Forecasted trends in vaccination coverage and correlations with socioeconomic factors: a global time-series analysis over 30 years

Alexandre de Figueiredo, Iain G Johnston, David M D Smith, Sumeet Agarwal, Heidi J Larson*, Nick S Jones*

Summary

Background Incomplete immunisation coverage causes preventable illness and death in both developing and developed countries. Identification of factors that might modulate coverage could inform effective immunisation programmes and policies. We constructed a performance indicator that could quantitatively approximate measures of the susceptibility of immunisation programmes to coverage losses, with an aim to identify correlations between trends in vaccine coverage and socioeconomic factors.

Methods We undertook a data-driven time-series analysis to examine trends in coverage of diphtheria, tetanus, and pertussis (DTP) vaccination across 190 countries over the past 30 years. We grouped countries into six world regions according to WHO classifications. We used Gaussian process regression to forecast future coverage rates and provide a vaccine performance index: a summary measure of the strength of immunisation coverage in a country.

Findings Overall vaccine coverage increased in all six world regions between 1980 and 2010, with variation in volatility and trends. Our vaccine performance index identified that 53 countries had more than a 50% chance of missing the Global Vaccine Action Plan (GVAP) target of 90% worldwide coverage with three doses of DTP (DTP3) by 2015. These countries were mostly in sub-Saharan Africa and south Asia, but Austria and Ukraine also featured. Factors associated with DTP3 immunisation coverage varied by world region: personal income (Spearman's $\rho=0\cdot66$, $p=0\cdot0011$) and government health spending ($0\cdot66$, $p<0\cdot0001$) were informative of immunisation coverage in the Eastern Mediterranean between 1980 and 2010, whereas primary school completion was informative of coverage in Africa ($0\cdot56$, $p<0\cdot0001$) over the same period. The proportion of births attended by skilled health staff correlated significantly with immunisation coverage across many world regions.

Interpretation Our vaccine performance index highlighted countries at risk of failing to achieve the GVAP target of 90% coverage by 2015, and could aid policy makers' assessments of the strength and resilience of immunisation programmes. Weakening correlations with socioeconomic factors show a need to tackle vaccine confidence, whereas strengthening correlations point to clear factors to address.

Funding UK Engineering and Physical Sciences Research Council.

Department of Mathematics, Imperial College London, London, UK
(A de Figueiredo MSc, I G Johnston DPhil, N S Jones PhD); **School of Biosciences, University of Birmingham, Birmingham, UK** (I G Johnston); **London School of Hygiene & Tropical Medicine, London, UK** (D M D Smith DPhil, H J Larson PhD); **Department of Global Health, University of Washington, Seattle, WA, USA** (H J Larson); **and Department of Electrical Engineering, Indian Institute of Technology, New Delhi, India** (S Agarwal DPhil)

Correspondence to:
Dr Nick S Jones, Department of Mathematics, Imperial College London, London SW7 2AZ, UK
nick.jones@imperial.ac.uk
or
Dr Heidi J Larson, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK
heidi.larson@lshtm.ac.uk

Introduction

Rates of vaccine-preventable diseases have decreased in many parts of the world in the past few decades, but many children remain unvaccinated. In 2013, UNICEF reported that $21\cdot8$ million children younger than 1 year had not completed the diphtheria, tetanus, and pertussis (DTP) immunisation series, with a similar number not receiving a single measles vaccination.[1] Although access to vaccinations is the main barrier in many settings, growing numbers of parents do not immunise their children based on their own personal attitudes.[2,3]

Both socioeconomic and attitudinal barriers to vaccination coverage have been identified in several countries. A set of recurring socioeconomic and demographic correlates of coverage have been reported, with parental education level,[4–8] age,[6,9,10] employment status and workplace,[4,6,7] religion,[11] ethnic origin,[12] gender of the child,[13] poverty,[14–18] and distance to health-care facilities[4,6,11,14,17,18] all linked to vaccine uptake, although marked differences often exist between countries.[11] Repeating themes have likewise been identified in examination of personal reasons for vaccine acceptance or delay, which range from perceived risks about potential adverse events, to religious or political beliefs that are external to, albeit influential in, vaccination.[9,19–27] Trust in health-care professionals[19,20,23,24] and the government[19,23,26] also features highly.

Recommendations to address gaps in vaccine coverage are context dependent. Targeting of mothers with low education,[28] low health literacy,[29] and dissemination of more information to communicate vaccine benefits and risks transparently[2] have been suggested to aid in improvement of vaccine hesitancy, as has the customisation of messages and engagement efforts for specific groups.[30] Efficient identification of clusters of non-vaccinators through monitoring of local immunisation rates has been identified as a key public

Research in context

Evidence before this study
We searched Google Scholar between Jan 1, 1980, and Jan 1, 2016, for studies assessing socioeconomic correlates of vaccine coverage. Our search terms included the words "socio-economic", "demographic", "correlates", and "vaccine coverage", with variations in each individual term allowed. For example, "vaccine" could be with "vaccination" or "immunise" with "immunisation", in addition to specific names of vaccine-preventable diseases such as "DTP" and "MMR"; alternative spellings, such as "immunize", were also allowed. For attitudinal correlates of vaccine coverage, we used search terms such as "vaccine hesitancy", "vaccine acceptance", and "vaccine delay". We identified few publications that examined differences in both attitudinal and socioeconomic correlates between countries, although we found some recent reviews that assessed differing attitudes towards vaccination uptake across Europe, and differing socioeconomic correlates in east African countries. Both attitudinal and socioeconomic surveys reported various correlates of vaccine uptake. Attitudes range from personal, religious, and political beliefs, to trust in health care and the government. Socioeconomic correlates vary between countries, but education level, religion, ethnic origin, and distance to health-care facilities are all recurring themes.

Added value of this study
We studied correlations between 190 socioeconomic factors and immunisation coverage in 190 countries, and report variations in the strength of socioeconomic correlates between world regions and across time. To our knowledge, this work represents the broadest analysis of socioeconomic links to vaccine coverage yet available, allowing insights into how socioeconomic correlates modulate coverage and how these correlates vary by world region. Forecasting of vaccine coverage time series enabled us to summarise the recent trend and variability in uptake rates and construct a vaccine performance index, which is, as far as we know, the first quantitatively derived marker of vaccine performance. Our predictive performance index represents a distinctive, interpretable measure to assess likely future vaccine coverage behaviour and resilience of immunisation programmes to volatile changes triggered by external shocks, whether driven by political rumour or natural disaster.

