rabiu MM, Gudlavalleti MV, Gilbert CE, Sivasubramaniam S, Kyari F, et al. (2011) Ecological determinants of blindness in Nigeria: the Nigeria National Blindness and Visual Impairment Survey. S Afr Med J 101: 53-58.

250. Ngondi J, Gebre T, Shargie EB, Graves P, Ejigsemahu Y, et al. (2008) Risk factors for active trachoma in children and trichiasis in adults: a household survey in Amhara Regional State, Ethiopia. Trans R Soc Trop Med Hyg 102: 432-438.

251. Goldstein H. (1995) Multilevel statistical models: second edition. London: John Wiley & Sons Inc.

252. Boyd HA, Flanders WD, Addiss DG, Waller LA (2005) Residual spatial correlation between geographically referenced observations: a Bayesian hierarchical modeling approach. Epidemiology 16: 532-541.

253. Pullan RL, Bukirwa H, Staedke SG, Snow RW, Brooker S (2010) Plasmodium infection and its risk factors in eastern Uganda. Malar J 9: 2.

254. Walker M, Hall A, Basanez MG (2011) Individual predisposition, household clustering and risk factors for human infection with Ascaris lumbricoides: new epidemiological insights. PLoS Negl Trop Dis 5: e1047.

255. Yang J, Zhao Z, Li Y, Krewski D, Wen SW (2009) A multi-level analysis of risk factors for Schistosoma japonicum infection in China. Int J Infect Dis 13: e407-412.

256. Tallo VL, Carabin H, Alday PP, Balolong E, Jr., Olveda RM, et al. (2008) Is mass treatment the appropriate schistosomiasis elimination strategy? Bull World Health Organ 86: 765-771.

257. Goldstein H, Browne W, Rasbash J (2002) Partitioning Variation in Multilevel Models. Understanding Statistics 4.

258. Browne WJ, Subramanian SV, Jones K, Goldstein H (2005) Variance partitioning in multilevel logistic models that exhibit overdispersion. Journal of the Royal Statistical Society: Series A (Statistics in Society) 168: 599-613.

259. World Bank Washington, DC:World Bank. http://hdl.handle.net/10986/6014 Accessed February 8, 2013.

260. FAO:Food and Agriculture Organization of the United Nations. http://www.fao.org/nr/water/aquastat/countries_regions/NGA/index.stm Accessed 26 July.

261. Global Atlas of Trachoma. www.trachomaatlas.org Accessed February 8.

262. Rabiu MM, Kyari F, Ezelum C, Elhassan E, Sanda S, et al. (2012) Review of the publications of the Nigeria national blindness survey: methodology, prevalence, causes of

blindness and visual impairment and outcome of cataract surgery. Ann Afr Med 11: 125-130.

263. Dineen B, Gilbert CE, Rabiu M, Kyari F, Mahdi AM, et al. (2008) The Nigerian national blindness and visual impairment survey: Rationale, objectives and detailed methodology. BMC Ophthalmol 8: 17.

264. Howe LD, Galobardes B, Matijasevich A, Gordon D, Johnston D, et al. (2012) Measuring socio-economic position for epidemiological studies in low- and middle-income countries: a methods of measurement in epidemiology paper. Int J Epidemiol 41: 871-886.

265. UNICEF Wa (2010) Progress on Sanitation and Drinking Water; 2010 update. . WHO/UNICEF Joint Monitoring Programme for Water Supply and Sanitation.

266. Bartholomew DJ, Steele F, Moustaki I, Galbraith JI (2008) Chapter 5: Prinicpal Components Analysis. In: A G, editor. Analysis of multivariate social science data. Boca Raton: Taylor & Francis Group, LLC.

267. Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, et al. (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36: 027-046.

268. Graham MH (2003) Confronting multicollinearity in ecological multiple regression. Ecology 84: 2809-2815.

269. King JD, Jip N, Jugu YS, Othman A, Rodgers AF, et al. (2010) Mapping trachoma in Nasarawa and Plateau States, central Nigeria. Br J Ophthalmol 94: 14-19.

270. Khandekar R, Nga NH, Mai P (2006) Blinding trachoma in the northern provinces of Vietnam--a cross sectional survey. Ophthalmic Epidemiol 13: 183-189.

271. Durkin SR, Casson RJ, Newland HS, Aung TH, Shein WK, et al. (2007) Prevalence of trachoma-related trichiasis and corneal opacity in rural Myanmar: the Meiktila Eye Study. Ophthalmology 114: e7-11.

272. Roba AA, Patel D, Zondervan M (2013) Risk of trachoma in a SAFE intervention area. Int Ophthalmol 33: 53-59.

273. Goulson D, Derwent LC, Hanley ME, Dunn DW, Abolins SR (2005) Predicting calyptrate fly populations from the weather, and probable consequences of climate change. Journal of Applied Ecology 42: 795-804.

274. Katz J, West KP, Jr., Khatry SK, LeClerq SC, Pradhan EK, et al. (1996) Prevalence and risk factors for trachoma in Sarlahi district, Nepal. Br J Ophthalmol 80: 1037-1041.

275. Blum D, Feachem RG, Huttly SR, Kirkwood BR, Emeh RN (1987) The effects of distance and season on the use of boreholes in northeastern Imo State, Nigeria. J Trop Med Hyg 90: 45-50.

276. Tawfik MS (1969) Ecological studies on some desert flies in Egypt. Indian Journal of Entomology 31: 201-221.

277. Clements AC, Lwambo NJ, Blair L, Nyandindi U, Kaatano G, et al. (2006) Bayesian spatial analysis and disease mapping: tools to enhance planning and implementation of a schistosomiasis control programme in Tanzania. Trop Med Int Health 11: 490-503.

278. Karimurio J, Gichangi M, Ilako DR, Adala HS, Kilima P (2006) Prevalence of trachoma in six districts of Kenya. East Afr Med J 83: 63-68.

279. Loewenthal R, Pe'er J (1990) A prevalence survey of ophthalmic diseases among the Turkana tribe in north-west Kenya. Br J Ophthalmol 74: 84-88.

280. Whitfield R, Jr., Schwab L, Bakker NJ, Bisley GG, Ross-Degnan D (1983) Cataract and corneal opacity are the main causes of blindness in the Samburu tribe of Kenya. Ophthalmic Surg 14: 139-144.

281. Kenya GotRo (2012) Vision 2030 Development Strategy for Northern Kenya and other Arid Lands. Ministry of State for Development of Northern Kenya and other Arid Lands, Government of the Republic of Kenya.

282. Government of Kenya (2004) Draft National Policy for the Sustainable Development of Arid and Semi-arid Lands of Kenya.

283. Okwi PO, Ndeng'e G, Kristjanson P, Arunga M, Notenbaert A, et al. (2007) Spatial determinants of poverty in rural Kenya. Proc Natl Acad Sci U S A 104: 16769-16774.

284. Kristjanson P, Radeny M, Baltenweck I, Ogutu J, Notenbaert A (2005) Livelihood mapping and poverty correlates at a meso-level in Kenya. Food Policy: 568-583.

285. Kristjanson P, Mango N, Krishna Kumari A, Radeny M, Johnson N (2010) Understanding poverty dynamics in Kenya. Journal of International Development 22: 978-996.

286. Ministry of Agriculture and Livestock Development (MoALD) (2010) Agricultural sector development strategy. Government of Kenya.

287. Kenya National Bureau of Statistics National Socio-Economic and Poverty Atlas.

288. Kenya National Bureau of Statistics (1997) Kenya welfare monitoring survey 1997. Government of Kenya.

289. Kenya National Bureau of Statistics (1999) Kenya Population and Housing Census. Government of Kenya.

290. Wagstaff A, Paci P, van Doorslaer E (1991) On the measurement of inequalities in health. Soc Sci Med 33: 545-557.

291. Pullan RL, Freeman MC, Gething PW, Brooker S ((under review)) Mapping inequalities in access to improved drinking water supply and sanitation across sub-Saharan Africa. PLoS Negl Trop Dis.

292. Kenya National Bureau of Statistics (2006) Kenya Integrated Household Budget Survey 2005-2006. Nairobi, Kenya: Kenya National Bureau of Statistics - Ministry of Planning and National development.

293. Wint G, Robinson T (2007) Gridded livestock of the world 2007. Rome: FAO.

294. Benson T, Chamberlin J, Rhinehart I (2005) Why the poor in rural Malawi are where they are: an analysis of the spatial determinants of the local prevalence of poverty. Washington, DC.

295. Minot N, Baulch B, Epprecht M (2003) Poverty and Inequality in Vietnam: Spatial Patterns and Geographic Determinants. Washington, DC: International Food Policy Research Institute in collaboration with the Institute of Developmental Studies.

296. United Nations SALB The Second Administrative Level Boundaries data set project: http://www.unsalb.org/.

297. Gelman A, Rubin D (1992) Inference from Iterative Simulation Using Multiple Sequences. Statistical Science 7: 457.

298. Brooks SP, Gelman A (1998) General Methods for Monitoring Convergence of Iterative Simulations. Journal of Computational and Graphical Statistics 7: 434-455.

299. Brooker S, Hay SI, Bundy DA (2002) Tools from ecology: useful for evaluating infection risk models? Trends Parasitol 18: 70-74.

300. Dolin PJ, Faal H, Johnson GJ, Minassian D, Sowa S, et al. (1997) Reduction of trachoma in a sub-Saharan village in absence of a disease control programme. Lancet 349: 1511-1512.

301. Fouad D, Mousa A, Courtright P (2004) Sociodemographic characteristics associated with blindness in a Nile Delta governorate of Egypt. Br J Ophthalmol 88: 614-618.

302. Montgomery MA, Desai MM, Elimelech M (2010) Assessment of latrine use and quality and association with risk of trachoma in rural Tanzania. Trans R Soc Trop Med Hyg 104: 283-289.

303. Harris P, Fotheringham AS, Crespo R, Charlton M (2010) The Use of Geographically Weighted Regression for Spatial Prediction: An Evaluation of Models Using Simulated Data Sets. Mathematical Geosciences 42: 657-680.

304. Mwandawiro CS, Nikolay B, Kihara JH, Ozier O, Mukoko DA, et al. (2013) Monitoring and evaluating the impact of national school-based deworming in Kenya: study design and baseline results. Parasit Vectors 6: 198.

305. Soares Magalhaes RJ, Biritwum NK, Gyapong JO, Brooker S, Zhang Y, et al. (2011) Mapping helminth co-infection and co-intensity: geostatistical prediction in Ghana. PLoS Negl Trop Dis 5: e1200.

306. Ashton RA, Kefyalew T, Tesfaye G, Pullan RL, Yadeta D, et al. (2011) School-based surveys of malaria in Oromia Regional State, Ethiopia: a rapid survey method for malaria in low transmission settings. Malar J 10: 25.

307. Gitonga CW, Karanja PN, Kihara J, Mwanje M, Juma E, et al. (2010) Implementing school malaria surveys in Kenya: towards a national surveillance system. Malar J 9: 306.

308. Deribe K, Meribo K, Gebre T, Hailu A, Ali A, et al. (2012) The burden of neglected tropical diseases in Ethiopia, and opportunities for integrated control and elimination. Parasit Vectors 5: 240.

309. Kabatereine NB, Malecela M, Lado M, Zaramba S, Amiel O, et al. (2010) How to (or not to) integrate vertical programmes for the control of major neglected tropical diseases in sub-Saharan Africa. PLoS Negl Trop Dis 4: e755.

310. Linehan M, Hanson C, Weaver A, Baker M, Kabore A, et al. (2011) Integrated implementation of programs targeting neglected tropical diseases through preventive chemotherapy: proving the feasibility at national scale. Am J Trop Med Hyg 84: 5-14.

311. Brooker S, Kabatereine NB, Gyapong JO, Stothard JR, Utzinger J (2009) Rapid mapping of schistosomiasis and other neglected tropical diseases in the context of integrated control programmes in Africa. Parasitology 136: 1707-1718.

312. WHO. Report of the fifteenth meeting of the WHO alliance for the Global Elimination of Blinding Trachoma by 2020; 2011; Geneva. WHO.

313. Mabey D, Solomon A, Foster A (2003) Trachoma. Lancet 362.

314. Minetti A, Riera-Montes M, Nackers F, Roederer T, Koudika MH, et al. (2012) Performance of small cluster surveys and the clustered LQAS design to estimate local-level vaccination coverage in Mali. Emerg Themes Epidemiol 9: 6.

315. Pezzoli L, Andrews N, Ronveaux O (2010) Clustered lot quality assurance sampling to assess immunisation coverage: increasing rapidity and maintaining precision. Trop Med Int Health 15: 540-546.

316. Ministry of Health of Kenya (2010) District level prevalence estimates of trachoma in Kenya. [Data provided by the Ministry of Health Kenya for this project] ed.

317. UNESCO (United Nations Educational S, and Cultural Organization), (2011) International Standard Classification of Education (ISCED). Paris: United Nations Educational, Scientific and Cultural Organization.

318. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33: 159-174.

319. Dawson CR, Jones BR, Tarizzo ML, WHO (1981) Guide to trachoma control in programmes for the prevention of blindness: Geneva : World Health Organization.

320. Jovani R, Tella JL (2006) Parasite prevalence and sample size: misconceptions and solutions. Trends Parasitol 22: 214-218.

321. Brooker S, Kolaczinski JH, Gitonga CW, Noor AM, Snow RW (2009) The use of schools for malaria surveillance and programme evaluation in Africa. Malar J 8: 231.

322. Lietman TM, Gebre T, Ayele B, Ray KJ, Maher MC, et al. (2011) The epidemiological dynamics of infectious trachoma may facilitate elimination. Epidemics 3: 119-124.

