Determinants of basic childhood vaccination coverage in European and OECD countries

Vladimira Varbanova, Frederik Verelst, Niel Hens & Philippe Beutels

View supplementary material

Published online: 29 Sep 2022.

Submit your article to this journal

Article views: 698

View related articles

View Crossmark data

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

🔓 OPEN ACCESS | Check for updates

Determinants of basic childhood vaccination coverage in European and OECD countries

Vladimira Varbanova [ID][a], Frederik Verelst [ID][a], Niel Hens [ID][a,b], and Philippe Beutels [ID][a]

[a]Centre for Health Economics Research and Modelling Infectious Diseases, Vaccine & Infectious Disease Institute, University of Antwerp, Antwerp, Belgium; [b]Center for Statistics (CenStat, Interuniversity Institute of Biostatistics and statistical Bioinformatics (I-BioStat) and Data Science Institute, Hasselt University, Hasselt, Belgium

ABSTRACT

Vaccination coverage varies between countries and over time. Using official databases, we extracted data on 50 national-level immunization, socio-economic, demographic, healthcare, and cultural factors, and the uptake of the third dose of diphtheria toxoid, tetanus toxoid, and pertussis vaccines (DTP3) and the first dose of measles-containing vaccines (MCV1) for 61 countries between 1990 and 2019. The main branch of the analysis included all covariates, while a secondary branch excluded life-expectancy and child mortality. The statistical analysis was completed in three stages: a variable-selection stage via random forests; multilevel multiple imputation for missing data in the reduced dataset; and generalized estimating equations (GEE) over all imputed datasets with pooled results. Less than 20 covariates were retained after variable-selection. Among a relatively small number of statistically significant (p-value <.05) effects in the pooled GEE results of our main branch, under-5 mortality and long-term orientation culture showed negative associations with both uptake outcomes and GDP per capita a positive association. For MCV1, whether a second dose was integrated into routine immunization appeared as the overall strongest negative correlate. In the secondary analytical branch, results were largely consistent, with a few additional statistically significant effects emerging, mainly related to immunization and healthcare system characteristics. These insights improve our understanding of the main factors influencing vaccine uptake, some of which are broadly contextual (e.g., GDP, socio-cultural factors), requiring bespoke vaccine program approaches, in order to maximize childhood vaccine uptake over time.

Introduction

Vaccination has been lauded as one of the greatest achievements of public health.[1,2] The World Health Organization (WHO) estimates that 2–3 million deaths are prevented yearly thanks to immunization.[3] Benefits of vaccines go beyond prevention of infection in the vaccinated and reduction of pathogens' circulation in the broader community. Indeed, the economic and societal impact has been shown to be far-reaching: from saving healthcare costs and averting productivity losses, to better cognitive development, educational achievement, and subsequent productivity and wealth,[4–7] as well as greater health equity and social integration.[8] Despite tremendous successes in recent decades (such as smallpox eradication and regional polio elimination), outbreaks of vaccine-preventable diseases continue to occur even in countries with relatively high vaccination coverage rates,[8] at least in part due to the underimmunization of specific subgroups in the population.[9,10] In order to maintain health, economic, and societal gains, continued surveillance of vaccine uptake and a better understanding of the drivers behind it are needed.

Much scientific research to date has focused on the topics of vaccination status, immunization timeliness, and vaccine hesitancy. The latter has been identified among the top 10 threats to global health by the WHO[11] and, while generally more common in wealthy nations,[8] this phenomenon has been observed in both "developed" and "developing" countries.[12] Recent (systematic) reviews[13–20] summarize empirical findings with regard to a large number of vaccination determinants, such as immunization program and logistics (e.g., vaccine availability and storage, accessibility, affordability), as well as child, parental (education, knowledge, attitudes), and family (socio-economic status, location) characteristics. Using survey, interview, or administrative data, previous research has investigated determinants and outcomes at the individual-level, with multi-country studies, for the most part, focusing on low- and middle-income countries (LMIC). The challenges in achieving and maintaining high vaccination coverage differ fundamentally between countries of completely different income levels. It is therefore of interest to focus on countries within a limited range of income variation. Indeed, important variations in vaccination coverage are observed among relatively wealthy countries, even within Europe. Despite many historical similarities, European countries tend to fall into different groups of sociocultural contexts [21,22] and have made different choices in the establishment of their health systems and implementation of their vaccination programs. Since vaccination is widely considered as one of the most important tools in public health,