Implications of all the available evidence
By use of a probabilistic forecast of future vaccination levels, we have provided a world map of a vaccine performance index that is informative of both the likelihood of stagnated or substandard coverage levels and the chance coverage levels will decline. These forecasts can easily identify countries that are likely to fail to achieve the Global Vaccine Action Plan target of 90% coverage of diphtheria, tetanus, and pertussis vaccine by 2015. We speculate that regions with strong socioeconomic ties to vaccination coverage can be interpreted as targets for intervention (because socioeconomic factors present barriers to vaccination), but that some regions (namely, Europe and the Americas), with low ties to socioeconomic factors, have increased attitudinal barriers to vaccination.

health challenge for the prevention of outbreaks when local groups adopt a non-vaccination status (as was the case in the 2014–15 measles outbreak at Disneyland, Anaheim, CA, USA).[31] Monitoring of trust in immunisation programmes through various indices has also been proposed as a way to better understand the circumstances in which vaccine hesitancy might arise.[32,33]

To build on the current literature, we constructed a vaccine performance index: a summary measure of the strength of immunisation coverage in a country. This index could be used to describe the variability in vaccine coverage induced by confidence and access issues, allowing for an intercountry analysis of trends. We interpreted the coverage values in view of the Global Vaccine Action Plan (GVAP)'s target of attaining 90% vaccination coverage in all countries by 2015.[34] We discuss the implications of both the correlative study and the performance index on immunisation strategies. We undertook a large-scale, worldwide correlative analysis to investigate the association between a range of socioeconomic factors and vaccine coverage, and to identify temporal trends and differences between WHO world regions.

Methods

Data
We examined a broad range of quantitative data for 190 descriptive socioeconomic factors for 190 countries from 1980 to 2010. Factors spanned economics, health care, industry, demographics, communications, infrastructure, physical geography, trade, and education. This breadth of socioeconomic factors was used to expose, in an unbiased way, which indicators might be related to immunisation coverage. Countries were grouped into six world regions according to WHO classifications: Africa, the Americas, the Eastern Mediterranean, Europe, South-East Asia, and the Western Pacific.[35] We obtained data from the Gapminder website, which draws from sources including the World Bank, the International Labour Organization, and WHO. A data curation approach was used to filter and impute missing data (appendix).

Data for vaccine coverage (the proportion of a target group immunised) were obtained for BCG tuberculosis vaccine, DTP1 and DTP3 (representing the first and third doses, respectively, and typically scheduled for 6 months after birth), measles-containing vaccine (MCV), and polio (POL3, representing the third dose of the polio vaccination

programme) from the 2011 WHO–UNICEF estimates of vaccine coverage[36] for the same countries and across the same time period, although we used the latest estimates in our vaccine performance index. Immunisation rates have increased across all WHO regions in the past three decades: coverage is highest in Europe and the Americas, with other regions achieving less complete coverage (figure 1A). Joint investigation of all vaccine time series showed that strong mutual correlations exist between the DTP1, DTP3, MCV, and POL3 vaccines, but that BCG falls outside this grouping (figure 1B). Time series of DTP3 coverage in individual countries differ substantially in magnitude, variability, and trends (figure 1C). We henceforth focused on DTP3 coverage because of this high correlation and in accordance with previous literature in which DTP3 is regarded a marker of the strength of a country's immunisation programmes (because the vaccine requires three different administrations).[37] A correlative analysis of the other vaccines is presented in the appendix.

We illustrate time-series data for socioeconomic factors with two examples: female education (figure 1D) and rural access to water (figure 1E). Figure 1F shows an example scatter plot of average DTP3 coverage against female education levels for three regions. We further explored the dataset by use of t-distributed stochastic neighbour embedding (t-SNE), which is a local structure-preserving clustering technique that can improve visualisation quality of high-dimensional data over other clustering algorithms.[38] We used t-SNE for visualisation of a similarity matrix, constructed by taking correlations across all countries between all pairs of socioeconomic factors and vaccines in a single year (x_{it} [for socioeconomic factor or vaccine i in year t]) and averaging the correlation across all 31 years:

$$S_{ij} = 1 - \frac{1}{T}\left|\sum_t \rho(x_{it}, x_{jt})\right|$$

where S_{ij} are elements of the similarity matrix $\mathbf{S}$, T is the total number of years, and ρ is the Spearman's rank correlation between two socioeconomic factors (or vaccines). Application of t-SNE showed strong links between socioeconomic factors in similar categories and revealed a separation of the BCG vaccine, which is closer in t-SNE space to factors relating to health spending, from the other four vaccines, which seem to be more correlated with schooling variables (figure 1G).

Correlative analysis

We used Spearman's rank correlation to explore links between coverage and socioeconomic factors (figure 1F) because coverage data are confined to the 0–100 interval and often non-linear relations arise between coverage and socioeconomic factors (which might be ordinal). To investigate the strengths of correlations with coverage over time, we computed a time-averaged correlation between each factor (x_t) and coverage (y_t) in a given year

(t) for all countries in a particular region $\rho(x_t, y_t)$, and we dropped the index i referring to socioeconomic indicator; we then took the time-average of this value:

$$\bar{\rho} = \frac{1}{t_2 - t_1 + 1}\sum_{t=t_1}^{t_2}\rho(x_t, y_t)$$

where t_1 is the most distant year and t_2 is the most recent year. For historic trends, we considered 1980 as t_1 and 2010 as t_2 and, for recent trends, we considered 2001 as t_1. To focus on the strongest signals, we agglomerated the top-ten correlating socioeconomic factors in each region across both time ranges for a cross comparison of informative correlates between regions; we restricted these socioeconomic factors to those with a Bonferroni-corrected meta p value (against the null hypothesis that the set of correlations were uninformative) of less than 0·01. In addition to these socioeconomic factors, we included other factors that are related and useful for comparison: primary school completion rate (male), improved sanitation access (overall and rural), and improved urban water access (appendix).