323. Legendre P (1993) Spatial autocorrelation: trouble or new paradigm? Ecology 74: 1659-1673.

324. Basáñez M-G, McCarthy JS, French MD, Yang G-J, Walker M, et al. (2012) A Research Agenda for Helminth Diseases of Humans: Modelling for Control and Elimination. PLoS Negl Trop Dis 6: e1548.

# Appendices

# Appendix 3.1: National blindness surveys in Africa

| Reference | Country | Survey characteristics (Design; year; age) | Number examined | Crude Prevalence (per 1000) | |
|---|---|---|---|---|---|
| | | | | TB | LV |
| [1] | Benin | PBPS; 1990; 0-99 | 7047 | 0.18 | 1.04 |
| [2] | Botswana | RAAB; 2007; 50-99 | 2127 | 3.76 | - |
| [3] | Chad | PBPS; 1985; 0-99 | 5002 | 5.26 | - |
| [4] | Eritrea | RAAB; 2008; 50-99 | 3163 | 0.95 | 0.63 |
| [5] | Ethiopia | PBPS; 2005; 0-99 | 71066 | 1.84 | 2.84 |
| [6] | Morocco | PBPS; 1992; 0-99 | 25061 | 0.31 | - |
| [7] | Nigeria | RAAB; 2005; 40-99 | 13591 | 1.76 | 0.9 |
| [8] | The Gambia | PBPS; 1986; 0-99 | 8174 | 1.19 | - |
| [9] | The Gambia | PBPS; 1996; 5-99 (NOT USED) | 13046 | 0.23 | 0.69 |

## References


1. Negrel AD, Avognon Z, Minassian DC, Babagbeto M, Oussa G, et al. (1995) [Blindness in Benin]. Med Trop (Mars) 55: 409-414.
2. Nkomazana O (2007) A national survey of visual impairment in Botswana. Community Eye Health 20: 9.
3. Pascolini D, Mariotti SP, Pokharel GP, Pararajasegaram R, Etya'ale D, et al. (2004) 2002 global update of available data on visual impairment: a compilation of population-based prevalence studies. Ophthalmic Epidemiol 11: 67-115.
4. Muller A (2008) RAAB in Ethiopia.
5. Berhane Y, Worku A, Bejiga A, (2006) National Survey on Blindness, Low Vision and Trachoma in Ethiopia. Addis Ababa, Ethiopia: Federal Ministry of Health of Ethiopia.
6. World Health Organization (1993) Report on the intercountry meeting for evaluating national prevention of blindness programmes.
7. Abdull MM, Sivasubramaniam S, Murthy GV, Gilbert C, Abubakar T, et al. (2009) Causes of blindness and visual impairment in Nigeria: the Nigeria national blindness and visual impairment survey. Invest Ophthalmol Vis Sci 50: 4114-4120.
8. Dolin PJ, Faal H, Johnson GJ, Ajewole J, Mohamed AA, et al. (1998) Trachoma in The Gambia. Br J Ophthalmol 82: 930-933.
9. Faal H, Minassian DC, Dolin PJ, Mohamed AA, Ajewole J, et al. (2000) Evaluation of a national eye care programme: re-survey after 10 years. Br J Ophthalmol 84: 948-951.

# Appendix 3.2: Data used to model prevalence of trachomatous blindness and low vision

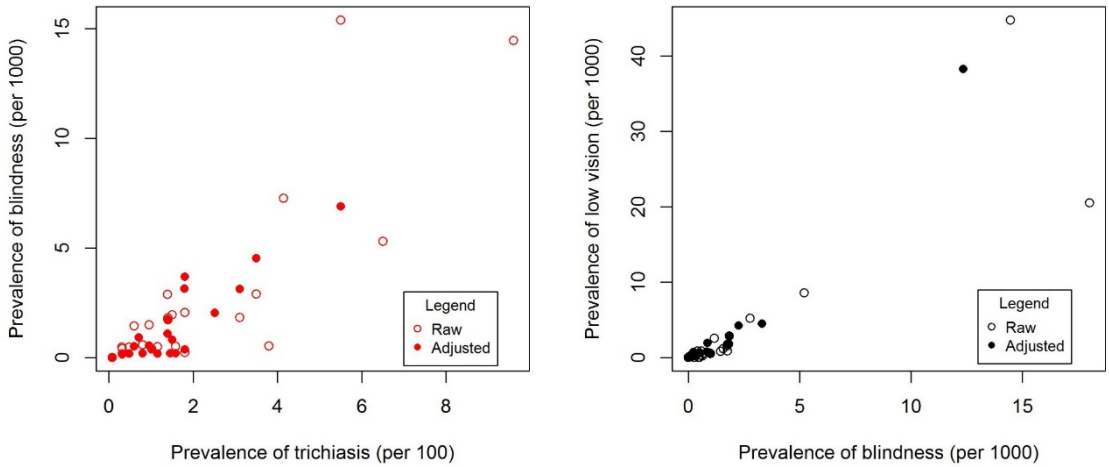



Figure A3.2.1 Plots of the raw and age-adjusted prevalence data for trachomatous trichiasis and trachomatous blindness (A) and trachomatous blindness and trachomatous low vision (B)

Table A3.2.1 Characteristics of 23 studies used to model the relationship between trachomatous trichiasis (TT) and trachomatous blindness (TB)

| Reference | Country | Setting (coverage; location; data) | Notes | Survey characteristics (year; age group; Nex) | | Prevalence (per 100 TT/1000 TB) | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Crude | | Adjusted (≥ 15 years) | |
| | | | | TT | TB | TT | TB | TT | TB |
| [1] | Australia | National; TB and TT | | 2008[a]; ≥ 15 years; 1,189 | 2008[a]; ≥ 40 years; 1,189 | 1.4 | 1.73 | 1.4 | 1.73 |
| [2] | Cameroon | Sub-national; Far-North; TB and LV | | 2010; ≥ 15 years; 41,533 | 2010; ≥ 15 years; 41,533 | 1.02 | 0.4 | 1.02 | 0.4 |
| [3] | Egypt | Sub-national; Menofiya; TB and TT | | 2000; ≥ 50 years; 2,426 | 2000; ≥ 50 years; 2,426 | 6.5 | 5.3 | 2.5 | 2.05 |
| [4,5] | Ethiopia | Sub-national; Gurage; matched | Matched to zone | 1999; ≥ 15 years; 369 | 1999; ≥ 40 years; 2,693 | 5.5 | 15.4 | 5.5 | 6.9 |
| [6] | Ethiopia | National; TB and TT | | 2005; ≥ 15 years; 16,874 | 2005; 0-99 years; 25,650 | 3.1 | 1.84 | 3.1 | 3.13 |
| [7] | Gambia | National; TB and TT | Post control | 1996; ≥ 15 years; 6,647 | 1996; 0-99 years; 13,046 | 1.8 | 0.23 | 1.8 | 0.38 |
| [8,9] | Malawi | Sub-national; Southern districts; matched | Matched to zone | 2008; ≥ 15 years; 1,135 | 2009; ≥ 50 years; 3,430 | 0.6 | 1.45 | 0.6 | 0.52 |
| [10,11] | Mali | Sub-national; Segou; matched | | 1996; ≥ 15 years; 1,630 | 1990; 0-99 years; 5,871 | 1.8 | 2.06 | 1.8 | 3.69 |
| [12,13] | Myanmar | Sub-national; Monywa; matched | Post control | 2004; ≥ 15 years; 1,005 | 2001; ≥ 50 years; 2,975 | 1.39 | 2.89 | 1.39 | 1.1 |
| [13,14] | Myanmar | Sub-national; Shwebo; matched | Post control | 2004; ≥ 15 years; 1,666 | 2001; ≥ 50 years; 2,997 | 0.32 | 0.41 | 0.32 | 0.15 |
| [15] | Myanmar | Sub-national; Mandalay Division, Meiktila; TB and TT | Post control | 2006; ≥ 40 years; 2,076 | 2006; ≥ 40 years; 2,076 | 1.4 | 1.82 | 0.7 | 0.92 |
| [16,17] | Nepal | Sub-national; Lumbini & Bheri zones; matched | Matched to zone | 1996; ≥ 15 years; 4,000 | 1995; ≥ 45 years; 4,602 | 1.5 | 1.97 | 1.5 | 0.81 |
| [18,19] | Oman | National; TB and TT | | 1997; ≥ 15 years; 4,805 | 1997; 0-99 years; 1,524 | 3.5 | 2.9 | 3.5 | 4.53 |
| [20,21] | Oman | National; TB and TT | Post control | 2005; ≥ 40 years; 2,359 | 2005; ≥ 40 years; 2,339 | 4.14 | 7.27 | 1.79 | 3.14 |
| [22] | South Sudan | Sub-national; Mankien; TB and TT | | 2005; ≥ 5 years; 2,449 | 2005; ≥ 5 years; 2,449 | 9.6 | 14.47 | 13.27 | 20.00 |
| [23] | Sudan | Sub-national; Gezira; matched | | 2008; ≥ 15 years; 5,596 | 2010; ≥ 50 years; 2,103 | 0.95 | 1.5 | 0.95 | 0.56 |

Table continued on next page

Table A3.2.1 continued

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| [23] | Sudan | Sub-national; Kassala; matched | | 2009; ≥ 15 years; 5,225 | 2010; ≥ 50 years; 2,050 | 0.48 | 0.488 | 0.48 | 0.18 |
| [23,24] | Sudan | Sub-national; North Kordofan; matched | | 2010; ≥ 15 years; 4,924 | 2010; ≥ 50 years; 2,032 | 0.08 | 0 | 0.08 | 0 |
| [24,25] | Sudan | Sub-national; Northern; matched | | 2009; ≥ 15 years; 2,770 | 2010; ≥ 50 years; 1,998 | 1.16 | 0.501 | 1.16 | 0.19 |
| [24,25] | Sudan | Sub-national; Sinnar; matched | | 2009; ≥ 15 years; 3,987 | 2010; ≥ 50 years; 1,938 | 1.58 | 0.516 | 1.58 | 0.19 |
| [24,25] | Sudan | Sub-national; White Nile; matched | | 2010; ≥ 15 years; 5,135 | 2010; ≥ 50 years; 2,097 | 0.31 | 0.477 | 0.31 | 0.18 |
| [25,26] | Tanzania | Sub-national; Kilimanjaro; matched | | 2004; ≥ 15 years; 3,393 | 2007; ≥ 50 years; 3,436 | 0.8 | 0.582 | 0.8 | 0.21 |
| [27,28] | Vietnam | National; TB and TT | Post control | 2007; ≥ 50 years; 28,033 | 2007; ≥ 50 years; 28,033 | 3.8 | 0.527 | 1.45 | 0.20 |

TB: trachomatous blindness; TT: trichiasis; Nex: number examined

[a] Indigenous populations

Table A3.2.2 Characteristics of 30 studies used to model the relationship between trachomatous blindness and trachomatous low vision

| Reference | Country | Setting (coverage; location; data) | Notes | Survey characteristics (year; age group; Nex) | Prevalence (per 1000) | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | Crude | | Adjusted (0-99 years) | |
| | | | | | TB | LV | TB | LV |
| [1] | Australia | National; TB and LV | | 2008[a]; ≥ 15 years; 1,189 | 1.73 | 1.55 | 0.83 | 0.79 |
| [29] | China | Sub-national; TB and LV | | 2008; ≥ 50 years; 2,842 | 0 | 0 | 0 | 0 |
| [2] | Cameroon | Sub-national; TB and LV | | 2010; ≥ 15 years; 41,533 | 0.4 | 0.9 | 0.25 | 0.54 |
| [3] | Egypt | Sub-national; Menofiya; TB and LV | | 2000; ≥ 50 years; 2,426 | 5.3 | 8.6 | 0.85 | 1.94 |
| [30] | Eritrea | National; TB and LV | | 2008; ≥ 50 years; 3,163 | 0.32 | 0.95 | 0.10 | 0.11 |
| [31] | Ethiopia | Sub-national; Jimma zone; TB and LV | | 1995; 0-99 years; 7,423 | 1.75 | 1.77 | 1.75 | 1.77 |
| [4] | Ethiopia | Sub-national; Gurage; TB and LV | | 1999; ≥ 40 years; 2,693 | 18.0 | 20.5 | 3.30 | 4.52 |
| [6] | Ethiopia | National; TB and LV | | 2005; 0-99; 30,022 | 1.84 | 2.84 | 1.84 | 2.84 |
| [32] | Ghana | Sub-national; | | 1991; 1,425 | 0 | 0 | 0 | 0 |
| [33] | Ghana | Sub-national; | | 2001; 2,298 | 0 | 0 | 0 | 0 |
| [34] | Gambia | National; TB and LV | | 1996; ≥ 15 years; 13,046 | 0.23 | 0.69 | 0.19 | 0.57 |
| [35] | India | Sub-national; TB and LV | | 1999; ≥ 50 years; 4,280 | 1.57 | 1.22 | 0.26 | 0.28 |
| [35] | India | Sub-national; TB and LV | | 2002; ≥ 50 years; 64,343 | 0.64 | 0.23 | 0.11 | 0.06 |
| [36] | Iran | Sub-national; TB and LV | | 2006; 0-99 years; 6,960 | 0 | 0 | 0 | 0 |
| [37] | Iran | Sub-national; TB and LV | | 2009; ≥ 50 years; 2,819 | 0.35 | 0.35 | 0.06 | 0.09 |
| [38] | Kenya | Sub-national; TB and LV | | 2002; ≥ 50 years; 3,503 | 1.16 | 2.55 | 0.13 | 0.43 |
| [39] | Cambodia | Sub-national; TB and LV | | 1996; 0-99 years; 6,558 | 0.33 | 0.46 | 0.30 | 0.46 |
| [40] | Mexico | Sub-national; TB and LV | | 2005; ≥ 50 years; 3,780 | 0 | 0 | 0 | 0 |
| [15,41] | Myanmar | Sub-national; TB and LV | | 2005; ≥ 40 years; 2,076 | 1.82 | 1.82 | 0.54 | 0.61 |
| [9] | Malawi | Sub-national; TB and LV | | 2010; ≥ 50 years; 3,430 | 1.45 | 0.86 | 0.17 | 0.15 |
| [42] | Nigeria | Sub-national; TB and LV | | 1988; 5-99 years; 6,381 | 2.77 | 5.2 | 2.25 | 2.23 |
| [43] | Nigeria | National; TB and LV | | 2007; ≥ 40 years; 13,591 | 1.76 | 0.9 | 0.33 | 0.20 |
| [44] | Nepal | Sub-national; TB and LV | | 2008; ≥ 50 years; 3,613 | 0.27 | 0 | 0.04 | 0 |
| [45] | Pakistan | National; TB and LV | | 2004; ≥ 30 years; 16,507 | 0.003 | 0.13 | 0.008 | 0.04 |
| [46] | PSE[b] | Sub-national; TB and LV | | 2008; ≥ 50 years; 3,579 | 0.58 | 0.32 | 0.06 | 0.05 |
| [38] | Rwanda | Sub-national; TB and LV | | 2006; ≥ 50 years; 2,206 | 0.45 | 0 | 0.05 | 0 |
| [47] | Sudan | Sub-national; TB and LV | | 2005; 5-99 years; 2,499 | 14.5 | 44.7 | 12.3 | 38.3 |
| [23,24] | Sudan | Sub-national; TB and LV | | 2010; ≥ 50 years; 2,050 | 0.49 | 0 | 0.07 | 0 |
| [25,26] | Tanzania | Sub-national; TB and LV | | 2007; ≥ 50 years; 3,436 | 0.58 | 0.86 | 0.07 | 0.16 |