---

CONTACT Vladimira Varbanova ✉ vladimira.varbanova@uantwerpen.be 🏛 University of Antwerp, Campus Drie Eiken (D.S. 243), Universiteitsplein 1, Antwerp B-2610, Belgium.

there is a continued need to examine which country and time-dependent factors may influence childhood vaccination coverage the most. This need seems obvious: we can learn from country differences to improve our understanding of the relative importance of different factors. Examining rich datasets to investigate the influence of a wide array of potential determinants of vaccination coverage rates in European and in other predominantly wealthy countries over time should therefore be of high interest to vaccine program managers and policy-makers.

Materials & methods

We were interested in performing analyses on a diverse and large sample of relatively wealthy countries that belong to Europe or to the Organization for Economic Co-operation and Development (OECD). Appendix A presents a table with the rankings of the 61 countries included in this study according to gross national income (GNI) for 2019, the last year we considered. As per the World Bank's classification by income level[23] from the 1st of July 2020, all but one country fall either in the "high income" or the "upper-middle income" brackets. This selection was thought to allow us to collate a rich dataset from existing databases for both the outcomes and potential determinants, allowing for a thorough analytical approach.

Outcome variables

We considered the nation-wide coverage of the third dose of diphtheria toxoid, tetanus toxoid, and pertussis vaccines (DTP3) and the first dose of measles-containing vaccines (MCV1) as outcomes, which are traditionally used as proxy performance measures of childhood immunization programs.[24–29] Because of data-availability concerns and COVID-19's negative impact on routine immunization efforts worldwide,[29,30] we decided to take into consideration the annual WHO and United Nations International Children's Emergency Fund (UNICEF) coverage estimates for these two vaccines[31] for the 30-year period leading up to 2019 (i.e., 1990–2019). Figure 1 shows that among 2019 coverage rates for the 61 European or OECD member countries included in our study, all but a few stand above the global average of 85%.[26] The interested reader can also find an interactive visualization of the longitudinal national coverage profiles of the countries included in Appendix B or at https://www.simid.be/software/dtp3/.

Covariates

A full list of the 50 covariates included in this study can be found in Table 1. Ten WHO immunization coverage indicators applicable to one or both of the vaccines of interest were taken into account.[32] In addition, we developed a new financial variable in order to measure relative expenditure on immunization activities. For this purpose, we used the WHO indicator "What is the total expenditure (from all sources) on routine immunization?," whose original values were reported either in USD or in local currency (LCU). If reported in local currency, values were first converted to USD according to the official

exchange rate (LCU per USD, period average) of the International Monetary Fund, as reported by the World Bank.[33] All amounts were then divided by the respective country's nominal GDP[33] in order to obtain a measure reflecting immunization expenditure relative to GDP. Furthermore, 20 general economic, demographic, population health, political stability, and healthcare system indicators were collected from two World Bank databases,[33,34] and data on the average number of completed years of education was retrieved from the United Nations Educational, Scientific and Cultural Organization (UNESCO) Institute for Statistics (UIS).[35] The World Values Survey (WVS) was used as a source of information on social, political, cultural, economic, and religious values. WVS is a widely used and cited noncommercial long-term project investigating human beliefs and values across the world. As of May 2022, the WVS has been executed in over 120 societies, in 7 waves. For the purposes of our study, we extracted 12 aggregate indicators from the "WVS Wave 1 to 6 Key Aggregates" dataset.[36] Finally, we also considered information from the 6-dimensional Hofstede national culture model.[37] Inspired by an IBM International personnel survey in the 1960s, this model has been continuously developed and currently includes six indices, based on data from over 100 countries worldwide. The six dimensions of national culture and their definitions,[38] as conceived by Hofstede and colleagues, can also be found in Table 1.