Forecasting

We first logit-transformed vaccine coverage data:

$$y' = -\log\left(\frac{100}{y} - 1\right)$$

where y' is transformed vaccine coverage and y is raw coverage, to map the coverage values to the real line, so that Gaussian distributions over datapoints were not truncated at values of 0 or 100. We used Gaussian processes, with a squared exponential covariance function and linear mean function (to account for the broadly linear increases in trend in this transformed space), to forecast vaccine coverage data, based on previous coverage values. The appendix summarises the prediction accuracy of this method, which is quantified by forecasting on a test dataset. We used the predictive distribution over forecasted values to define a summary measure—the vaccine performance index—which balances the (desirable) probability of attainment of high future coverage against the (undesirable) probability of experiencing a future negative perturbation to coverage, informed by preceding trends and variability. We defined the vaccine performance index as the probability of achieving of coverage in excess of a threshold value (v), less the probability of a drop in coverage of size at least d:

$$VPI_{v\tau,d}(t_*) = P[y(t_*) > v] - P[y(t_*) < y(t_N) - d]$$

where VPI is the value of the vaccine performance index at time t_*, t_N is the most recent point for which coverage exists, and $t_* = t_N + \tau$ is the forecasted point (so that, for example, $t_N = 1999$ and $t_* = 2000$ for a country's 2000 vaccine
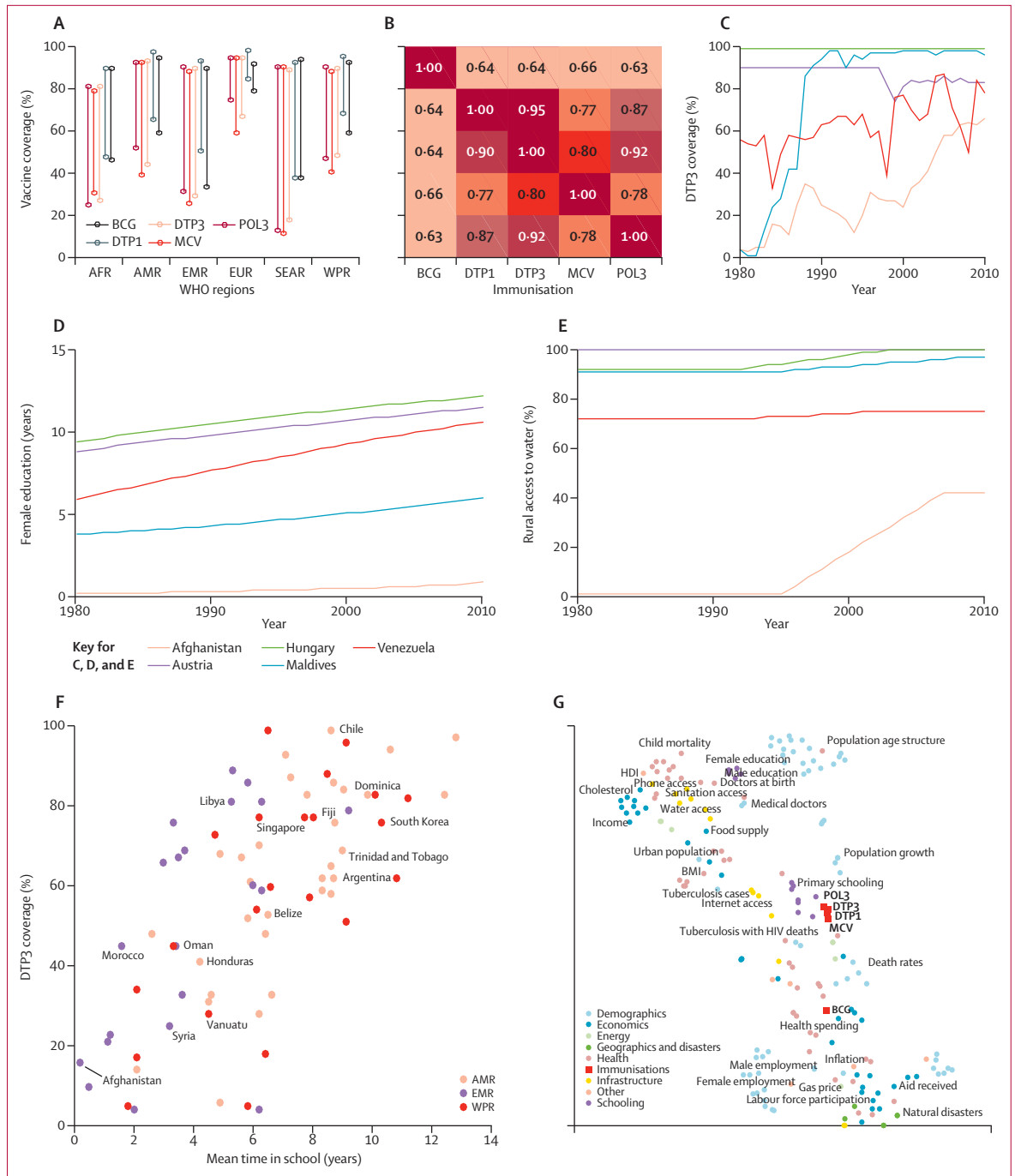
***Figure 1*: Dynamics of socioeconomic factors and DTP3 coverage**
(A) Increasing trends in vaccine coverage across WHO regions with time (1980–2010). Leftmost circles represent 1980 coverage rates and rightmost circles represent 2010 rates, averaged across all countries in a region. (B) Spearman correlation between all vaccines across all countries and years. (C) Diversity of behaviours in DTP3 coverage with time. (D) Time series of female education. (E) Time series of rural access to water. (F) Scatter plot showing the link between DTP3 coverage and female education in 1982. (G) Visualisation of similarity matrix between all pairs of socioeconomic factors by use of *t*-distributed stochastic neighbour embedding (*t*-SNE). Heuristic categories were assigned before *t*-SNE was applied. DTP1=diphtheria, tetanus, and pertussis first dose. DTP3=diphtheria, tetanus, and pertussis third dose. MCV=measles-containing vaccine. POL3=polio vaccine third dose. AFR=Africa. AMR=Americas. EMR=Eastern Mediterranean. EUR=Europe. SEAR=South-East Asia Region. WPR=Western Pacific region. HDI=Human Development Index.
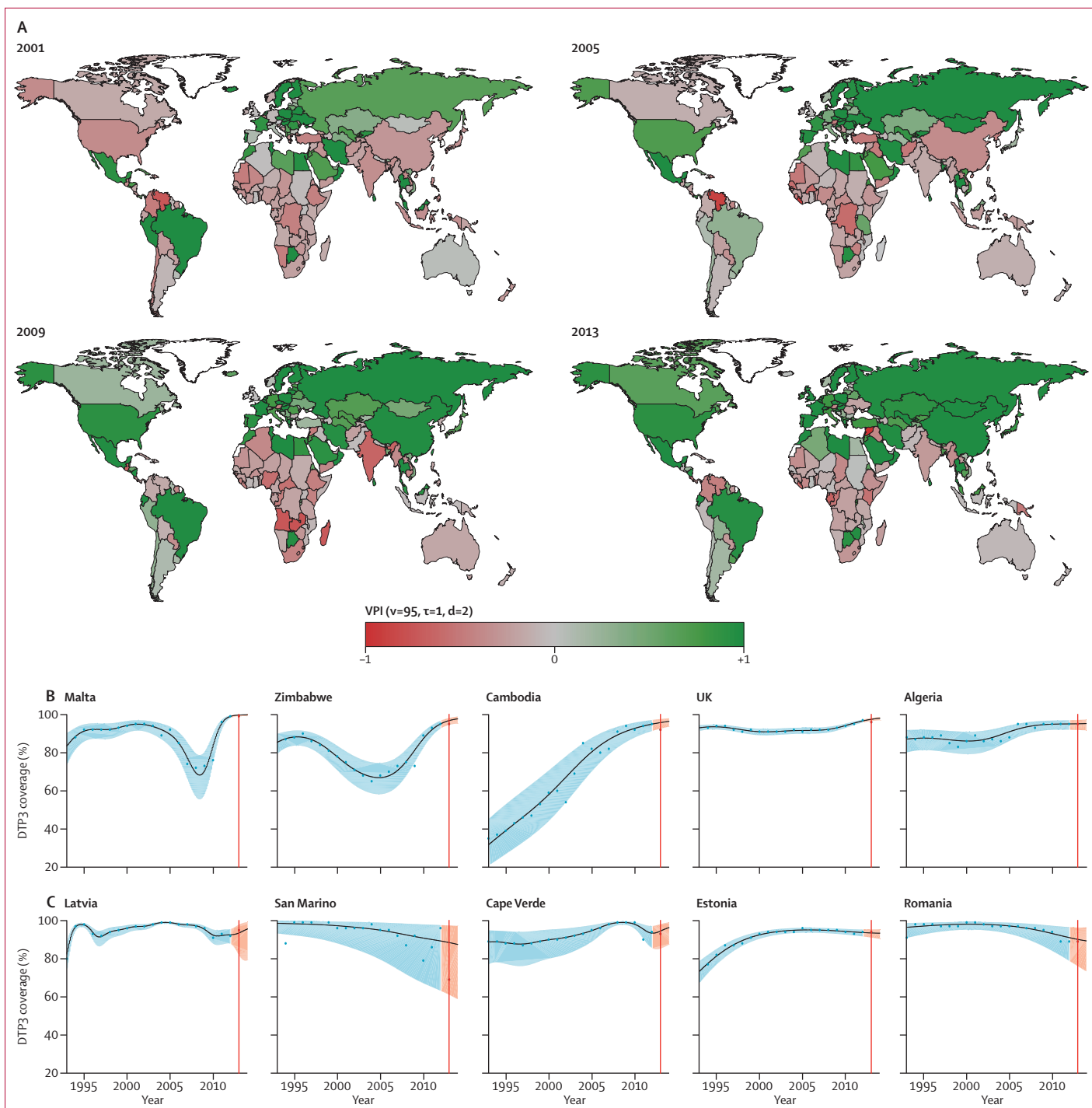
***Figure 2*: VPI values**
(A) Sub-Saharan Africa, parts of southern and South-East Asia, and some European countries have low VPI. White areas signify missing data. (B) Countries with the largest improvement in VPI between 2009 and 2013 (Malta is most improved). (C) Countries with the largest decrease in VPI between 2009 and 2013 (Romania has the most deteriorated performance). WHO–UNICEF coverage estimates are in blue with the observed (test) value in 2013 in red; the mean of the Gaussian process in shown in black and shaded areas represent 95% CIs for training (blue) and forecasted (red) data. The red vertical line indicates 2013. DTP3=diphtheria, tetanus, and pertussis third dose. VPI=vaccine performance index.