TB: trachomatous blindness; LV: trachomatous low vision; Nex: number examined; [a] Indigenous populations; [b] Palestinian Occupied Territories

## References


1. Taylor HR, National Indigenous Eye Health Survey Team (2009) National Indigenous eye health survey: minum barreng (tracking eyes): full report. Melbourne: Indigenous Eye Health Unit, The University of Melbourne.
2. Noa Noatina B, Kagmeni G, Mengouo MN, Moungui HC, Tarini A, et al. (2013) Prevalence of trachoma in the far north region of cameroon: results of a survey in 27 health districts. PLoS Negl Trop Dis 7: e2240.
3. Ezz al Arab G, Tawfik N, El Gendy R, Anwar W, Courtright P (2001) The burden of trachoma in the rural Nile Delta of Egypt: a survey of Menofiya governorate. Br J Ophthalmol 85: 1406-1410.
4. Melese M, Alemayehu W, Bayu S, Girma T, Hailesellasie T, et al. (2003) Low vision and blindness in adults in Gurage Zone, central Ethiopia. Br J Ophthalmol 87: 677-680.
5. Bejiga A, Alemayehu W (2001) Prevalence of trachoma and its determinants in Dalocha District, Central Ethiopia. Ophthalmic Epidemiol 8: 119-125.
6. Berhane Y, Worku A, Bejiga A, (2006) National Survey on Blindness, Low Vision and Trachoma in Ethiopia. Addis Ababa, Ethiopia: Federal Ministry of Health of Ethiopia.
7. Dolin PJ, Faal H, Johnson GJ, Ajewole J, Mohamed AA, et al. (1998) Trachoma in The Gambia. Br J Ophthalmol 82: 930-933.
8. Kalua K, Chirwa T, Kalilani L, Abbenyi S, Mukaka M, et al. (2010) Prevalence and risk factors for trachoma in central and southern Malawi. PLoS One 5: e9067.
9. Kalua K, Lindfield R, Mtupanyama M, Mtumodzi D, Msiska V (2011) Findings from a rapid assessment of avoidable blindness (RAAB) in Southern Malawi. PLoS One 6: e19226.
10. Kortlang C, Koster JC, Coulibaly S, Dubbeldam RP (1996) Prevalence of blindness and visual impairment in the region of Segou, Mali. A baseline survey for a primary eye care programme. Trop Med Int Health 1: 314-319.
11. Ministry of Health of Mali (2010) District- level prevalence estimates of trachoma in Mali. [Data provided by the International Trachoma Initiative for this project]
12. Limburg H MN, Khin Aye Soe, Thuzar Han, (unpublished) Study report on rapid assessment of cataract surgical services in Monywa District, Union of Myanmar. Ministry of Health, Department of Health, Trachoma Control and Prevention
13. Kyaw TA, Ko Ko U (2007) A forty-year battle against blinding trachoma in Myanmar. New Delhi: World Health Organization. 1-3 p.
14. Limburg H MN, Khin Aye Soe, Thuzar Han, (unpublished) Study report on rapid assessment of cataract surgical services in Shwebo District, Union of Myanmar. Ministry of Health, Department of Health, Trachoma Control and Prevention
15. Casson RJ, Newland HS, Muecke J, McGovern S, Durkin S, et al. (2007) Prevalence and causes of visual impairment in rural myanmar: the Meiktila Eye Study. Ophthalmology 114: 2302-2308.
16. Pokharel GP, Regmi G, Shrestha SK, Negrel AD, Ellwein LB (1998) Prevalence of blindness and cataract surgery in Nepal. British Journal of Ophthalmology 82: 600-605.
17. Community PBPS surveys in Nepal, 1995-1996. [Data provided by Mr. BB Thapa, Programme Manager, Ministry of Health Nepal]
18. Khandekar R, Mohammed AJ, Negrel AD, Riyami AA (2002) The prevalence and causes of blindness in the Sultanate of Oman: the Oman Eye Study (OES). Br J Ophthalmol 86: 957-962.
19. Community trachoma survey in Oman, 1996-1997. [Data provided by Dr Rajiv Khandekar, Program Manager, Eye Health Care, Ministry of Health Oman]
20. Khandekar R, Mohammed AJ (2007) The prevalence of trachomatous trichiasis in Oman (Oman eye study 2005). Ophthalmic Epidemiol 14: 267-272.

21. Khandekar R, Mohammed AJ, Raisi AA (2007) Prevalence and causes of blindness & low vision; before and five years after 'VISION 2020' initiatives in Oman: a review. Ophthalmic Epidemiol 14: 9-15.
22. Ngondi J, Reacher M, Matthews F, Ole-Sempele F, Onsarigo A, et al. (2007) The epidemiology of low vision and blindness associated with trichiasis in southern Sudan. BMC Ophthalmol 7: 12.
23. Hassan A, Ngondi JM, King JD, Elshafie BE, Al Ginaid G, et al. (2011) The prevalence of blinding trachoma in northern states of Sudan. PLoS Negl Trop Dis 5: e1027.
24. RAAB in six states of Northern Sudan. [Data provided by Mr. AS Hassan, National Coordinator, Sudan Trachoma Control Program, Federal Ministry of Health Sudan]
25. Habiyakire C, Kabona G, Courtright P, Lewallen S (2010) Rapid assessment of avoidable blindness and cataract surgical services in kilimanjaro region, Tanzania. Ophthalmic Epidemiol 17: 90-94.
26. Results from population based prevalence surveys in Tanzania. [Data provided by Dr. Nkundwe Mwakyusa, Programme Manager, Ministry of Health Tanzania for the Global Atlas of Trachoma]
27. Vietnam Institute of Ophthalmology (2007) Vietnam national plan of blindness prevention and eye care towards "Vision 2020".
28. Hanutsaha P (2009) Eye Health Situation in Vietnam.
29. Wu M (2008) RAAB in Luliang County, China.
30. Muller A (2008) RAAB in Ethiopia.
31. Zerihun N, Mabey D (1997) Blindness and low vision in Jimma Zone, Ethopia: results of a population-based survey. Ophthalmic Epidemiol 4: 19-26.
32. Moll AC, van der Linden AJ, Hogeweg M, Schader WE, Hermans J, et al. (1994) Prevalence of blindness and low vision of people over 30 years in the Wenchi district, Ghana, in relation to eye care programmes. Br J Ophthalmol 78: 275-279.
33. Guzek JP, Anyomi FK, Fiadoyor S, Nyonator F (2005) Prevalence of blindness in people over 40 years in the Volta Region of Ghana. Ghana Med J 39: 55-62.
34. Faal H, Minassian DC, Dolin PJ, Mohamed AA, Ajewole J, et al. (2000) Evaluation of a national eye care programme: re-survey after 10 years. Br J Ophthalmol 84: 948-951.
35. National Program for Control of Blindness (2007) Rapid Assessment of Avoidable Blindness in India. New Delhi: Ministry of Health and Family Welfare.
36. Feghhi M, Khataminia G, Ziaei H, Latifi M (2009) Prevalence and causes of blindness and low vision in Khuzestan Province, Iran. J Ophthalmic Vis Res 4: 29-34.
37. Rajavi Z, Katibeh M, Ziaei H, Fardesmaeilpour N, Sehat M, et al. (2011) Rapid assessment of avoidable blindness in Iran. Ophthalmology 118: 1812-1818.
38. Mathenge W, Nkurikiye J, Limburg H, Kuper H (2007) Rapid assessment of avoidable blindness in Western Rwanda: blindness in a postconflict setting. PLoS Med 4: e217.
39. Rutzen AR, Ellish NJ, Schwab L, Graham PJ, Pizzarello LD, et al. (2007) Blindness and eye disease in Cambodia. Ophthalmic Epidemiol 14: 360-366.
40. Limburg H, Barria von-Bischhoffshausen F, Gomez P, Silva JC, Foster A (2008) Review of recent surveys on blindness and visual impairment in Latin America. Br J Ophthalmol 92: 315-319.
41. Durkin SR, Casson RJ, Newland HS, Aung TH, Shein WK, et al. (2007) Prevalence of trachoma-related trichiasis and corneal opacity in rural Myanmar: the Meiktila Eye Study. Ophthalmology 114: e7-11.
42. Abiose A, Murdoch I, Babalola O, Cousens S, Liman I, et al. (1994) Distribution and aetiology of blindness and visual impairment in mesoendemic onchocercal communities, Kaduna State, Nigeria. Kaduna Collaboration for Research on Onchocerciasis. Br J Ophthalmol 78: 8-13.
43. Abdull MM, Sivasubramaniam S, Murthy GV, Gilbert C, Abubakar T, et al. (2009) Causes of blindness and visual impairment in Nigeria: the Nigeria national blindness and visual impairment survey. Invest Ophthalmol Vis Sci 50: 4114-4120.
44. Gurung R (2008) RAAB in Bagmati and Janakpur zone, Nepal. Unpublished.

45. Dineen B, Bourne RR, Jadoon Z, Shah SP, Khan MA, et al. (2007) Causes of blindness and visual impairment in Pakistan. The Pakistan national blindness and visual impairment survey. Br J Ophthalmol 91: 1005-1010.
46. Chiang F, Kuper H, Lindfield R, Keenan T, Seyam N, et al. (2010) Rapid assessment of avoidable blindness in the Occupied Palestinian Territories. PLoS One 5: e11854.
47. Ngondi J, Ole-Sempele F, Onsarigo A, Matende I, Baba S, et al. (2006) Prevalence and causes of blindness and low vision in southern Sudan. PLoS Med 3: e477.

# Appendix 3.3: Age standardisation of prevalence data

## Overview

Age standardisation of data was required at three points in the analysis: i) to standardise matched prevalence estimates of trachomatous trichiasis (TT) and trachomatous blindness to ages ≥15 years to use in the regression model, ii) to standardise district/region level TT prevalence data to ages ≥15 years and iii) to generate age and sex-specific prevalence estimates of trachomatous blindness from model output, accounting for variability in the age-prevalence curve. The same methodology was used for each step, which accounted for the demographic structure of the country population at corresponding years, but the respective analyses were done in R (i. and ii.) and WinBUGS (iii.).

## Methods

*Age-prevalence curve*

Age stratified data on TT, trachomatous blindness and trachomatous low vision were extracted from published and unpublished data and used to model age-prevalence curves. Where possible, data were disaggregated into yearly intervals and otherwise age was entered as the median of each age category presented. An upper bound of 85 years was used to calculate the median for last prevalence category where the range not presented.

A grouped binomial generalised linear model was used to model the relationship between age and prevalence of each disease state in R, using a logit link. This followed the form:

$$Y \sim Binomial(n, \pi)$$

$$\pi = alpha + \beta_1 \times age \times \beta_2 \times age^2$$

Where $Y$ is the number of individuals with trachomatous blindness (or trichiasis or low vision), from $n$ observations with the probability $\pi$ of having trachomatous blindness (or trichiasis or low vision). The binomial probability $\pi$ was modelled as a function of *age* and $age^2$ in order to produce parameter estimates for the intercept (alpha), $\beta_1$ and $\beta_2$.

Post-hoc, the prevalence of blindness in individuals under the age of 15 years and the prevalence of low vision in individuals under the age of 5 years was set to be zero, based on a lack of cases even in hyper endemic areas [6].

*Age standardisation*

The resulting curve was used to calculate the proportion of the cumulative prevalence in each age category (here abbreviated as AR). Along with information on the proportion of the population in each category, this can then be used to estimate any age-specific prevalence by the following methodology. The age-prevalence curve was used to standardise input to population ≥15 years in R to export for WinBUGS models and also incorporated into the WinBUGS code to generate final age and sex-specific estimates with associated model uncertainty.