Statistical analysis

Our initial dataset consisted of two outcome variables and 50 covariates, for 61 countries, over a 30-year time-span (1990–2019). The dataset had a two-level structure, as the six Hofstede cultural indices are considered time-invariant and thus have just one value per country. The relationship between 48 and 49 covariates and DTP3 and MCV1 national vaccination rates, respectively, was investigated separately for each outcome in the main branch of the analysis. In an additional, secondary, branch, we excluded life-expectancy at birth and under-5 mortality, because of a possible bi-directional causal relationship, leaving 46 covariates for DTP3 and 47 for MCV1.

The statistical analysis was completed in three stages, executed in turn for each outcome in each branch (with and without life-expectancy and under-5 mortality). In the first stage, we employed the random forests (RF)[39] procedure as a variable-selection technique. Fine-tuning model parameters, we ran the procedure a total of 11 times (refer to Appendix C for full details on the statistical methods employed). In order to determine the most important covariates, we looked at the variable importance indicator and ranked covariates according to the number of times each appeared among the "top 20" variable importance scores. For a reduction of over 60%, down from 48–49 to 18–19 covariates, only those with a frequency of at least 8 out of 11 were retained for further investigation. In the second stage of the analysis, we performed multilevel multiple imputation in order to address the issue of missing data. Missingness after variable selection was 19% in the main branch and 24% in the secondary branch, over the total of 42,090 data points in both reduced dataset (30 time points for 61 countries for 23 variables).
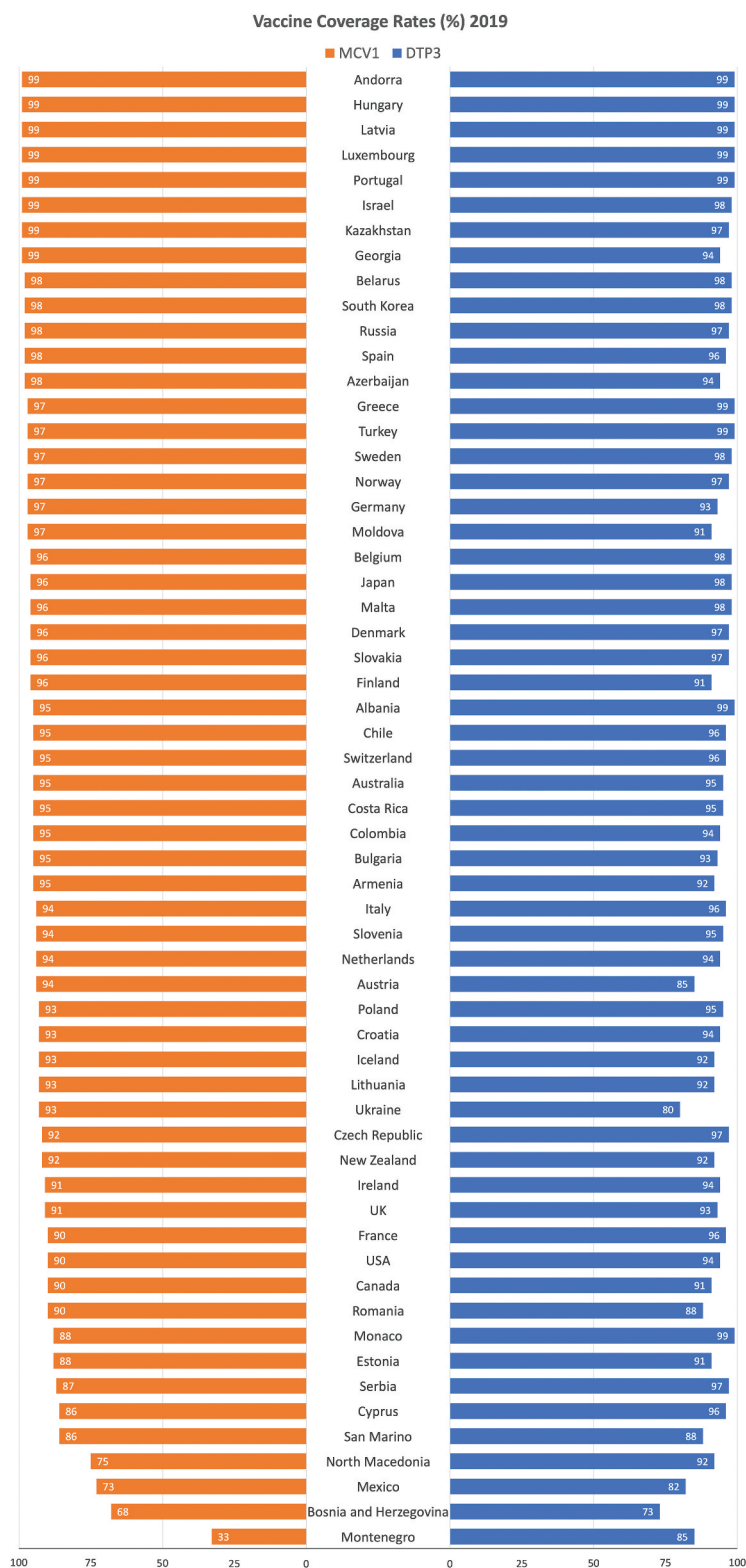
Figure 1. First dose measles-containing vaccine (MCV1) and third dose diphtheria-tetanus-pertussis (DTP3) vaccination coverage rates for the 61 countries included in the study (only 2019 shown).