performance index value) and for which we mapped the predictive distributions back to the range 0–100. This expression is reweightable if the reader or policy maker wishes to focus on a different combination of strategic properties. There are natural alternative forms to this index; for example, to assess the GVAP goal of 90% coverage in all countries by 2015, we set τ equal to 2 years, v equal to 90, and d equal to infinity, thus using our model to forecast the probability that coverage in 2015 will exceed 90%. We used MATLAB (version 2015a) for all analyses.

Role of the funding source
The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results
We assessed countries' trends and variability by considering the probability of coverage exceeding 95% minus the probability of coverage falling by more than 2% 1 year in the future (ie, v=95, τ=1, and d=2); the appendix details the justification for these parameter values and alternative forms of this index, which might be natural alternatives that address different policy-making goals. Figure 2A shows worldwide vaccine performance index values for these parameters in 2001, 2005, 2009, and 2013. DTP3 coverage time series and Gaussian process fits for the five countries with the highest increases and decreases in vaccine performance index values from 2009 to 2013 are displayed in figure 2B and 2C, respectively. Malta had the most improved vaccine performance index between 2009 and 2013, since, in 2009, it had sharply decreasing coverage rates, which have since fully recovered (figure 2B). Romania had the most deteriorated coverage values, since high and steady values have gradually worsened (figure 2C). The appendix presents vaccine performance index values for all countries.

Consistently high-performing countries included Saudi Arabia, Russia, Hungary, and Slovakia, which have had consistently high DTP3 coverage rates since at least 2000 (appendix). The UK, USA, Canada, Australia, and New Zealand had historically lower vaccine performance index values, which have improved in the UK, USA, and Canada, but remain low in Australia and New Zealand (figure 2, appendix).

More generally, Europe, North and Central America, and much of Asia have shown steady progress since 2005; by contrast, South America, the Indian subcontinent, and sub-Saharan Africa have shown little improvement (figure 2, appendix). Surprisingly, many European countries have struggled to increase coverage to top levels: Norway, Iceland, Ukraine, and Romania, despite reasonably high DTP3 coverage rates in 2013 (83%, 91%, 94%, and 89%, respectively) all have vaccine

performance index values far from +1 (figure 2, appendix), signifying coverage rates that are below 95%, and show little sign of improvement beyond this threshold. Findings for Africa are mixed: north African countries have high vaccine performance index values, whereas countries in sub-Saharan Africa generally have low values due to high variability and low coverage, with some exceptions, notably Rwanda, Burundi, Botswana, and Zimbabwe (figure 2, appendix). These low and volatile coverage rates are likely to be a sign of the vast number of people in sub-Saharan Africa with poor continued access to latter doses in the DTP programme.

South-East Asia generally performs poorly; many countries have vaccine performance values close to –1 (figure 2, appendix), suggesting that chances of a decline in coverage in the following year are high. Many Eastern Mediterranean countries have shown clear improvement in the past decade, with countries typically presenting high, sustained coverage rates, although Syria, Iraq, Pakistan, Afghanistan, and Yemen are notable exceptions (figure 2, appendix).