The prevalence in the total population is algebraically equal to the sum of the number of cases in each age group divided by the total population:

$$P_{Total} = \frac{\sum_A P_A \, x \, Pop_A}{Pop_{Total}}$$

where $P$ is the prevalence and $Pop$ the population in each of $A$ age groups. Each age-specific prevalence can be re-defined by using the ratio between the proportion of the prevalence in each age category, as estimated from the age-prevalence curve. In this example we will use age-categories used by Ranson and Evans, although in practice the

equations are adapted according to the age ranges of available prevalence estimates (5+, 50+, 45+, 40+, 35+, etc).

It follows that the age specific prevalence in each group can be defined by its relationship to another age group, as shown below:

$$P_{45-59} \times \frac{AR_{60plus}}{AR_{45-59}} = P_{60plus}$$

$$P_{60plus} \times \frac{AR_{45-59}}{AR_{60plus}} = P_{45-59}$$

Thus, the age-specific prevalence in all of $A$ age groups (45-59, 15-44, 5-15, 0-4 years) can be defined in relation to $P_{60plus}$:

$$P_{Total} = \frac{P_{60plus} \times Pop_{60plus} + \sum_{\mathrm{A}} \left( (P_{60plus} \times \frac{AR_{\mathrm{A}}}{AR_{60plus}}) \times Pop_{\mathrm{A}} \right)}{Pop_{Total}}$$

Which is equivalent to:

$$P_{Total} = \frac{P_{60plus} \times \left( Pop_{60plus} + \sum_{\mathrm{A}} \left( Pop_{\mathrm{A}} \times \frac{AR_{\mathrm{A}}}{AR_{60plus}} \right) \right)}{Pop_{Total}}$$

And the equation rearranged as needed to solve for $P_{Total}$ or $P_{60plus}$ :

$$P_{60plus} = \frac{P_{Total} \times Pop_{Total}}{Pop_{60plus} + \sum_{\mathrm{A}} \left( \frac{Pop_{\mathrm{A}} \times AR_{\mathrm{A}}}{AR_{60plus}} \right)}$$

Finally, the age-specific population estimates = $Pop_{Total} \times D_{Agegroup}$ where $D_{Agegroup}$ equals the proportion of the population in that age group. Thus, $Pop_{Total}$ cancels out of the right side of the equation, leaving:

$$P_{60plus} = \frac{P_{Total}}{D_{60plus} + \sum_{\mathrm{A}} \left( \frac{D_{\mathrm{A}} \times AR_{\mathrm{A}}}{AR_{60plus}} \right)}$$

The resulting prevalence estimate in this age group can then be used to calculate the age-specific prevalence in other age groups, as shown previously.

*Sex-specific prevalence estimates*

Similarly, sex-specific prevalence estimates are calculated for each age group (using 60plus as an example):

$$P_{60plus} \times Pop_{60plus} = P_{male} \times Pop_{60plus} \times D_{male} + P_{female} \times Pop_{60plus} \times D_{female}$$

As $P_{female}$ is equivalent to $P_{male} \times SexRatio$, we can substitute this into the equation and solve for $P_{male}$:

$$P_{male} = \frac{P_{60plus} \times Pop_{60plus}}{Pop_{60plus} \times D_{male} + SexRatio \times Pop_{60plus} \times D_{female}}$$

Which reduces to:

$$P_{male} = \frac{P_{60plus}}{D_{male} + SexRatio \times D_{female}}$$

Table A3.3.1 Comparison of the published and modelled estimates of the proportion of the burden (cumulative prevalence) of trachomatous blindness

| Published by Ranson & Evans (1996)[a] | Estimated from age-prevalence curve[b] |
|---|---|
| $AR_{60plus} = 80\ (71,90)$ | $AR_{60plus} = 74.1\ (56,98)$ |
| $AR_{45-59} = 18\ (9,23)$ | $AR_{45-59} = 24.5\ (18,32)$ |
| $AR_{15-44} = 2\ (1,4)$ | $AR_{15-44} = 1.0\ (0.5,1.9)$ |
| $AR_{5-14} = 0\ (0,1)$ | $AR_{5-14} = 0\ (0,0)$ |
| $AR_{10-4} = 0\ (0,1)$ | $AR_{0-4} = 0\ (0,0)$ |

AR: age distribution of the prevalence of trachomatous blindness

[a] Based on four studies with corresponding upper and lower bounds

[b] No model uncertainty was incorporated when adjusting input model data for the relationship between trichiasis and trachomatous blindness, but ARs were directly estimated within winbugs for final age-specific prevalence estimates and so incorporated variability in the age-prevalence curve.

## References


1. Dolin PJ, Faal H, Johnson GJ, Ajewole J, Mohamed AA, et al. (1998) Trachoma in The Gambia. Br J Ophthalmol 82: 930-933.
2. Kolaczinski J (unpublished data) Population-based prevalence surveys in Western Equatoria, South Sudan.
3. Madani MO, Huguet P, Mariotti SP, Dezoumbe D, Tosi C, et al. (2003) [Trachoma in Chad: results of an epidemiological survey]. Sante 13: 9-15.
4. Melese M, Alemayehu W, Bayu S, Girma T, Hailesellasie T, et al. (2003) Low vision and blindness in adults in Gurage Zone, central Ethiopia. Br J Ophthalmol 87: 677-680.
5. Mesfin MM, de la Camera J, Tareke IG, Amanual G, Araya T, et al. (2006) A community-based trachoma survey: prevalence and risk factors in the Tigray region of northern Ethiopia. Ophthalmic Epidemiol 13: 173-181.
6. Ngondi J, Reacher M, Matthews F, Ole-Sempele F, Onsarigo A, et al. (2007) The epidemiology of low vision and blindness associated with trichiasis in southern Sudan. BMC Ophthalmol 7: 12.
7. Noa Noatina, B., et al. (2013). "Prevalence of trachoma in the far north region of Cameroon: results of a survey in 27 health districts." PLoS Negl Trop Dis **7**(5): e2240.
8. Regassa K, Teshome T (2004) Trachoma among adults in Damot Gale District, South Ethiopia. Ophthalmic Epidemiol 11: 9-16.
9. Saal MB, Schemann JF, Saar B, Faye M, Momo G, et al. (2003) [Trachoma in Senegal: results of a national survey]. Med Trop (Mars) 63: 53-59.
10. Sahlu T, Larson C (1992) The prevalence and environmental risk factors for moderate and severe trachoma in southern Ethiopia. J Trop Med Hyg 95: 36-41.
11. Schemann JF, Sacko D, Banou A, Bamani S, Bore B, et al. (1998) [Cartography of trachoma in Mali: results of a national survey]. Bull World Health Organ 76: 599-606.
12. Solomon A (unpublished data) Community census survey in Kahe, Tanzania.
13. Tabbara KF, Ross-Degnan D (1986) Blindness in Saudi Arabia. JAMA 255: 3378-3384.
14. Wondimu A, Bejiga A (2003) Prevalence of trachomatous trichiasis in the community of Alaba District, Southern Ethiopia. East Afr Med J 80: 365-368.

# Appendix 3.4: Summary of prevalence data used in modelling

Table A.3.4.1 Data availability from national trachomatous blindness surveys (TB), Rapid Avoidable Blindness Surveys (RAAB) and population-based district-level trichiasis (TT) surveys including years of surveys, geographical coverage of prevalence data and proportion of districts classed "non-Endemic" based on TRA or anecdotal evidence. TT data include district or regional level data.

| GBD Region | Country | ISO | 1990 | 2010 | Low Vision | Notes | Trachoma Control |
|---|---|---|---|---|---|---|---|
| North Africa / Middle East | Algeria | ALG | No data: assigned 0 or regional average | No data: assigned 0 or regional average | Modelled | No TT data | |
| Sub-Saharan Africa, West | Benin | BEN | TB: 1990<br>Coverage: 100%<br>Non-endemic: 0% | TB: 1990<br>Coverage: 100%<br>Non-endemic: 0% | Yes | No TT data | |
| Sub-Saharan Africa, Southern | Botswana | BWA | RAAB: 1997<br>Coverage: 100%<br>Non-endemic: 0% | RAAB: 1997<br>Coverage: 100%<br>Non-endemic: 0% | Modelled | No TT data | |
| Sub-Saharan Africa, West | Burkina Faso | BFA | TT: 1997<br>Coverage: 59%[a]<br>Non-endemic: 0% | TT: 2007-2012<br>Coverage: 100%<br>Non-endemic: 0%<br>Post-intervention | Modelled | | SAFE start date: 2007 |
| Sub-Saharan Africa, East | Burundi | BDI | TF data only<br>Coverage: 0%<br>Non-endemic: 100% | TF data only<br>Coverage: 0%<br>Non-endemic: 100% | Modelled | Set to zero as only four districts classed as endemic for TF (2009-2010) | SAFE start date: 2011 |
| Sub-Saharan Africa, West | Cameroon | CMR | TT: 2008-2011<br>Coverage: 23%<br>Non-endemic: 42% | TT: 2008-2011<br>Coverage: 23%<br>Non-endemic: 42% | Modelled | Southern Cameroon is classified non-endemic (ie set to zero) | SAFE start date: 2011 |
| Sub-Saharan Africa, Central | CAR | CAF | TT: 2008-2011<br>Coverage: 52.9%<br>Non-endemic: 0% | TT: 2008-2011<br>Coverage: 52.9%<br>Non-endemic: 0% | Modelled | | |
| Sub-Saharan Africa, West | Chad | TCD | TB: 1985<br>Coverage: 100%<br>Non-endemic: 0% | TT: 2002<br>Coverage: 57.1%<br>Non-endemic: 0% | Modelled | Also had TT data from 2002 for 57.1% of districts and 7% of districts were classified as non-endemic. | |
| Sub-Saharan Africa, West | Cote d'Ivoire | CIV | TT: 2008<br>Coverage: 10.3%<br>Non-endemic: 72% | TT: 2008<br>Coverage: 10.3%<br>Non-endemic: 72% | Modelled | Only northern areas suspected endemic, southern areas excluded (ie set to zero) | |
| North Africa / Middle East | Djibouti | DJI | TT: 1985<br>Coverage: 36.4%<br>Non-endemic: 0% | No data: assigned 0 or regional average | Modelled | | Major SES change, classed as suspected endemic now. |

Table continued on next page

Table A3.4.1 continued

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| North Africa / Middle East | Egypt | EGY | TT: 1999-2002<br>Coverage: 7.7%<br>Non-endemic: 0% | TT: 1999-2002<br>Coverage: 7.7%<br>Non-endemic: 0% | Modelled | | |
| Sub-Saharan Africa, East | Eritrea | ERI | RAAB: 2008<br>Coverage: 100%<br>Non-endemic: | RAAB: 2008<br>Coverage: 100%<br>Non-endemic: | Yes | Also have TT data for 62.1% of the country from 2006. | SAFE start date: 2010 |
| Sub-Saharan Africa, East | Ethiopia | ETH | TB: 2005<br>Coverage: 100%<br>Non-endemic: | TB: 2005<br>Coverage: 100%<br>Non-endemic: | Yes | Further to 2005 national survey, also have TT data for 100% of the country at regional, wereda and kebele-levels from 1999-2012. | SAFE start date: 2003 |
| Sub-Saharan Africa, West | Ghana | GHA | TT: 1997-2003<br>Coverage: 18.8%<br>Non-endemic: 75% | TT: 2007-2008<br>Coverage: 24.5%<br>Non-endemic: 75%<br>Post-intervention | Modelled | | SAFE start date: 1999 |
| Sub-Saharan Africa, West | Guinea | GIN | TT: 2001-2002<br>Coverage: 39.5%<br>Non-endemic: 0% | TT: 2001-2011<br>Coverage: 39.5%<br>Non-endemic: 0% | Modelled | | |
| Sub-Saharan Africa, West | Guinea Bissau | GNB | TT: 2005<br>Coverage: 100%<br>Non-endemic: | TT: 2005<br>Coverage: 100%<br>Non-endemic: | Modelled | | SAFE start date: 2009 |
| Sub-Saharan Africa, East | Kenya | KEN | TT: 2004-2012<br>Coverage: 17.3%<br>Non-endemic: 75.4% | TT: 2004-2012<br>Coverage: 17.3%<br>Non-endemic: 75.4% | Modelled | | SAFE start date: 2007 |
| Sub-Saharan Africa, East | Malawi | MWI | TT: 2008-2012<br>Coverage: 9.4%<br>Non-endemic: 15.6% | TT: 2008-2012<br>Coverage: 9.4%<br>Non-endemic: 15.6% | Modelled | | SAFE start date: 2011 |
| Sub-Saharan Africa, West | Mali | MLI | TT: 1996<br>Coverage: 100%<br>Non-endemic: | TT: 2008-2010<br>Coverage: 100%<br>Non-endemic:<br>Post-intervention | Modelled | | SAFE start date: 2000 |
| Sub-Saharan Africa, West | Mauritania | MRT | TT: 2000-2005<br>Coverage: 100%<br>Non-endemic: | TT: 2004-2011<br>Coverage: 67.4%<br>Non-endemic: 32.6% | Modelled | | SAFE start date: 2004 |
| North Africa / Middle East | Morocco | MAR | TB: 1992<br>Coverage: 100%<br>Non-endemic: | TT: 2003<br>Coverage: 8.7%<br>Non-endemic: 91.3% | Modelled | | SAFE start date: 1999 |