Before implementing multiple imputation, the arcsine transformation was applied to the two coverage rates expressed as proportions, and then all variables were standardized to have a mean of 0 and a standard deviation of 1. Using multivariate imputation by chained equations (MICE)[40] predictive mean matching (PMM), we obtained 20 imputed datasets, which were then analyzed individually using generalized estimating equations (GEE),[41] in the third stage of the analysis. Finally, results from the 20 individual GEE models were pooled together to get an estimate for each covariate's effect. See

Table 1. Initial list of covariates.

	Variable	Possible values	Sources	Last accessed
1.	Number of adverse events following immunization (AEFI) reported	numeric	WHO[c]	25.11.2021
2.	Percentage of total expenditure on routine immunization financed by government funds*	numeric (0–100)	WHO[c]	25.11.2021
3.	What percentage (%) of the cold chain equipment (at all sub-national levels of the system) is equipped with electronic continuous temperature monitoring systems?*	numeric (0–100)	WHO[c]	25.11.2021
4.	Has the country a multi-year plan (MYP) for immunization?	Yes/No	WHO[c]	26.11.2021
5.	Is Measles 2nd dose vaccine integrated into the routine immunization systems?	Yes/No/Yes[P]	WHO[c]	29.11.2021
6.	Vaccination services interrupted because of lack of vaccine for Measles?*	Yes/No	WHO[c]	25.11.2021
7.	Vaccination services interrupted because of lack of vaccine for DTP?*	Yes/No	WHO[c]	25.11.2021
8.	Have districts (or levels below) an electronic vaccine stock management system?*	Yes/No	WHO[c]	25.11.2021
9.	Does the immunization programme have a dedicated immunization supply chain manager at national level?*	Yes/No	WHO[c]	25.11.2021
10.	Has the country a standing technical advisory group on immunization (NITAG)?	Yes/No	WHO[c]	24.11.2021
11.	Immunization expenditure relative to GDP*^	numeric	WHO[c] TWB[d]	29.11.2021 30.11.2021
12.	GDP per capita, PPP (current international $)	numeric	TWB[d]	30.11.2021
13.	Current health expenditure (% of GDP)	numeric (0–100)	TWB[d]	30.11.2021
14.	Life-expectancy at birth, total (years)	numeric	TWB[d]	30.11.2021
15.	Population, total	numeric	TWB[d]	30.11.2021
16.	Unemployment, total* (% of total labor force) (modeled ILO estimate)	numeric (0–100)	TWB[d]	30.11.2021
17.	Birth rate, crude* (per 1,000 people)	numeric	TWB[d]	30.11.2021
18.	Births attended by skilled health staff (% of total)	numeric (0–100)	TWB[d]	30.11.2021
19.	Community health workers* (per 1,000 people)	numeric	TWB[d]	30.11.2021
20.	Hospital beds (per 1,000 people)	numeric	TWB[d]	30.11.2021
21.	Nurses and midwives (per 1,000 people)	numeric	TWB[d]	30.11.2021
22.	Physicians (per 1,000 people)	numeric	TWB[d]	30.11.2021
23.	Population ages 15-64* (% of total)	numeric (0–100)	TWB[d]	30.11.2021
24.	Population ages 65 and above (% of total)	numeric (0–100)	TWB[d]	30.11.2021
25.	Population growth* (annual %)	numeric (0–100)	TWB[d]	30.11.2021
26.	Pregnant women receiving prenatal care* (%)	numeric (0–100)	TWB[d]	30.11.2021
27.	Surface area (sq. km)	numeric	TWB[d]	30.11.2021
28.	Specialist surgical workforce* (per 100,000 population)	numeric	TWB[d]	30.11.2021
29.	Mortality rate, under-5 (per 1,000 live births)	numeric	TWB[d]	30.11.2021
30.	Urban population (% of total)	numeric (0–100)	TWB[d]	30.11.2021
31.	Political stability and absence of violence/terrorism	numeric (−2.5–2.5)	TWB[e]	06.12.2021
32.	Average number of completed years of education of a country's population aged 25 years and older, excluding years spent repeating individual grades.*	numeric	UNESCO[f]	24.11.2021
33.	Emancipative values index* (national culture's emphasis on freedoms in the domains of reproductive choice, gender equality, people's voice, and personal autonomy)	numeric (0–1)	WVS[g]	24.11.2021