To assess the GVAP goal of 90% coverage in all countries by 2015, we used our vaccine performance index to forecast the probability that coverage in 2015 (2 years after the most recent available datapoints, in 2013) will exceed 90%. This analysis yielded a bimodal distribution (figure 3), whereby countries will either comfortably reach this goal or fall rather short of it (with the exception of South America, which has many borderline countries). European countries generally perform better than other regions (with the notable exceptions of Denmark, Iceland, Romania, Austria, Moldova, San Marino, and Ukraine), with an almost certain chance of reaching the GVAP goal (on the basis of their recent trend; figure 3). Countries in sub-Saharan Africa and the Indian subcontinent (with only few exceptions) seem likely to drastically miss the GVAP target (figure 3). This inference and forecasting approach highlights not only regions that require stronger action, but also borderline countries where comparatively minor action might realise the GVAP goal.

From the broad range of 190 socioeconomic factors investigated, those shown in figure 4A can be interpreted as the most likely to have a genuine (and consistent) link with DTP3 coverage over time. The strength of correlation between socioeconomic factors and DTP3 coverage varies between region, over time, and across socioeconomic factors (figure 4A). Births attended by skilled health staff and access to sanitation are the two factors that have mean-averaged p values of less than 0·05 across the largest number of regions (figure 4A). In Africa, the Americas, and Europe, the magnitude of correlations are high historically, but have decreased recently, and display patterns that are broadly similar: the importance of access to water and sanitation, births attended by skilled health staff, and economic factors in determining coverage is historically high, but diminishing (particularly in the
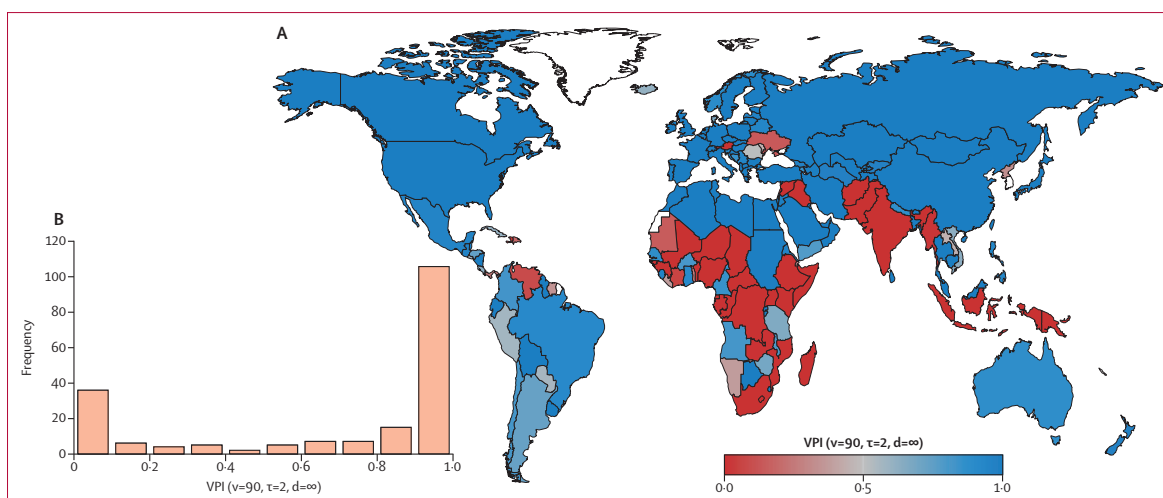
Figure 3: Extrapolation of current coverage trends based on the Gaussian process model

(A) 53 countries have more than a 50% chance of failing to meet Global Vaccine Action Plan 2015 targets if current diphtheria, tetanus, and pertussis (DTP3) coverage trends continue. 0·5 ≤ VPI < 0·9: Angola, Argentina, Armenia, Australia, Bosnia and Herzegovina, Brunei Darussalam, Burkina Faso, Cameroon, Chile, Colombia, Costa Rica, Cuba, Denmark, El Salvador, Honduras, Iceland, Jamaica, Kiribati, Malawi, Nauru, Nepal, Norway, Palau, Paraguay, Peru, Romania, Samoa, Senegal, Trinidad and Tobago, Tuvalu, Tanzania, Vietnam, Yemen, Zimbabwe. 0·1 ≤ VPI < 0·5: Barbados, Comoros, Korea, Djibouti, Dominican Republic, Laos, Liberia, Mauritania, Micronesia, Namibia, Panama, Moldova, San Marino, Solomon Islands, Suriname, Togo, Ukraine. VPI <0·1: Afghanistan, Austria, Benin, Central African Republic, Chad, Congo, Côte d'Ivoire, Democratic Republic of the Congo, Equatorial Guinea, Ethiopia, Gabon, Guinea, Guinea-Bissau, Haiti, India, Indonesia, Iraq, Kenya, Lebanon, Madagascar, Mali, Marshall Islands, Mozambique, Burma, Niger, Nigeria, Pakistan, Papua New Guinea, Somalia, South Africa, Syria, Timor-Leste, Uganda, Vanuatu, Venezuela, Zambia. VPI=vaccine performance index. (B) Histogram of index values.

Americas; figure 4A). An exception to this finding is the correlation with the service industry (the service socioeconomic factor is the net output of the service industry, defined as wholesale and retail trade, transport, and other governmental services such as education and health care[39]), which is reasonably high and increasing in Africa and the Americas. By contrast, in the Eastern Mediterranean and the Western Pacific region, socioeconomic factors have remained correlated over time (figure 4A). The South-East Asia region has a more mixed pattern of recent changes in socioeconomic connections with, for example, a decreasing link to rates of primary school completion and an increasing link to water access and sanitation (figure 4A).

In both Africa and the Eastern Mediterranean, educational variables and access to water are particularly informative about DTP3 coverage, whilst in the Eastern Mediterranean alone, government health spending, income metrics, access to sanitation, phone use, and CO_2 emissions are strongly linked with coverage (figure 4A). Socioeconomic correlates are generally low (both recently and historically) in Europe, and are reasonably low in the Americas (figure 4A). In South-East Asia, education ratio and access to water are notably high correlates that have not decreased in importance (figure 4A). In the Western Pacific region, births attended by skilled health staff is the strongest correlate (and one that appears to be becoming more informative of DTP3 coverage; figure 4A).