Table continued
Table A3.4.1 continued

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Sub-Saharan Africa, Southern | Mozambique | MOZ | No data: assigned 0 or regional average | No data: assigned 0 or regional average | Modelled | No TT data | |
| Sub-Saharan Africa, West | Niger | NER | TT: 1997-1999<br>Coverage: 100%<br>Non-endemic: | TT: 2007-2011<br>Coverage: 76.7%<br>Non-endemic: 7.0%<br>Post-intervention | Modelled | | SAFE start date: 2002 |
| Sub-Saharan Africa, West | Nigeria | NGA | RAAB: 2010<br>Coverage: 100%<br>Non-endemic: | RAAB: 2010<br>Coverage: 100%<br>Non-endemic: | Yes | Also TT data for 22.6% of the country from 2000-2011; 52% classified non-endemic | SAFE start date: 2010 |
| Sub-Saharan Africa, West | Senegal | SEN | TT: 2000-2003<br>Coverage: 73.5%<br>Non-endemic: 0% | TT: 2000-2003<br>Coverage: 73.5%<br>Non-endemic: 0% | Modelled | | SAFE start date: 2004 |
| Sub-Saharan Africa, East | Somalia | SOM | No data: assigned 0 or regional average | No data: assigned 0 or regional average | Modelled | No TT data | |
| Sub-Saharan Africa, East | South Sudan | SSD | TT: 1999-2009<br>Coverage: 17.2%<br>Non-endemic: 0% | TT: 1999-2009<br>Coverage: 17.2%<br>Non-endemic: 0% | Modelled | | |
| Sub-Saharan Africa, East | Sudan | SDN | TT: 2005-2010<br>Coverage: 61.3%<br>Non-endemic: 0% | TT: 2005-2010<br>Coverage: 61.3%<br>Non-endemic: 0% | Modelled | | SAFE start date: 1999 |
| Sub-Saharan Africa, East | Tanzania | TZA | TT: 2001-2006<br>Coverage: 45.8%<br>Non-endemic: 8.3% | TT: 2004-2009<br>Coverage: 45.8%<br>Non-endemic: 8.3% | Modelled | A few districts have post-intervention data | SAFE start date: 1999 |
| Sub-Saharan Africa, West | The Gambia | GMB | TB: 1986<br>Coverage: 100%<br>Non-endemic: | TT: 2006-2009<br>Coverage: 90.7%<br>Non-endemic: 9.3%<br>Post-intervention | Modelled | | SAFE start date: 2006 |
| Sub-Saharan Africa, West | Togo | TGO | TT: 2012<br>Coverage: 93.3%<br>Non-endemic: 6.7% | TT: 2012<br>Coverage: 93.3%<br>Non-endemic: 6.7% | Modelled | | |
| Sub-Saharan Africa, East | Uganda | UGA | TT: 2006-2010<br>Coverage: 31.3%<br>Non-endemic: 59.8% | TT: 2006-2010<br>Coverage: 31.3%<br>Non-endemic: 59.8% | Modelled | | SAFE start date: 2007 |
| Sub-Saharan Africa, East | Zambia | ZMB | TT: 1985-2012<br>Coverage: 36.9%<br>Non-endemic: 0% | TT: 1985-2012<br>Coverage: 36.9%<br>Non-endemic: 0% | Modelled | | SAFE start date: 2010 |

TB: trachomatous blindness; LV: low vision; TT: trachomatous trichiasis; TF: Trachomatous inflammation, follicular

# Appendix 4.1: Climatic and environmental data

Table A4.1.1 Description of climatic and environmental variables

| Variable | Description and source |
|---|---|
| **Climate** | |
| Mean annual temperature (°C) | Interpolation of average monthly mean annual temperature data (~1950-2000)<br>Derived: Meteorological stations<br>Resolution: 2.5 arc-minute (~5 km)<br>Source: WorldClim BioClim variables |
| Maximum warmest temperature (°C) | Interpolation of average maximum temperature in the warmest month (~1950-2000)<br>Derived: Meteorological stations<br>Resolution: 2.5 arc-minute (~5 km)<br>Source: WorldClim BioClim variables |
| Mean annual precipitation (mm) | Interpolation of average monthly mean annual precipitation data (~1950-2000)<br>Derived: Meteorological stations<br>Resolution: 2.5 arc-minute (~5 km)<br>Source: WorldClim BioClim variables |
| Precipitation of driest month (mm) | Interpolation of average monthly mean annual precipitation data (~1950-2000)<br>Derived: Meteorological stations<br>Resolution: 2.5 arc-minute (~5 km)<br>Source: WorldClim BioClim variables |
| Annual aridity index | Interpolation of mean Annual Precipitations/Mean Annual Potential Evapo-Transpiration (~1950-2000).<br>Derived: Meteorological stations (WorldClim)<br>Resolution: 30 arc second (~1 km)<br>Source: Consortium for Spatial Information (CGIAR-CSI) Global-Aridity and Global-PET Database [1,2] |
| Monthly average potential evapo-transpiration (PET) (mm/month) | Interpolation of monthly average measure of the amount of evaporation expected if a sufficient water source were available. Hargreaves model using monthly average geo-datasets of: mean temperature, daily temperature range and extra-terrestrial radiation (~1950-2000)<br>Derived: Meteorological stations (WorldClim)<br>Resolution: 30 arc second (~1 km)<br>Source: Consortium for Spatial Information (CGIAR-CSI) Global-Aridity and Global-PET Database [1,2] |
| **Environmental** | |
| Land surface temperature (LST) (°C) | Mean annual land surface temperature for the years 2005-2007.<br>Derived: Satellite remote sensing<br>Resolution: 2.5 arc-minute (~5 km)<br>Source: Moderate Resolution Imaging Spectroradiometer (MODIS) on NASA's Terra satellite |

Table continued

Table A4.1.1 continued

| | |
|---|---|
| Altitude (meters) | Elevation data<br>Derived: Radar<br>Resolution: 2.5 arc-minute (~5 km)<br>Source: Shuttle Radar Topography Mission (SRTM) |
| Enhanced vegetation index (EVI) | Index of the vegetation signal from surface reflectance.<br>Derived: Satellite remote sensing<br>Resolution: 2.5 arc-minute (~5 km)<br>Source: Moderate Resolution Imaging Spectroradiometer (MODIS) on NASA's Terra satellite |
| Global land cover classification[a] | Global land cover classification<br>Derived: Satellite remote sensing<br>Resolution: 2.5 arc-minute (~5 km)<br>Source: UN *Land Cover Classification* System (LCCS) using ENVISAT satellite mission's MERIS sensor at 5km$^2$ resolution |
| Ruminant density (animals per 5km cell) | Predicted distribution of livestock in 2005.<br>Derived: Observed livestock statistics and environmental variables<br>Resolution: 3 arc-minute (~5 km)<br>Source: FAO Global Livestock Densities [3] |
| Cost-distance to road network | Accessibility measure calculated as a cost-distance surface using a set of topographical variables set as constraints to access of the road network.<br>Derived: Distance to road network, slope, major water bodies, streams and land cover<br>Resolution: 30 arc second (~1 km)<br>Source: Generated for this analysis |
| Distance to river or water body | Derived: Distance to nearest river or surface water body<br>Resolution: 2.5 arc-minute (~5 km)<br>Source: FAO Rivers and Surface Water Bodies database |
| Urban classification | Gridded database of urban settlements with populations greater than 1000 persons.<br>Derived: Satellite night-light data and gridded population data<br>Resolution: 30 arc second (~1 km)<br>Source: Global Rural-Urban Mapping Project (GRUMP) |
| Population density | Gridded population data for the year 2010<br>Derived: NA<br>Resolution: 30 arc second (~1 km)<br>Source: SEDAC's Gridded Population of the World, Version 3 data set (GPWv3) |

m: meters; mm: millimetres; C: Celsius; km: kilometers

[a]Classified as binary variable indicating savannah/grasslands

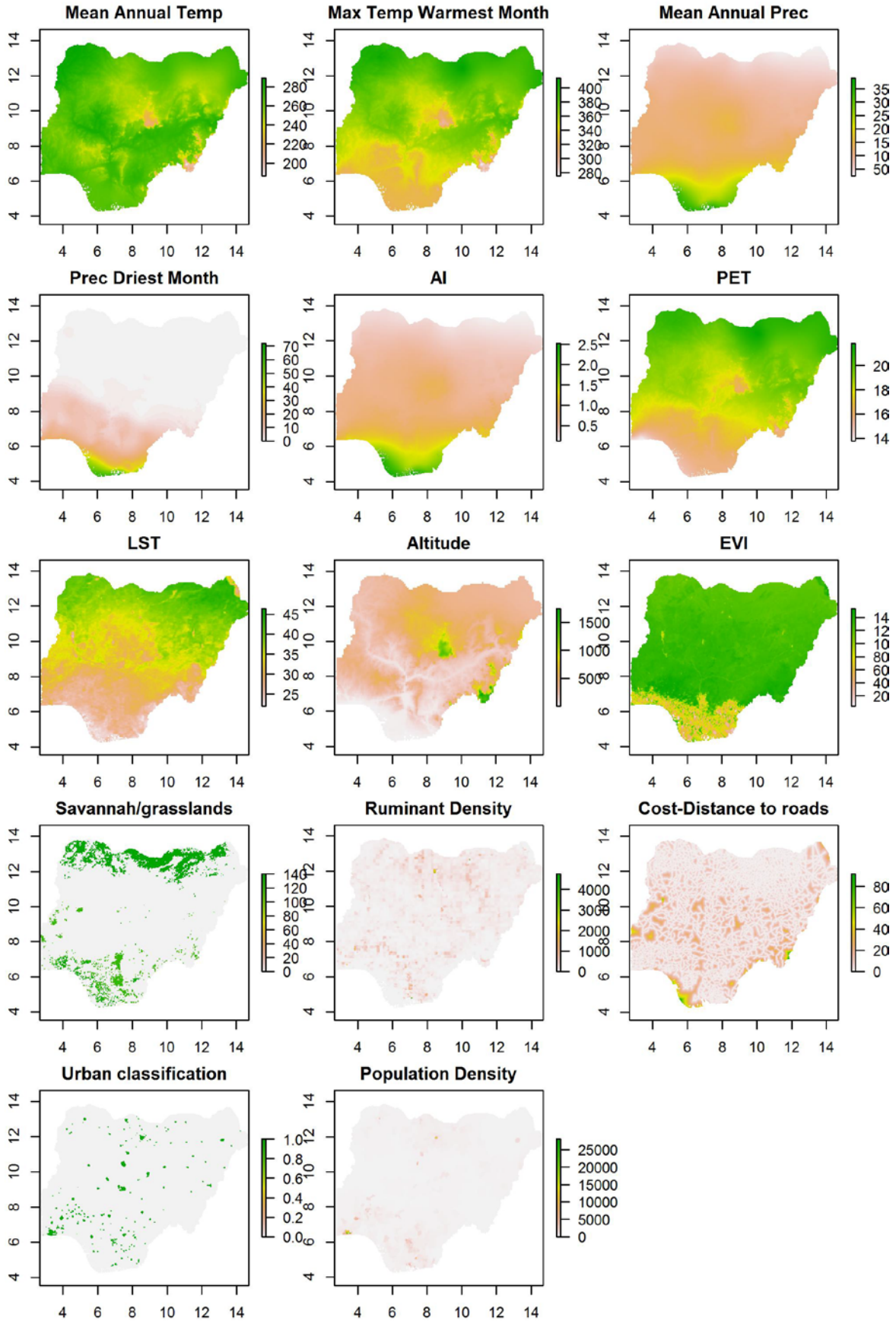



Figure A4.1.1 Maps of climatic and environmental factors within in Nigeria, as detailed in Table A4.1.1. AI: aridity index; PET: potential evapo-transpiration; LST: land-surface temperature; EVI: enhanced vegetation index; Prec: precipitation

Table A4.1.2 Summary statistics for each climate and environmental measure from 304 clusters in Nigeria

| Variable | Median (range) | SD |
|---|---|---|
| **Climate** | | |
| Land surface temperature (LST) (°C) | 31.7 (22.5, 42.8) | 5.0 |
| Mean annual temperature (°C) | 26.4 (21.8, 28.7) | 1.0 |
| Average maximum temperature in the warmest month (°C) | 34.6 (30.9, 41.0) | 2.7 |
| Mean annual precipitation (mm) | 1284.0 (407.0, 3833) | 639.2 |
| Average precipitation in driest month (mm) | 6.0 (0.0, 67.0) | 10.3 |
| Annual aridity index | 0.75 (0.19, 2.46) | 0.45 |
| Annual potential evapo-transpiration (PET) | 1738.5 (1385.0, 2157.0) | 193.9 |
| **Environmental** | | |
| Altitude (m) | 270.5 (4.0, 1287.0) | 226.1 |
| Enhanced vegetation index (EVI) | 1218.0 (88.1, 1500.8) | 270.8 |
| Global land cover classification[a] | 17.0 % | - |
| Ruminant density (animals per 5km cell) | 68.1 (0, 1051.4) | 144.8 |
| Cost-distance to road network | 1387.5 (0, 22842.7) | 2983.3 |
| Urban classification | 27.3 % urban | - |
| Population density | 285.0 (8.0, 27982.0) | 2570.4 |

m: meters; mm: millimetres; C: Celsius

[a]Classified as binary variable indicating savannah/grasslands

# Appendix 4.2: Principal components analysis & environmental variable selection

## Background

Collinearity is a special case of model non-identifiability, in which two or more highly correlated variables are associated with an outcome, causing inflation of the variance of regression parameters so that the "true" predictor cannot be identified without further information [1,2]. Multicollinearity is a common problem encountered developing models using environmental variables [3], and there are a number of approaches in ecological analyses to deal with resulting statistical instability. The simplest approach is to drop all but one collinear variable and assume that the retained variable captures functionally important variation [2]. However, in the absence of a strong evidence base to support this selection process, choice of the retained variable may be arbitrary and miss potentially important sources of variation in the dataset. While the objective of a principal components analsysis (PCA) is often to reduce dimensionality of a dataset by replacing correlated variables with a smaller number of uncorrelated variables, this approach can be used to help interpret the underlying structure of climatic variables, in terms of variance and covariance, and inform subsequent model building strategies [4].