34.	Secular values index* (national culture's secular distance to "sacred" sources of authority – religious, patrimonial, normative, and order institutions)	numeric (0–1)	WVS[g]	24.11.2021
35.	Social movement activities* (the extent to which the peaceful social movement activities of petitions, demonstrations, and boycotts are part of a national culture's action repertoire)	numeric (0–1)	WVS[g]	24.11.2021
36.	Informational connectedness* (the diversity of information sources used by the average individual in a nation)	numeric (0–1)	WVS[g]	24.11.2021
37.	Liberal understanding of democracy* (the extent to which people's understanding of democracy is liberal in the sense that they define it "correctly" by its liberal, including free elections, civil liberties and equal rights)	numeric (0–1)	WVS[g]	24.11.2021
38.	Perceived democraticness of own country*	numeric (0–1)	WVS[g]	24.11.2021
39.	Perceived fairness of other people*	numeric (0–1)	WVS[g]	24.11.2021
40.	Standard trust toward unspecified other people*	numeric (0–1)	WVS[g]	24.11.2021
41.	Associational activity* (in recreational/humanitarian/environmental, church/religious, political parties/labor unions/professional associations)	numeric (0–1)	WVS[g]	24.11.2021
42.	Perceived health*	numeric (0–1)	WVS[g]	24.11.2021
43.	Perceived choice in shaping one's life*	numeric (0–1)	WVS[g]	24.11.2021
44.	Life satisfaction*	numeric (0–1)	WVS[g]	24.11.2021
45.	Power distance ("the degree to which the less powerful members of a society accept and expect that power is distributed unequally")	numeric (0–100)	Hofstede 6-D model of national culture[h]	26.11.2021
46.	Individualism (v Collectivism; "a preference for a loosely-knit social framework in which individuals are expected to take care of only themselves and their immediate families")	numeric (0–100)	Hofstede 6-D model of national culture[h]	26.11.2021
47.	Masculinity (v Femininity; "a preference in society for achievement, heroism, assertiveness, and material rewards for success")	numeric (0–100)	Hofstede 6-D model of national culture[h]	26.11.2021
48.	Uncertainty avoidance ("the degree to which the members of a society feel uncomfortable with uncertainty and ambiguity")	numeric (0–100)	Hofstede 6-D model of national culture[h]	26.11.2021

(*Continued*)

Table 1. (Continued).

Variable	Possible values	Sources	Last accessed
49. Long-term orientation (v Short-term Orientation; the degree to which society holds the notion that the world "is in flux, and preparing for the future is always needed" as opposed to the view that "the world is essentially as it was created, so that the past provides a moral compass")	numeric (0–100)	Hofstede 6-D model of national culture[h]	26.11.2021
50. Indulgence (v Restraint; the extent to which society "allows relatively free gratification of basic and natural human drives related to enjoying life and having fun")	numeric (0–100)	Hofstede 6-D model of national culture[h]	26.11.2021

*Variable not retained after the first, random forests, stage of the analysis.
^See section Materials & methods (Covariates) for explanation on how this variable was derived.
[c]https://immunizationdata.who.int/listing.html.
[d]https://databank.worldbank.org/source/world-development-indicators.
[e]https://info.worldbank.org/governance/wgi/.
[f]http://data.uis.unesco.org/#.
[g]https://www.worldvaluessurvey.org/WVSNewsShow.jsp?ID=367&ID=367.
[h]https://geerthofstede.com/research-and-vsm/dimension-data-matrix/.
Abbreviations: WHO = World Health Organization; TWB = The World Bank; UNESCO = United Nations Educational, Scientific and Cultural Organization; WVS = World Values Survey.