The association between public health spending and coverage varies globally, with a strong link in the Eastern Mediterranean and weaker links in South-East Asia,

Africa, and the Western Pacific (where the link is increasing), and the Americas (where it is decreasing; figure 4A). Only in the Eastern Mediterranean is gross domestic product per capita the strongest correlate of vaccine coverage; in all other regions the strongest links are with less directly economic factors (figure 4A).

We noted a significant difference in the correlation strength between DTP3 coverage and male and female education in Africa and South-East Asia (figure 4B; p value from paired t tests <0·0001 for both); correlations were not significant for the other regions at the p=0·05 level, but paternal education was more informative in the Western Pacific.

We plotted these correlations for the other vaccinations we examined; some notable observations arose from the comparison of these other vaccines with DTP3 (appendix). One regional trend was that the effect of public health spending and access to sanitation in South-East Asia, which was moderate for DTP3, was very high and often increasing for other vaccines. Another observation was that the Eastern Mediterranean, where we recorded stable or decreasing correlations with several socioeconomic factors for DTP3, instead showed increasing links with all factors (except investments) for BCG. The Americas, which had decreasing correlations for most factors with DTP3, showed decreases to almost Europe-like absences of coverage correlation in DTP1 and MCV, although all other regions and factors had similar patterns for DTP3 and MCV coverage. Europe is broadly the same across all vaccines except BCG, for which correlations seem to be negative between factors such as government health
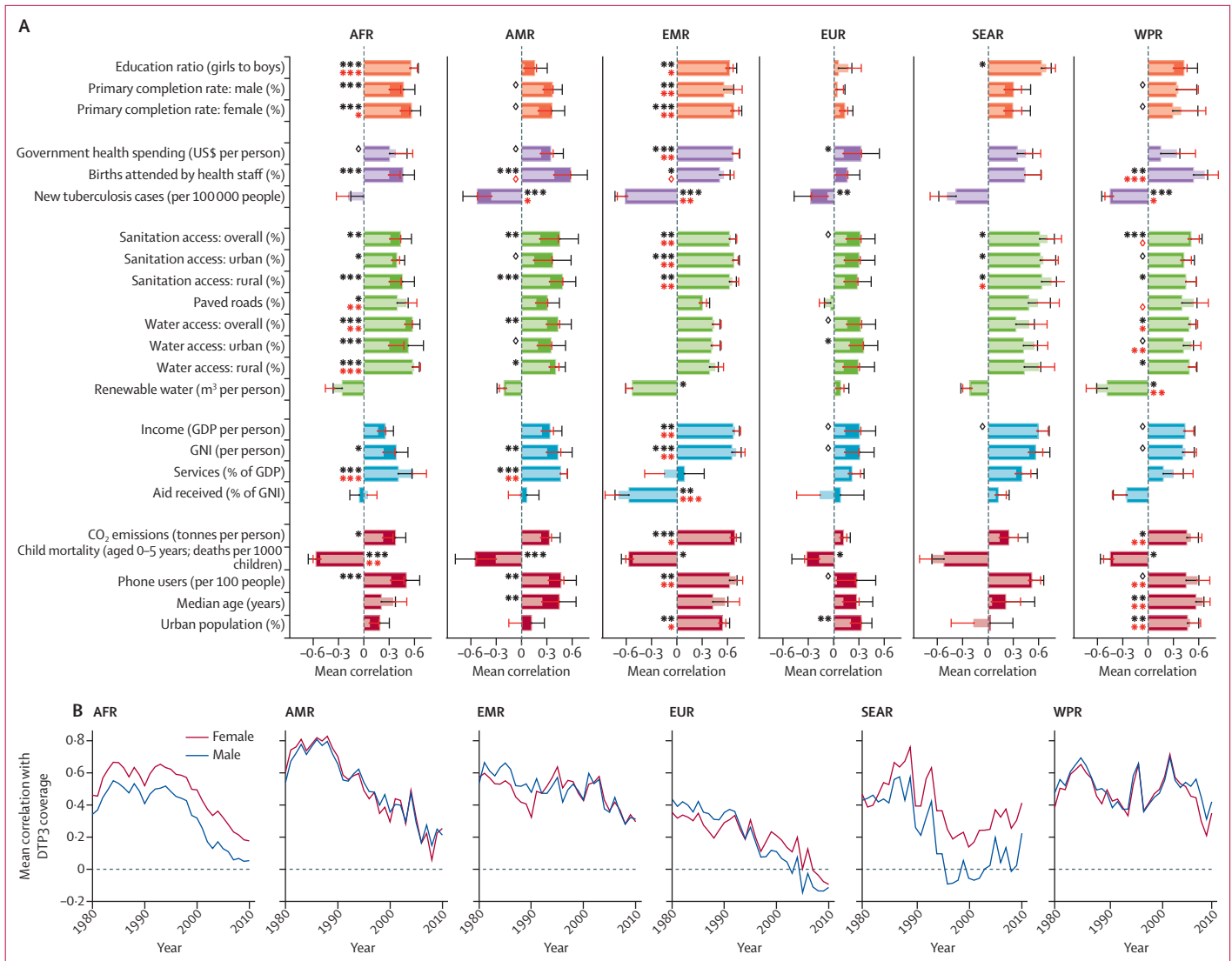
Figure 4: Historic and recent socioeconomic correlates with DTP3 coverage
(A) Groups of key socioeconomic factors, obtained by collecting the top correlating factors in each region across the time periods 1980–2010 and 2001–10. Only factors with the most consistent non-zero correlations with coverage are presented. Spearman's rho is represented by thick dark bars for 1980–2010, and by thin light bars for 2001–10. Mean p values are shown in black for 1980–2010 and in red for 2001–10: *** denotes p<0·001 , ** denotes p<0·01, * denotes p<0·05, and ◊ denotes p<0·10. Error bars represent 95% CIs. (B) Time series of the Spearman's ρ between DTP3 coverage and female (blue) and male (red) education. GDP=gross domestic product. GNI=gross national income. AFR=Africa. AMR=Americas. EMR=Eastern Mediterranean. EUR=Europe. SEAR=South-East Asia Region. WPR=Western Pacific region. DTP3=diphtheria, tetanus, and pertussis third dose.

spending, sanitation access, income, and BCG coverage. All estimated correlations in the study were robust to our imputation procedure (appendix).