Multicollinearity was found to be present between all included climatic variables in Nigeria to varying degrees, with particularly high (above 0.7) correlation between subsets of climatic indices related to precipitation and evaporation, and temperature and altitude. This appendix outlines the PCA conducted to support environmental variable selection in Nigeria by characterising patterns of variation in the data to support selection of a single variable from each identified group. Correlated predictors were then added into the model through sequential regression after identification of a sequence of ecological

importance, based on existing studies, in order to calculate the independent contribution of each explanatory variable.

## Methods

*Data*

Climatic indices were selected as described in the text, with a number of interpolated and satellite derived measures used to capture annual mean indices (including temperature, precipitation and evaporation trends) and potential environmental extremes (precipitation in the driest month and temperature in the hottest month) (Appendix 4.1). Enhanced vegetation index (EVI) and altitude were also included in the PCA, as they are often used as proxy measures for rainfall and temperature, and a lower prevalence of trachoma has been associated with higher altitudes in a number of trachoma studies [5-7]. These proxy measures are broad indicators, however, and may also be associated with other environmental characteristics (for example land cover and land use).

*Analysis*

Principal components (PC) were estimated using the function 'prcomp' in the "stats" package in R. This function performs a principal components analysis using a singular value decomposition on the data matrix, which is conceptually equivalent to doing an eigenvalue decomposition on the covariance matrix but numerically more stable [8]. In the PCA, variables are re-expressed through linear combinations (rotations) of the original variables through the decomposition process. This is accomplished by generating a covariance matrix and identifying groups of variables, such that each group is internally correlated but less correlated with other variables. Eigenvectors (analogous to singular values) are calculated that are perpendicular to observed associations and proportional to

the amount of variance explained by the principal component, which is subsequently calculated from these lines and the original variables. Each component thus represents an independent dimension of variability and together sum to the total variation in the variables [9].

Principal components that explained a significant proportion of the variance in the data were identified using the Cattell scree test and the Kaiser rule, which respectively identify a point of inflection on a scree plot and eigenvectors greater than one [10,11]. Each component was then associated with the outcome using a generalised linear model with binomial logit link, adjusted for clustering. Patterns in climatic data were investigated by 1) examination of which covariates were heavily loading and the relationship (positive or negative) of the PC with the outcome, 2) plots of the rotations to visually identify groupings exerting similar effects and 3) plots of the component values, to allow identification of how PCs work together to aggregate high risk (TT/CO≥1%) clusters.

Together with evidence from the literature, this analysis informed selection of a single climate covariate from each observed grouping. Variables from highly correlated and contrasting groups were then included in the model using sequential regression in order to identify independent variation associated with explanatory variables.

## Results

There was high correlation between a number of included climatic variables, with linear relationships tending to become more variable at lower land surface temperatures and moderate meteorological temperatures (Figure A4.2.1, Table A4.2.1). The majority (87.4%) of the total variation was captured by the first two PCs, which had eigenvalues greater than one (Table A4.2.2) and were below the elbow of the scree plot (Figure A4.2.2). All PCs were associated with the cluster level risk of TT/CO through regression.

Loadings for each component are displayed in Table A4.2.2 and plots of the loadings for the first three components presented in Figure A4.2.3. The first principal component explained 67% of the variance, and suggested a contrasting relationship between two internally correlated groups of climatic indicators with similar magnitudes of loadings. The first group included climatic indices related to rainfall: annual mean precipitation, precipitation in the driest month and aridity. The opposing group may also be thought of as related to water availability, and included indicators related to evaporation: land-surface temperature, the maximum temperature in the warmest month, and the potential for evapo-transpiration (PET). Altitude and mean annual temperature were internally correlated in the dataset with a contrasting relationship, but not highly correlated with variables in the above two groups. These variables loaded heavily on the second component, which accounted for 20.4% of the total variation. EVI did not group clearly with the other variables, although it was correlated with both meteorological temperature and rainfall. This suggests that the additional variation that it contributes to the third principal component, along with LST, may be related to landcover or urban extents, which was investigated during the modelling process.

The ability of the first three components to aggregate clusters of higher risk (TT/CO ≥ 1%) is displayed in Figure A4.2.4, and suggests that variation in these components is important to include in the model.

## Variable selection

Mean annual precipitation and land surface temperature were retained from the two groups identified from the first component for further analyses, based on previous studies of trachoma. As the literature provides the strongest evidence base for an association between precipitation and trachoma, mean annual precipitation was considered the

principal climatic factor in the regression and represents the shared variation observed in the first component [12-14].

Land-surface temperature was selected from the second group because it is satellite derived and may be more reliable than interpolated weather station data. LST was then regressed against precipitation and residuals included as new variables that are conditional on precipitation.

Mean annual temperature was prioritised from the second component as 1) it had the highest loading on the second component, 2) it was considered to be a more proximal risk factor than altitude, and 3) altitude may be associated with other environmental characteristics, like land cover. Any residual effect of altitude was checked by including it in the model after regressing it against temperature.

EVI was regressed against the collinear variables from the first grouping (mean annual precipitation and the residuals of LST) to assess whether it accounted for any residual variation in the model which may be captured by the third component.

## Discussion

The results from the exploratory PCA allowed identification of four key contrasting groups in the first two principal components, which explained approximately 87% of the total variation. The first PC accounted for the most variation and might be interpreted as variation common to different measures of water availability which formed two contrasting groups related to precipitation and evaporation measures. Where mean annual precipitation was very low, average precipitation in the driest month tended to plateau. As expected, aridity was highly collinear with mean annual precipitation, which is directly used to calculate the index. Similarly, LST and monthly temperature indices are commonly used to construct measures of evapo-transpiration, so a degree of collinearity is expected. In this case, PET was generated from meteorological data that were not collinear

(including mean annual temperature and daily temperature range) and extra terrestrial radiation (or solar radiation). Thus, it may be that this radiation accounts for the variation PET shares with LST in the first component.

The second set of variables identified in the PCA were altitude and mean annual temperature. These variables were not collinear with the groups from the first PC and so were included separately in subsequent model building. Temperature is hypothesized to have an effect on the life span of the punitive fly vector, *Musca sorbens*, which has shown to live from 12 days at 32° C to 35 days at 24°C [3,15].

Unique variation in LST, after accounting for variation captured by the first component, was distinguished in the third component along with EVI. These two indices are often (inversely) correlated due to reduced spectral emissivity associated with less vegetated areas corresponding to higher temperatures [16].

## References

1. Gelman A, Hill J (2007) Data analysis using regression and mutilevel/hierarchical models: Cambridge University Press.
2. Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, et al. (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36: 027-046.
3. Koukounari A, Toure S, Donnelly CA, Ouedraogo A, Yoda B, et al. (2011) Integrated monitoring and evaluation and environmental risk factors for urogenital schistosomiasis and active trachoma in Burkina Faso before preventative chemotherapy using sentinel sites. BMC Infect Dis 11: 191.
4. Bartholomew DJ, Steele F, Moustaki I, Galbraith JI (2008) Chapter 5: Prinicpal Components Analysis. In: A G, editor. Analysis of multivariate social science data. Boca Raton: Taylor & Francis Group, LLC.
5. Baggaley RF, Solomon AW, Kuper H, Polack S, Massae PA, et al. (2006) Distance to water source and altitude in relation to active trachoma in Rombo district, Tanzania. Trop Med Int Health 11: 220-227.
6. Alemayehu W, Melese M, Fredlander E, Worku A, Courtright P (2005) Active trachoma in children in central Ethiopia: association with altitude. Trans R Soc Trop Med Hyg 99: 840-843.
7. Haileselassie T, Bayu S (2007) Altitude-a risk factor for active trachoma in southern Ethiopia. Ethiop Med J 45: 181-186.
8. Team RDC (2011) R: A language and environment for statistical computing.: R Foundation for Statistical Computing.

9. Shlens J (2009) A tutorial on principal component analysis. New York City: Center for Neural Science, New York University
10. Cattell RB (1966) The Scree Test for the Number of Factors. Multivariate Behavioral Research 1: 245-276.
11. Kaiser HF (1960) The application of electronic computers in factor analysis. Educational and Psychological Measurement 20: 141-151.
12. Clements AC, Kur LW, Gatpan G, Ngondi JM, Emerson PM, et al. (2010) Targeting trachoma control through risk mapping: the example of Southern Sudan. PLoS Negl Trop Dis 4: e799.
13. Ramesh A, Kovats S, Haslam D, Schmidt E, Gilbert CE (2013) The impact of climatic risk factors on the prevalence, distribution, and severity of acute and chronic trachoma. PLoS Negl Trop Dis 7: e2513.
14. Hagi M, Schemann JF, Mauny F, Momo G, Sacko D, et al. (2010) Active trachoma among children in Mali: Clustering and environmental risk factors. PLoS Negl Trop Dis 4: e583.
15. Goulson D, Derwent LC, Hanley ME, Dunn DW, Abolins SR (2005) Predicting calyptrate fly populations from the weather, and probable consequences of climate change. Journal of Applied Ecology 42: 795-804.
16. Goetz SJ, Prince SD, J S Advances in Satellite Remote Sensing of Environmental Variables for Epidemiological Applications. In: Hay SI, Randoph SE, Rogers DJ, editors. Advances in Parasitology: Remote Sensing and Geographical Systems in Epidemiology. London: Academic Press. pp. 289-307.

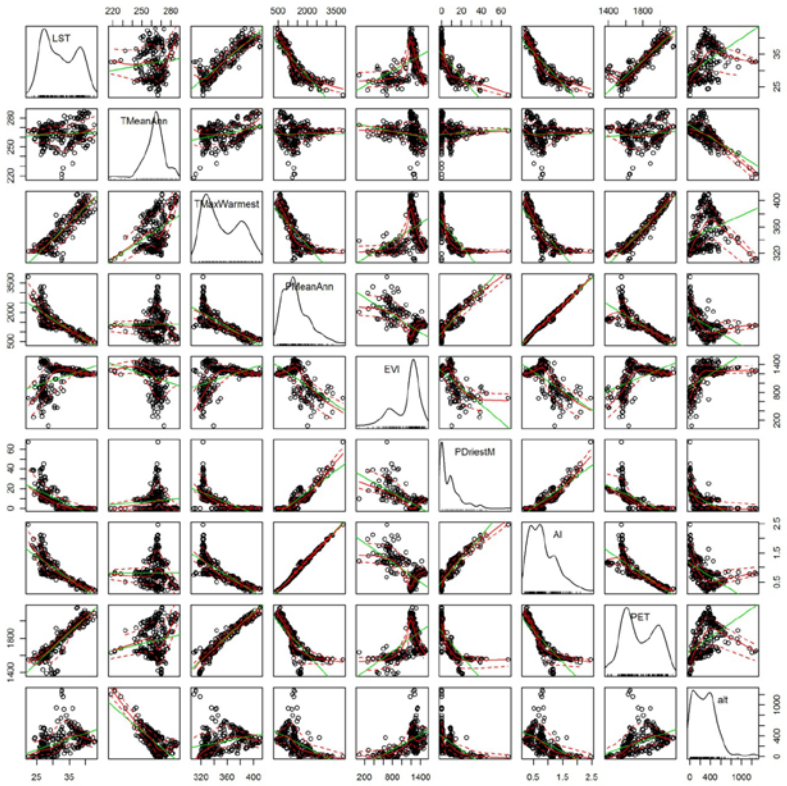



Figure A4.2.2 Scatter matrix of climatic and environmental indices

Table A4.2.1 Pearson correlation matrix between climatic and environmental indices

| Variable | LST | Mean annual rainfall | Mean annual temperature | Maximum temperature warmest month | Precipitation in driest month | EVI | Aridity index | PET | Altitude |
|---|---|---|---|---|---|---|---|---|---|
| LST | 1.000 | -0.852 | 0.105 | 0.904 | -0.735 | 0.422 | -0.861 | 0.910 | 0.474 |
| Mean annual rainfall | -0.852 | 1.000 | -0.001 | -0.799 | 0.897 | -0.646 | 0.995 | -0.814 | -0.545 |
| Mean annual temperature | 0.105 | -0.001 | 1.000 | 0.362 | 0.089 | -0.248 | 0.002 | 0.136 | -0.711 |
| Maximum temperature in warmest month | 0.904 | -0.799 | 0.362 | 1.000 | -0.697 | 0.446 | -0.825 | 0.965 | 0.291 |
| Precipitation in driest month | -0.735 | 0.897 | 0.089 | -0.697 | 1.000 | -0.681 | 0.916 | -0.749 | -0.606 |
| EVI | 0.422 | -0.646 | -0.248 | 0.446 | -0.681 | 1.000 | -0.678 | 0.549 | 0.578 |
| Aridity index | -0.861 | 0.995 | 0.002 | -0.825 | 0.916 | -0.678 | 1.000 | 0.851 | -0.566 |
| PET | 0.910 | -0.814 | 0.136 | 0.965 | -0.749 | 0.549 | -0.851 | 1.000 | 0.481 |
| Altitude | 0.474 | -0.545 | -0.711 | 0.291 | -0.606 | 0.578 | -0.566 | 0.481 | 1.000 |

LST: land surface temperature; EVI: enhanced vegetation index; PET: potential evapo-transpiration

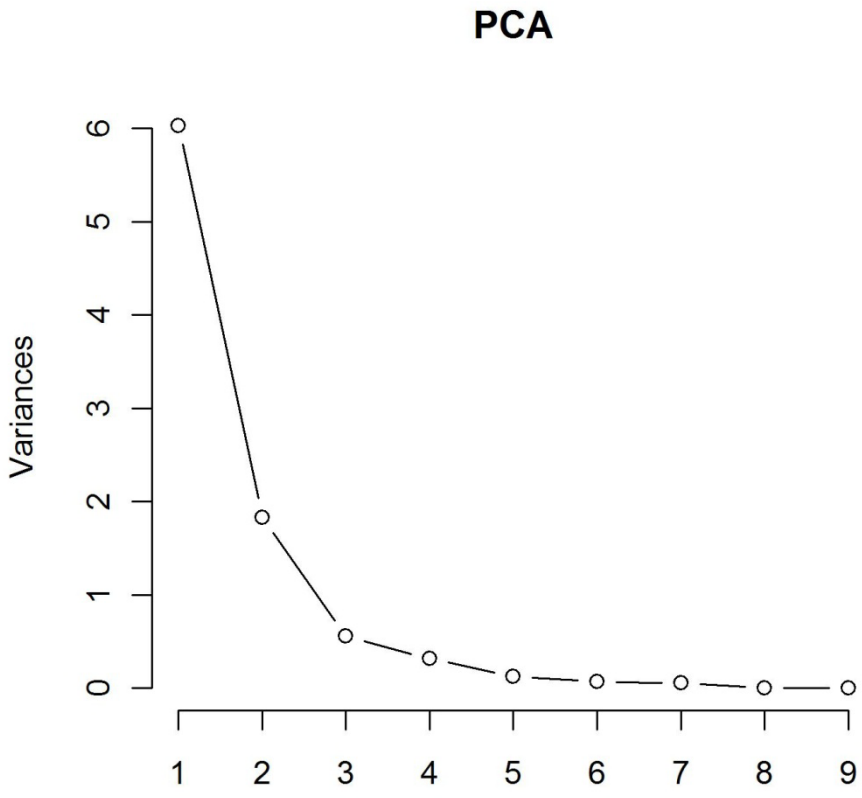



Figure A4.2.3 Scree plot of the variances of the principal components. The “elbow” of the plot appears near the third component, suggesting the first two components contribute the most.