Figure 2 for a diagram of the main branch of the analysis (the secondary one is identical, apart from the fact that it contains two covariates less – life-expectancy and under-5 mortality).

Results

Table 2 shows the covariates retained in the main branch of the analysis, after completing the first, variable-selection, stage. Covariates are listed according to the data source (i.e., in no particular order) and the number of RF models (out of 11 possible) in which they had importance scores among the "top 20" is shown as well, in case they were retained for a particular outcome. Nineteen covariates were retained for DTP3, 16 of which were also among the 18 covariates retained for MCV1. According to our established criterion, nurses and midwives, political stability and absence of violence/terrorism, and surface area turned out to be important only for DTP3, while births attended by skilled health staff (and, naturally, whether the second dose of the Measles vaccine is integrated into the routine immunization system) were only to be included with MCV1 as an outcome. Leaving life-expectancy and under-5 mortality out in the second, additional, branch of the analysis did not have a significant impact on the selection of variables, as can be seen in Table 3. Eighteen covariates were
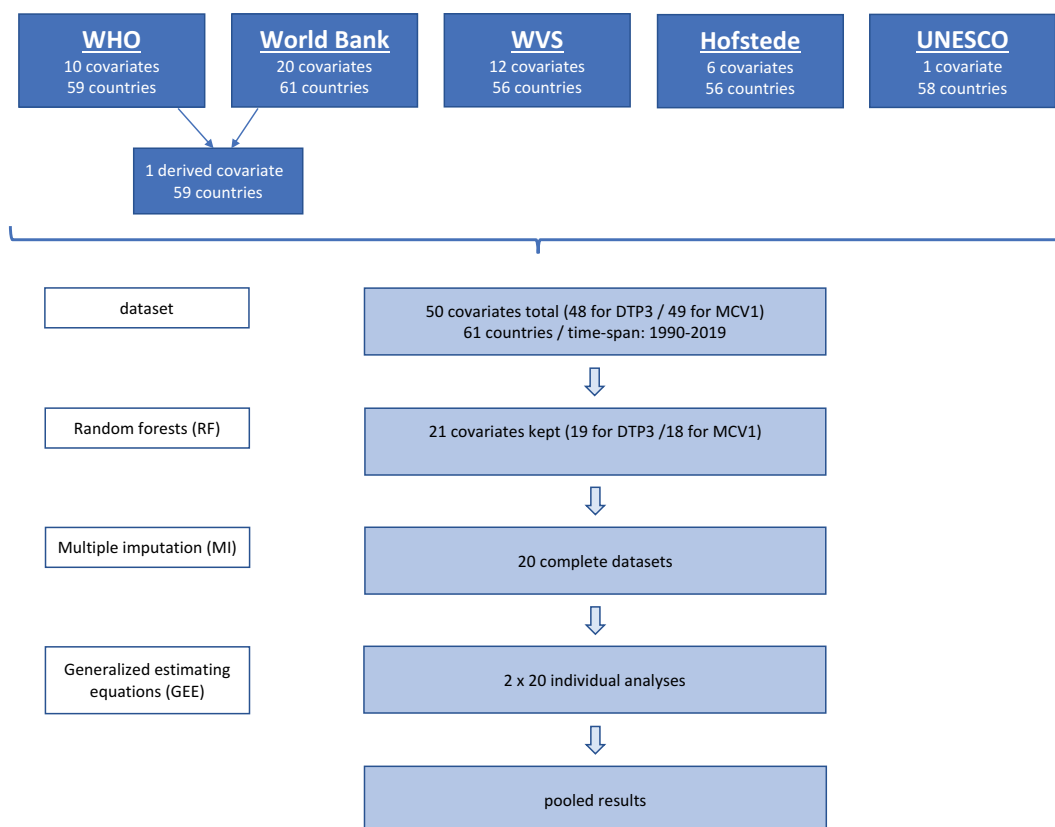

Figure 2. Flow chart of analytical steps (example of the main branch).