Discussion
In keeping with findings from previous studies, our study found several socioeconomic factors to be informative of immunisation levels, such as correlates related to health-care facilities[3] and educational variables.[4–8] Literature from the past few years has indeed identified various socioeconomic and attitudinal barriers to vaccination coverage for a number of countries, such as parental education level,[4–8] age,[6,9,10] employment status

and workplace,[4,6,7] poverty,[14–18] inequity,[40] and distance to health-care facilities.[4,6,11,14,17,18] Findings from a systematic review[41] emphasised the link between out-of-hospital birth and reduced immunisation rates in countries with medium or low Human Development Index (HDI) scores, and associated the use of private health-care services and no health insurance with low immunisation levels in countries with very high HDI scores. We likewise found a link between income and births attended by skilled health staff and immunisation levels. We believe we are the first to identify several other potential barriers to immunisation coverage; for example, the link between governmental health spending and

primary school completion rates (in addition to income) across Eastern Mediterranean countries, and access to water and primary school completion rates in Africa.

We suggest that having a set of low-magnitude socioeconomic correlations with vaccine coverage corresponds to an encouraging state wherein vaccine access is available to most of the population. Conversely, when strong socioeconomic correlations exist, they signal potential limiting factors in vaccine access, although further analyses would be needed to infer causation between these factors and coverage. Hence, Europe, which has displayed consistently low socioeconomic correlations, enjoys high access to all vaccines; however, it suffers from non-infrastructural barriers to vaccination that centre on religious or philosophical beliefs and perceived risks rather than access, which we suspect is by contrast with other regions in which access is the main barrier.[3,42] Notably, the Americas has shown low (and decreasing) socioeconomic correlations with DTP1, MCV, and BCG, but higher correlations (although currently decreasing) associated with DTP3. This pattern suggests a transitional stage wherein vaccines in the region are broadly accessible, but that some factors (including attendance of medical staff at births) are restricting uptake of DTP3. This notion seems to be supported in the literature, which cites a range of socioeconomic determinants (including refusal from groups with higher socioeconomic status,[43,44] personal beliefs, and low access to health-care facilities) as barriers to vaccination.[3] We note the strength and consistency of births attended by skilled health staff as an informer of DTP3 coverage rates, and highlight the potential of this factor for use as a proxy indicator of the condition of a health-care system.

We speculate that regions where correlations are strengthening can be interpreted as more pressing targets for intervention. For example, the ratio of girls to boys in primary education, public health spending, and sanitation in South-East Asia and the Eastern Mediterranean; rural access to clean water (which is widely believed to be able to prevent polio in some parts of Nigeria[27]) and the presence of skilled health staff in Africa, where a potential surrogate, children born in a public or private health-care facility, is consistently linked to completion of routine vaccinations.[11]

Although an aim of the present study was to provide a global overview of correlates of immunisation coverage, we note several limitations. Regional variability is likely to exist within countries and also between other sets of countries grouped by regions other than those defined by the WHO. A large-scale global analysis investigating regional trends and identifying factors associated both within and between countries could therefore uncover more nuanced trends. Higher-frequency data from social media sources might be able to predict trends in vaccination uptake behaviour more accurately than can lower frequency national coverage data. The vaccine performance index methodology can be naturally integrated with models incorporating high-frequency time-series predictive variables to more accurately forecast immunisation levels as high-frequency data for vaccine sentiments becomes increasingly available. Such fine-grained analyses could facilitate the use of time lags in exploration of predictive models of vaccine coverage using combinations of high-frequency social media data and low-frequency socioeconomic factors.

Review of historical trends and variability in vaccination rates can aid policy makers' assessments of the strength and resilience of immunisation programmes. The 2014 GVAP report notes that a third of the world's 194 countries were failing to reach 90% coverage in 2013,[45] but these static summaries conceal both the vaccination track record and the possible future immunisation levels. A WHO report[46] notes that 65 countries had failed to achieve either "90% coverage nationally or 80% coverage in every district or equivalent administrative unit against diphtheria, tetanus and whooping cough by 2015", which is in line with predictions from our vaccine performance index.

We thus present our vaccine performance index as a tool for future research, which could enhance GVAP assessment reports for future targets by providing predictive measures that account for coverage trends and variability.

References
1 WHO. Weekly epidemiological record. *Wkly Epidemiol Rec* 2012; **87:** 421–36.
2 Black S, Rappuoli RA. A crisis of public confidence in vaccines. *Sci Transl Med* 2010; **2:** 61mr1.
3 Larson HJ, Jarrett C, Eckersberger E, Smith DMD, Paterson P. Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: a systematic review of published literature, 2007–2012. *Vaccine* 2014; **32:** 2150–59.
4 Ushie BA, Fayehun OA, Ugal DB. Trends and patterns of under-5 vaccination in Nigeria, 1990–2008: what manner of progress? *Child Care Health Dev* 2014; **40:** 267–74.
5 Calhoun LM, van Eijk AM, Lindblade KA, et al. Determinants and coverage of vaccination in children in western Kenya from a 2003 cross-sectional survey. *Am J Trop Med Hyg* 2014; **90:** 234–41.
6 Hu Y, Chen Y, Guo J, Tang X, Shen L. Completeness and timeliness of vaccination and determinants for low and late uptake among young children in eastern China. *Hum Vaccin Immunother* 2014; **10:** 1408–15.
7 Ueda M, Kondo N, Takada M, Hashimoto H. Maternal work conditions, socioeconomic and educational status, and vaccination of children: a community-based household survey in Japan. *Prev Med* 2014; **66:** 17–21.

8   Desai S, Alva S. Maternal education and child health: is there a strong causal relationship? *Demography* 1998; **35:** 71–81.

9   Otieno NA, Nyawanda BO, Audi A, et al. Demographic, socio-economic and geographic determinants of seasonal influenza vaccine uptake in rural western Kenya, 2011. *Vaccine* 2014; **32:** 6699–704.

10   Danis K, Georgakopoulou T, Stavrou T, Laggas D, Panagiotopoulos T. Socioeconomic factors play a more important role in childhood vaccination coverage than parental perceptions: a cross-sectional study in Greece. *Vaccine* 2010; **28:** 1861–69.

11   Canavan ME, Sipsma HL, Kassie GM, Bradley EH. Correlates of complete childhood vaccination in East African countries. *PLoS One* 2014; **9:** e95709.

12   Wagner KS, van Wijgerden JCJ, Andrews N, Goulden K, White JM. Childhood vaccination coverage by ethnicity within London between 2006/2007 and 2010/2011. *Arch Dis Child* 2014; **99:** 348–53.

13   Pande RP, Yazbeck AS. What's in a country average? Wealth, gender, and regional inequalities in immunization in India. *Soc Sci Med* 2003; **57:** 2075–88.

14   Galarce EM, Minsky S, Viswanath K. Socioeconomic status, demographics, beliefs and A(H1N1) vaccine uptake in the United States. *Vaccine* 2011; **29:** 5284–89.