Table A4.2.2 Factor loadings of the principal components indicating correlation between observed variables and specific components

| | Loadings for Principal Components | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Variable** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** |
| LST | -0.368 | -0.155 | 0.401 | 0.029 | 0.353 | -0.196 | 0.715 | -0.059 | -0.013 |
| Mean Annual Rainfall | 0.389 | 0.021 | 0.089 | -0.001 | -0.239 | 0.517 | 0.245 | -0.190 | 0.046 |
| Mean Annual Temperature | 0.014 | -0.712 | -0.264 | 0.303 | -0.117 | 0.001 | -0.237 | 0.743 | -0.184 |
| Maximum Temperature in warmest month | -0.356 | -0.318 | 0.156 | 0.407 | -0.402 | -0.231 | 0.252 | -0.125 | -0.617 |
| Precipitation in driest month | 0.371 | -0.075 | 0.206 | 0.476 | 0.170 | -0.035 | 0.213 | 0.031 | 0.017 |
| EVI | -0.287 | 0.252 | -0.742 | 0.470 | 0.648 | 0.363 | -0.186 | -0.060 | -0.052 |
| Aridity Index | 0.397 | 0.017 | 0.099 | 0.301 | -0.243 | -0.133 | 0.260 | 0.278 | 0.721 |
| PET | -0.376 | -0.158 | 0.211 | 0.432 | -0.191 | -0.236 | -0.375 | -0.557 | 0.245 |
| Altitude | -0.258 | 0.522 | 0.300 | 0.133 | -0.314 | 0.659 | 0.141 | -0.021 | 0.006 |
| Eigenvalues | 6.031 | 1.833 | 0.557 | 0.317 | 0.129 | 0.070 | 0.057 | 0.004 | 0.001 |
| Cum. proportion of variance | 0.670 | 0.874 | 0.936 | 0.971 | 0.985 | 0.993 | 0.999 | 0.999 | 1.000 |

LST: land surface temperature; EVI: enhanced vegetation index; PET: potential evapo-transpiration

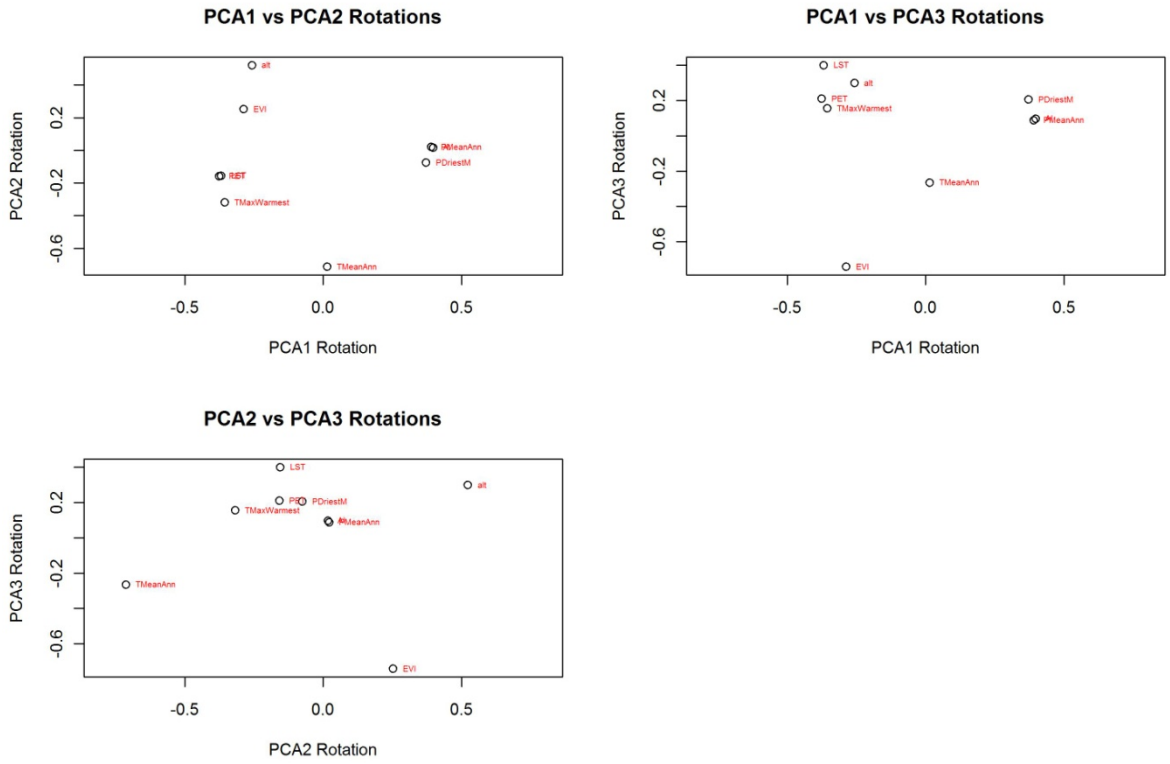



Figure A4.2.4 Rotated data for the first three components, contrasting groups which have high within-correlation along each component. Data are rotated so that the eigenvectors from each component are the axes, which effectively distributes variability more equally between factors.

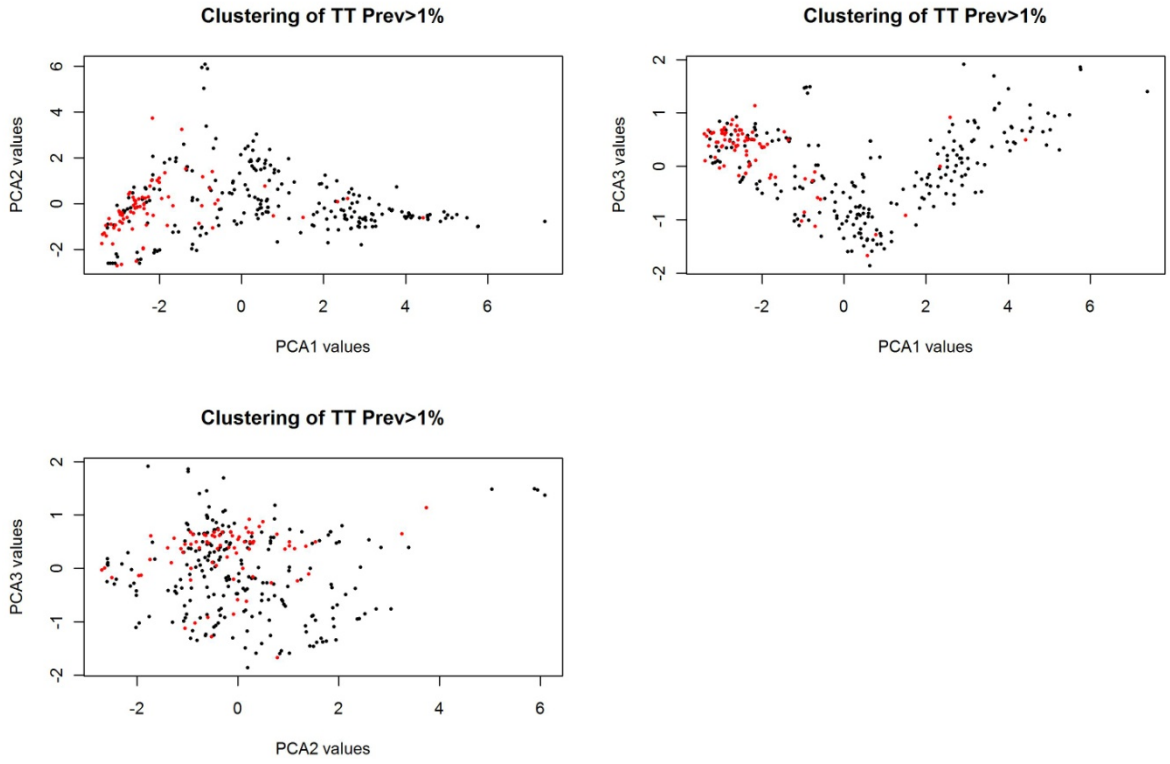



Figure A4.2.5 Calculated principal component (PC) values for each cluster, based on loadings for each component. Clusters with a prevalence of trichiasis (TT) >1% are shown in red, highlighting the ability of the three components to distinguish areas of higher prevalence.

# Appendix 4.3: Introduction to semi-variogram analysis

A widely-used approach in the description of spatial structures is semi-variogram analysis. This function characterises the spatial autocorrelation structure of a variable by defining semi-variance (a measure of expected dissimilarity between a given pair of observations) as a function of lag (the distance separating the observation locations). This method is used not only as an exploratory technique in spatial analysis, but to formally inform parameter estimates for incorporation into geostatistical models.

## Estimation of the semi-variogram

The empirical (or sample) semivariogram is estimated directly from the survey data by measuring the mean squared difference of pairs of observations that are separated by a lag [1,2]. While this can be plotted as a semi-variogram cloud (i.e. a scatterplot of the distance between and variogram ordinate) for every point, more commonly values are averaged that fall within a declared bin width in order to provide a smoothed visualisation of the underlying covariance structure.

$$\gamma(h) = \frac{1}{2\,W}\sum_{W}\left[\left(Z(s_i) - Z(s_j)\right)^2\right]$$

Where $Z(s_i)$ and $Z(s_j)$ are data values at locations $s_i$ and $s_j$, for all $W$ pairs of points separated by the Euclidean distance $h$. The smaller the variance in the difference, the greater the correlation between measurements taken a given distance apart.

A theoretical semi-variogram can then be fitted to the sample semi-variogram using ordinary or weighted-least squares in order to formally describe the structure and incorporate spatial dependency into models.

## Interpretation

The shape of the semivariogram can provide information about the spatial autocorrelation structure and the distance over which such correlation is present, and be used to define the covariance structure between points in regression models. In the presence of spatial autocorrelation (Figure A4.3.1), the semivariogram rises with distance over a *range,* whereupon it plateaus to a maximum value termed the *sill*. The range represents the maximum distance over which there is autocorrelation, while the *nugget* represents the variation that is aspatial or measurement error [3]. A flat semivariogram indicates the absence of spatial structure, while a constant rise in semivariance with no plateau suggests the presence of a large scale trend (ie spatial autocorrelation at distances larger than the study region).

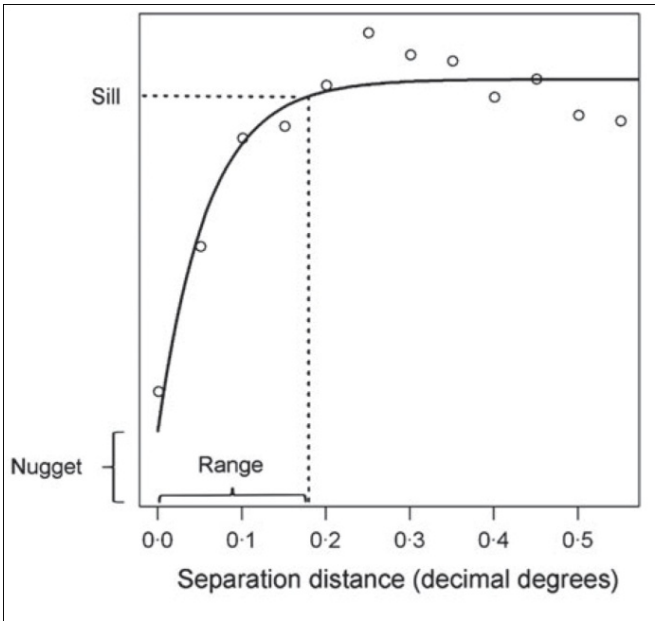



Figure A4.3.1 Semi-variogram plot showing key structural elements

As discussed in the introduction, a spatially varying mean can be easily incorporated into the analysis by regression. The residuals from this regression can be extracted, and used to explore second order spatial autocorrelation.

## References

1. Diggle PJ, Ribeiro PJ (2007) Model-based Geostatistics. New York: Springer.
2. Dale M, Dixon P, Fortin MJ, Legendre P, Myers DE, et al. (2002) Conceptual and mathematical relationships among methods for spatial analysis. Ecography 25: 558-577.
3. Cressie N (1991) Statistics for spatial data. New York: Wiley.