Table 2. Covariates retained in the main branch of the analysis based on variable importance scores in the random forests (RF) procedure (if a covariate was retained for a given outcome, its score here shows in how many of the 11 runs of the RF procedure it was among the "top 20"*).

covariate	DTP3	MCV1
Has the country a multi-year plan (MYP) for immunization?	9	10
Is Measles 2nd dose vaccine integrated into the routine immunization systems?**		11
Births attended by skilled health staff		10
Current health expenditure	10	11
GDP per capita	10	11
Hospital beds	10	10
Life-expectancy at birth	11	11
Nurses and midwives	8	
Physicians	9	11
Political stability and absence of violence/terrorism	8	
Population, total	11	11
Population ages 65 and above	11	9
Surface area	10	
Under-5 mortality rate	11	11
Urban population	10	11
Individualism	10	11
Indulgence	11	11
Long-term orientation	11	11
Masculinity	11	10
Power distance	10	8
Uncertainty avoidance	10	8

*See detailed explanation in section Materials & methods (Statistical analysis).
**Covariate only included in the MCV1 analysis.

Table 3. Covariates retained in the branch of the analysis without life-expectancy and under-5 mortality based on variable importance scores in the random forests (RF) procedure (if a covariate was retained for a given outcome, its score here shows in how many of the 11 runs of the RF procedure it was among the "top 20"*).

covariate	DTP3	MCV1
Number of adverse events following immunization (AEFI) reported?		9
Has the country a multi-year plan (MYP) for immunization?	11	10
Has the country a standing technical advisory group on immunization (NITAG)?		11
Is Measles 2nd dose vaccine integrated into the routine immunization systems?**		11
Births attended by skilled health staff	8	8
Current health expenditure	11	11
GDP per capita	11	10
Hospital beds	10	9
Nurses and midwives	11	
Physicians	9	11
Political stability and absence of violence/terrorism	11	11
Population, total	11	11
Population ages 65 and above	11	11
Surface area	10	10
Urban population	11	11
Individualism	11	
Indulgence	11	11
Long-term orientation	11	11
Masculinity	11	9
Power distance	11	
Uncertainty avoidance	11	9

*See detailed explanation in section Materials & methods (Statistical analysis).
**Covariate only included in the MCV1 analysis.

retained for both DTP3 and MCV1. In comparison with the main branch, surface area, political stability and absence of violence/terrorism, number of adverse events following immunization, and whether the country has a standing technical advisory group on immunization were retained for the analysis of MCV1 now, while individualism and power distance from

the Hofstede cultural indices were dropped. The only difference for the DTP3 selection in this branch compared to the main one was the inclusion of births attended by skilled health staff now.

Tables 4 and 5 show our final results for the main and additional branches of the analysis, respectively, after pooling the GEE effect estimates over the multiple imputed and separately analyzed datasets. Starting with DTP3 in the main analysis (Table 4), we observe the strongest overall and strongest negative statistically significant (at the conventional 0.05-level) association with under-5 mortality (coefficient of −0.413). The second strongest, this time positive, effect is that of power distance (0.374), followed by GDP per capita (0.259) and political stability and absence of violence/terrorism (0.185), both positively related as well. Long-term orientation (−0.165), surface area (−0.138), and whether the country has a multi-year plan for immunization (coded as 0=no/1=yes; coefficient of −0.097) all have smaller negative statistically significant effects. With regard to MCV1, whether the second dose of the vaccine is integrated into the routine immunization system shows a very strong negative association (−0.924), meaning that as the status of the second dose changes from "not integrated," to "partially integrated," to "integrated," MCV1 coverage rates tend to decrease. Furthermore, under-5 mortality shows a negative effect (−0.428), very similar in magnitude to the one we see for DTP3. Lastly, indulgence (−0.323) and long-term orientation (−0.218) both exhibit a negative relation to MCV1, while GDP (−0.319) and urbanization (0.274) show a positive relation.