15   Friederichs V, Cameron JC, Robertson C. Impact of adverse publicity on MMR vaccine uptake: a population based analysis of vaccine uptake records for one million children, born 1987–2004. *Arch Dis Child* 2006; **91:** 465–68.

16   Reading R, Colver A, Openshaw S, Jarvis S. Do interventions that improve immunisation uptake also reduce social inequalities in uptake? *BMJ* 1994; **308:** 1142–44.

17   Mitchell S, Andersson N, Ansari NM, Omer K, Soberanis JL, Cockcroft A. Equity and vaccine uptake: a cross-sectional study of measles vaccination in Lasbela District, Pakistan. *BMC Int Health Hum Rights* 2009; **9** (suppl 1)**:** S7.

18   Le Polain de Waroux O, Schellenberg JR, Manzi F, et al. Timeliness and completeness of vaccination and risk factors for low and late vaccine uptake in young children living in rural southern Tanzania. *Int Health* 2013; **5:** 139–47.

19   Atwell JE, Salmon DA. Pertussis resurgence and vaccine uptake: implications for reducing vaccine hesitancy. *Pediatrics* 2014; **134:** 602–04.

20   Benin AL, Wisler-Scher DJ, Colson E, Shapiro ED, Holmboe ES. Qualitative analysis of mothers' decision-making about vaccines for infants: the importance of trust. *Pediatrics* 2006; **117:** 1532–41.

21   MacDonald NE, Smith J, Appleton M. Risk perception, risk management and safety assessment: what can governments do to increase public confidence in their vaccine system? *Biologicals* 2012; **40:** 384–88.

22   Hak E, Schönbeck Y, De Melker H, Van Essen GA, Sanders EAM. Negative attitude of highly educated parents and health care workers towards future vaccinations in the Dutch childhood vaccination program. *Vaccine* 2005; **23:** 3103–07.

23   Marlow LAV, Waller J, Wardle J. Trust and experience as predictors of HPV vaccine acceptance. *Hum Vaccin* 2007; **3:** 171–75.

24   McMurray R, Cheater FM, Weighall A, Nelson C, Schweiger M, Mukherjee S. Managing controversy through consultation: a qualitative study of communication and trust around MMR vaccination decisions. *Br J Gen Pract* 2004; **54:** 520–25.

25   Allen JD, Othus MKD, Shelton RC, et al. Parental decision making about the HPV vaccine. *Cancer Epidemiol Biomarkers Prev* 2010; **19:** 2187–98.

26   Freimuth VS, Musa D, Hilyard K, Quinn SC, Kim K. Trust during the early stages of the 2009 H1N1 pandemic. *J Health Commun* 2014; **19:** 321–39.

27   Murele B, Vaz R, Gasasira A, Mkanda P, Erbeto T, Okeibunor J. Vaccine perception among acceptors and non-acceptors in Sokoto State, Nigeria. *Vaccine* 2014; **32:** 3323–27.

28   Nankabirwa V, Tylleskär T, Tumwine JK, Sommerfelt H, and the Promise-ebf Study Group. Maternal education is associated with vaccination status of infants less than 6 months in Eastern Uganda: a cohort study. *BMC Pediatr* 2010; **10:** 92–101.

29   Mira J, Subramanian SV, Sylvestre MP, et al. Association between maternal health literacy and child vaccination in India: a cross-sectional study. *J Epidemiol Community Health* 2015; **69:** 849–57.

30   Keane MT, Walter MV, Patel BI, et al. Confidence in vaccination: a parent model. *Vaccine* 2005; **23:** 2486–93.

31   Wolf E, Rowhani-Rahbar A, Duchin J, DeHart MP, Opel D. The challenges in measuring local immunization coverage: a statewide case study. *Pediatrics* 2016; **137:** e20153755.

32   Ozawa S, Stack ML. Public trust and vaccine acceptance– international perspectives. *Hum Vaccin Immunother* 2013; **9:** 1774–78.

33   Larson HJ, Schulz W. The state of vaccine confidence 2015. http://www.vaccineconfidence.org/The-State-of-Vaccine-Confidence-2015.pdf (accessed April 29, 2016).

34   WHO. Global vaccine action plan 2011–2020. 2013. http://apps.who.int/iris/handle/10665/78141 (accessed Aug 10, 2016).

35   WHO. WHO regional offices. www.who.int/about/regions/en (accessed Aug 8, 2013).

36   UNICEF. UNICEF statistics. www.childinfo.org/tables/Immunization/coverage/1980-2011.xls (accessed Aug 8, 2013).

37   Arevshatian L, Clements C, Lwanga S, et al. An evaluation of infant immunization in Africa: is a transformation in progress? *Bull World Health Organ* 2007; **85:** 449–57.

38   van der Maaten LJP, Hinton GE. Visualizing high-dimensional data using t-SNE. *J Mach Learn Res* 2008; **9:** 2579–605.

39   The World Bank. Services, etc, value added (% of GDP). http://data.worldbank.org/indicator/NV.SRV.TETC.ZS (accessed Aug 8, 2013).

40   Hinman AR, McKinlay MA. Immunization equity. *Am J Prev Med* 2015; **49** (6 suppl 4)**:** S399–405.

41   Tauil MC, Sato APS, Waldman EA. Factors associated with incomplete or delayed vaccination across countries: a systematic review. *Vaccine* 2016; **34:** 2635–43.

42   Yaqub O, Castle-Clarke S, Sevdalis N, Chataway J. Attitudes to vaccination: a critical review. *Soc Sci Med* 2014; **112:** 1–11.

43   Barata RB, Ribeiro MC, de Moraes JC, Flannery B, and the Vaccine Coverage Survey 2007 Group. Socioeconomic inequalities and vaccination coverage: results of an immunisation coverage survey in 27 Brazilian capitals, 2007–2008. *J Epidemiol Community Health* 2012; **66:** 934–41.

44   Kim SS, Frimpong JA, Rivers PA, Kronenfeld JJ. Effects of maternal and provider characteristics on up-to-date immunization status of children aged 19 to 35 months. *Am J Public Health* 2007; **97:** 259–66.

45   WHO, Strategic Advisory Group of Experts on Immunization. 2014 assessment of the Global Vaccine Action Plan. http://www.who.int/immunization/global_vaccine_action_plan/SAGE_DoV_GVAP_Assessment_report_2014_English.pdf (accessed July 11, 2015).

46   WHO. World immunization week 2016: immunization game-changers should be the norm worldwide. April 21, 2016. http://www.who.int/mediacentre/news/releases/2016/world-immunization-week/en/ (accessed April 29, 2016).