# Appendix A5.1: Description of socioeconomic, environmental and climatic variables

| Variable (level) | Description and source |
|---|---|
| **Socioeconomic** | |
| District level variables | |
| Access to improved water source | District level, year-specific predictions of the proportion of households with access to an improved water source<br>Resolution: Area (district)<br>Source: [1] |
| Access to water source w/in 1 km | District level, year-specific predictions of the proportion of households with access to water within 1km<br>Resolution: Area (district)<br>Source: [1] |
| Improved sanitation | District level, year-specific predictions of the proportion of households with access to improved sanitation<br>Resolution: Area (district)<br>Source: [1] |
| Open defecation | District level, year-specific predictions of the proportion of households reporting open defecation<br>Resolution: Area (district)<br>Source: [1] |
| Crowded living quarters | District level, year-specific predictions of the proportion of households with more than 5 individuals per room<br>Resolution: Area (district)<br>Source: [1] |
| Average school attendance | Net attendance ratio reported for 2005/2006.<br>Resolution: Area (district)<br>Source: Kenya Open Data (from Kenya Integrated Household Budget Survey) |
| Constituency level variables | |
| Poverty Incidence | Measures of poverty modelled using small area estimation techniques with data from the 1997 Welfare Monitoring Survey and the 1999 Population and Housing Census<br>Resolution: Area (constituency)<br>Source: Central Bureau of Statistics and The World Bank [2] |
| Gini Index | Measures of poverty modelled using small area estimation techniques with data from the 1997 Welfare Monitoring Survey and the 1999 Population and Housing Census<br>Resolution: Area (constituency)<br>Source: Central Bureau of Statistics and The World Bank [2] |

Table continued

Table A5.1 continued

| Cluster level variables | |
|---|---|
| Distance to nearest primary school (km) | Derived: Gridded surface calculated from 2007 Kenya primary schools database<br>Resolution: 30 arc second (~1 km)<br>Source: Generated for this analysis from schools database at Kenya OpenData |
| Schools per 1000 population | Derived: Number of primary schools per 1000 population within each sublocation<br>Resolution: Sublocations (small area)<br>Source: Generated for this analysis from 2007 Kenya primary schools database at Kenya OpenData and Afripop 2010 |
| Cattle density (animals per 5km cell) | Gridded density of cattle, based on the spatial disaggregation of sub-national statistical data based on empirical relationships with environmental variables in similar ag-ecological zones.<br>Resolution: 2.5 arc-minute (~5 km)<br>Source: Food and Agriculture Organization's Gridded Livestock of the World [3] |
| Distance to small scale irrigation | Derived: Distance small scale irrigation sites<br>Resolution: 30 arc second (~1 km)<br>Source: Generated for this analysis from database of small irrigation points at World Resources Institute [4] http://www.wri.org/resources/data-sets/kenya-gis-data |
| **Environmental** | |
| Urban classification | Gridded database of urban settlements with populations greater than 1000 persons.<br>Derived: Satellite night-light data and gridded population data<br>Resolution: 30 arc second (~1 km)<br>Source: Global Rural-Urban Mapping Project (GRUMP) |
| Land cover | Land Cover Type 2 (UMD) Annual Averages Data Set for Africa averaged each year from 2001 to 2009.<br>Derived: Satellite remote sensing<br>Resolution: 500 m<br>Source: Africa Soil Information Service (AfSIS) using the Moderate Resolution Imaging Spectroradiometer (MODIS) imagery from the National Aeronautics and Space Administration (NASA). |
| Distance to river or water body | Derived: Distance to nearest river or surface water body<br>Resolution: 2.5 arc-minute (~5 km)<br>Source: FAO Rivers and Surface Water Bodies database |

Table continued

Table A5.1 continued

| | |
|---|---|
| Distance to road (km) | Accessibility measure calculated as the distance to roads.<br>Derived: Digital Chart of the World<br>Resolution: Vector<br>Source: DIVA |
| Distance to primary road (km) | Accessibility measure calculated as the distance to major (primary) roads.<br>Derived: Digital Chart of the World<br>Resolution: Vector<br>Source: DIVA |
| Population density | Gridded population data for the year 2010<br>Derived: NA<br>Resolution: 30 arc second (~1 km)<br>Source: SEDAC's Gridded Population of the World, Version 3 data set (GPWv3) |
| Altitude (meters) | Elevation data<br>Derived: Radar<br>Resolution: 2.5 arc-minute (~5 km)<br>Source: Shuttle Radar Topography Mission |
| Enhanced vegetation index (EVI) | Index of the vegetation signal from surface reflectance<br>Derived: Satellite remote sensing<br>Resolution: 2.5 arc-minute (~5 km)<br>Source: Moderate Resolution Imaging Spectroradiometer NASA's Terra satellite |
| **Climate** | |
| Land surface temperature (LST) (°C) | Mean and variance of annual land surface temperature for the years 2005-2007.<br>Derived: Satellite remote sensing<br>Resolution: 2.5 arc-minute (~5 km)<br>Source: Moderate Resolution Imaging Spectroradiometer (MODIS) on NASA's Terra satellite |
| Mean annual temperature (°C) | Interpolation of average monthly mean annual temperature data (~1950-2000)<br>Derived: Meteorological stations<br>Resolution: 2.5 arc-minute (~5 km)<br>Source: WorldClim BioClim variables |
| Mean annual precipitation (mm) | Interpolation of average monthly mean annual precipitation data (~1950-2000)<br>Derived: Meteorological stations<br>Resolution: 2.5 arc-minute (~5 km)<br>Source: WorldClim BioClim variables |
| Annual aridity index | Interpolation of mean Annual Precipitations/Mean Annual Potential Evapo-Transpiration (~1950-2000).<br>Derived: Meteorological stations (WorldClim bioclimatic variables)<br>Resolution: 30 arc second (~1 km)<br>Source: Consortium for Spatial Information (CGIAR-CSI) Global-Aridity and Global-PET Database [5,6] |

m: meters; mm: millimetres; C: Celsius; km: kilometers

[a]Classified as binary variable indicating savannah/grasslands

## References

1. Pullan RL, Freeman MC, Gething PW, S B (under review) Mapping inequalities in access to improved drinking water supply and sanitation across sub-Saharan Africa.
2. Central Bureau of Statistics (2003) Geographic dimensions of well-being in Kenya. Where are the poor? From districts to locations. . Nairobi, Kenya: Ministry of Planning and National Development, in collaboration with International Livestock Research Institute.
3. Robinson T, Fanceschini G, William W The Food and Agriculture Organization's Gridded Livestock of the World. Veterinaria Italiana 43: 745-751
4. Ministry of Agriculture and Livestock Development (MoALD) (1995) District Profiles of Irrigation and Drainage in Kenya (All Districts). Nairobi: Ministry of Agriculture and Livestock Development (MoALD), Irrigation and Drainage Branch.
5. Zomer R, Trabucco A, Bossio D, van Straaten O, Verchot L (2008) Climate Change Mitigation: A Spatial Analysis of Global Land Suitability for Clean Development Mechanism Afforestation and Reforestation. Agriculture, Ecosystems & Environment 126: 67-80.
6. Zomer R, Bossio D, Trabucco A, Yuanjie L, Gupta D, et al. (2007) Trees and Water: Smallholder Agroforestry on Irrigated Lands in Northern India. Colombo, Sri Lanka: International Water Management Institute. 45 p.

# Appendix 5.2: Summary statistics for each covariate

| | Surveyed sites | |
|---|---|---|
| **Variable** | **Median (range)** | **SD** |
| **Socioeconomic** | | |
| Constituency | | |
| Poverty incidence | 53.5% (33, 76) | 11.8 |
| Gini Index | 33% (29, 42) | 2.7 |
| District: | | |
| Access to improved water source | 63.1% (51.4, 88.9) | 9.2 |
| Access to water source w/in 1km | 44.0% (29.9, 66.4) | 11.4 |
| Improved sanitation | 46.5% (40.5, 81.3) | 10.6 |
| Open defecation | 31.9% (16.6, 46.9) | 9.8 |
| Crowded living quarters (>5 per room) | 10.1% (3.6, 19.5) | 4.8 |
| Net attendance | 74.3% (44.1, 91.9) | 15.9 |
| Proportion SAC never attended | 13.8% (0.0, 47.9) | 14.0 |
| Cluster/Sublocation | | |
| Distance to nearest school (km) | 2.2 (0.0, 45.0) | 7.1 |
| Schools per 1000 population | 1.0 (0.0, 5.3) | 1.0 |
| Cattle density (animals per 5km cell) | 18.0 (0, 650) | 71.8 |
| Distance to small scale irrigation | 18.7 (0.0, 305.1) | 63.8 |
| **Environmental** | | |
| Distance to water bodies (km) | 32.6 (0, 169.5) | 28.6 |
| Urban classification | 8.3% urban | - |
| Landcover : Savannah/grasslands | 63.3% | - |
| Barren/sparsely vegetated | 3.6% | - |
| Population density (persons per 1km cell) | 34.5 (0, 1463) | 225.7 |
| Distance to road (km) | 3.2 (0, 39.1) | 6.2 |
| Distance to major road (km) | 77.6 (0, 407.8) | 98.0 |
| **Climate** | | |
| Altitude (m) | 1052 (214, 2816) | 535.5 |
| Enhanced vegetation index (EVI) | 0.3 (-0.0, 0.5) | 0.1 |
| Variance land surface temperature (LST) (°C) | 13.0 (1.0, 39.0) | 6.1 |
| Average land surface temperature (LST) (°C) | 36.4 (19.0, 52.6) | 6.4 |
| Mean annual temperature (°C) | 22.0 (12.0, 29.0) | 3.5 |
| Mean annual precipitation (mm) | 512.9 (86.9, 1831.7) | 351.8 |
| Annual aridity index | 0.4 (0.1, 1.4) | 0.2 |

m: meters; mm: millimetres; km: kilometres; C: Celsius; SD: standard deviation; SAC: school-age children

# Appendix 6: Technical notes on data simulation

## Parameterisation & expansion of dataset

Simulated data were parameterised using full census data from Kahe. The relative risk of TF between children aged 1-5 and 6-9 years was estimated ($RR_{age}$ = 2.0) from the full dataset. Using data from individuals for which enrolment data were available (n=421) the relative risk of TF was estimated between enrolled and non-enrolled children aged 6-9 years ($RR_{enrol}$ = 0.5). In addition, an analysis of variance estimator was used to estimate the household intraclass correlation (ICC) present in Kahe, which was equal to 0.26 [1]. Initially, the enrolment rate ($R_{enroll}$) was set as 0.7 and assumed to not cluster within households. This assumption was based on results from a logistic regression model looking at the effects of these indicators on risk of TF in children aged 6-9 years. This model showed that individual level school attendance was the main factor in this context (odds ratio=0.52, p= 0.05), and having a school-going sibling was not associated with any additional risk (p=0.71). Similarly, in children 1-5 years old, having a school-going sibling was not associated with risk of TF after adjusting for household clustering (p=0.13).

Between and within district variation was estimated by fitting a beta distribution to the district-level and cluster-level TF prevalence data, respectively. The average variance and simulated mean prevalence were then used to define the sampling distributions for the two levels, with the mean of prevalences set as 0.2 at the district-level and equal to the simulated district prevalence for the cluster sampling distribution, from which cluster prevalence values were randomly drawn.

The dataset was subsetted, retaining data for the relevant age group (1-9 years). This dataset was then replicated to generate identical household and demographic structures for the specified number of communities ($n$) within each of $k$ districts.

## Simulation of cluster data

We outline a method in which individual disease status was then simulated for an *n* x *k* dataset using a methodology based on previous work for generating correlated binary random vectors outlined in Olives et al [2]. This approach was modified to use household as a grouping factor and simulated data for each subgroup separately, so that each subgroup maintained household-level clustering, the specified $RR_{age}$ and $RR_{enrol}$ and summed to the overall cluster-level prevalence.

For each cluster within each district, the enrolment was simulated so that a specified proportion of children set by $R_{enroll}$ were randomly selected and assigned as enrolled. TF cases were simulated by the two groups (age and enrolment) with the goal of ensuring the $RR_{age}$ of TF in unenrolled and $RR_{enroll}$ in 6-9s was preserved. To do so, we assumed that prevalence in enrolled 1-5s was zero (since we assume there are no 1-5s in school) and solved the following system of equations for the marginal prevalence of TF by age and enrolment:

1. $P_{cluster} = P(TF|Age = 1-5, Enroll = 0) \times P(Age = 1-5, Enroll = 0) +$
   $P(TF|Age = 6-9, Enroll = 0) \times P(Age = 6-9, Enroll = 0) +$
   $P(TF|Age = 6-9, Enroll = 1) \times P(Age = 6-9, Enroll = 1)$
2. $RR_{Age\ in\ nonenroll} =$
   $Pr(TF|Age = 1-5, Enroll = 0)/Pr(TF|Age = 6-9, Enroll = 0)$
3. $RR_{Enrolled\ in\ 6-9} =$
   $Pr(TF|Age = 6-9, Enroll = 1)/Pr(TF|Age = 6-9, Enroll = 0)$

Once the marginal TF prevalences by age and enrolment were derived, the household ICC was introduced by using the methodology outlined in Olives et al [1].

## References

1. Ridout MS, Demetrio CG, Firth D (1999) Estimating intraclass correlation for binary data. Biometrics 55: 137-148.
2. Olives C, Pagano M, Deitchler M, Hedt BL, Egge K, et al. (2009) Cluster designs to assess the prevalence of acute malnutrition by lot quality assurance sampling: a validation study by computer simulation. J R Stat Soc Ser A Stat Soc 172: 495-510.