Examining the pooled estimates from the secondary analysis (Table 5), we see a larger number of statistically significant effects. The additional effects we found here can be assumed to have been previously "absorbed" by those of life-expectancy and under-5 mortality, which were included as covariates in the main analysis. Overall, the results are consistent with those of the main analysis, especially regarding the effects of the Hofstede cultural indices, GDP per capita, and the indicator pertaining to the second MCV dose. Births attended by skilled health staff now show a statistically significant positive relation with both outcomes (coefficient of 0.190 for DTP3 and 0.156 for MCV1), while, on the contrary, whether the country has a multi-year immunization plan is no longer significant for DTP3 (the effect was significant, but very small before). In addition, the proportion of the population living in urban areas now shows a positive effect on DTP3 (coefficient of 0.223, being barely non-significant in the main analysis before with a *p*-value of 0.06), while the density of nurses and midwives exhibits a smaller, negative, association (−0.108). With regard to MCV1, the indicator on whether the country has a standing technical advisory group on immunization (not selected for retention in the main analysis) shows a relatively strong negative effect (−0.563), meaning that countries with such an advisory group tend to have lower MCV1 coverage rates. Finally, the current health expenditure and density of physicians also showed a statistically significant negative, albeit much weaker, relation (coefficients of −0.155 and −0.104, respectively).

Table 4. Pooled generalized estimating equations (GEE) results from the main branch of the analysis (statistically significant effects at the 0.05-level are shown in bold).

Covariate	DTP3 effect (std error; p-value)	MCV1 effect (std error; p-value)
Has the country a Multi-Year Plan (MYP) for immunization?	**−0.097 (0.039; 0.016)**	−0.054 (0.044; 0.230)
Is Measles 2nd dose vaccine integrated into the routine immunization systems?*	n/a	**−0.924 (0.408; 0.028)**
Births attended by skilled health staff	n/a	0.056 (0.049; 0.263)
Current health expenditure	−0.118 (0.060; 0.058)	−0.085 (0.073; 0.252)
GDP per capita	**0.259 (0.056; <0.0001)**	**0.319 (0.065; <0.0001)**
Hospital beds	0.049 (0.050; 0.338)	0.026 (0.063; 0.681)
Life-expectancy at birth	−0.141 (0.091; 0.130)	−0.278 (0.141; 0.060)
Nurses and midwives	−0.049 (0.052; 0.350)	n/a
Physicians	0.056 (0.041; 0.181)	−0.066 (0.047; 0.163)
Political stability and absence of violence/terrorism	**0.185 (0.045; 0.0001)**	n/a
Population	0.080 (0.042; 0.064)	0.003 (0.042; 0.935)
Population ages 65 and above	−0.049 (0.058; 0.398)	−0.032 (0.070; 0.654)
Surface area	**−0.138 (0.040; 0.001)**	n/a
Under-5 mortality rate	**−0.413 (0.065; <0.0001)**	**−0.428 (0.092; <0.0001)**
Urban population	0.170 (0.088; 0.062)	**0.274 (0.093; 0.006)**
Individualism	0.127 (0.069; 0.065)	−0.215 (0.122; 0.092)
Indulgence	−0.135 (0.118; 0.260)	**−0.323 (0.125; 0.016)**
Long-term orientation	**−0.165 (0.063; 0.013)**	**−0.218 (0.063; 0.001)**
Masculinity	−0.052 (0.056; 0.361)	0.081 (0.084; 0.347)
Power distance	**0.374 (0.080; <0.0001)**	0.103 (0.097; 0.297)
Uncertainty avoidance	0.065 (0.071; 0.377)	−0.079 (0.096; 0.413)

*Covariate only included in the MCV1 analysis.

Table 5. Pooled generalized estimating equations (GEE) results from the branch of the analysis without life-expectancy and under-5 mortality (statistically significant effects at the 0.05-level are shown in bold).

Covariate	DTP3 effect (std error; p-value)	MCV1 effect (std error; p-value)
Has the country a Multi-Year Plan (MYP) for immunization?	−0.037 (0.054; 0.504)	0.074 (0.040; 0.069)
Has the country a standing technical advisory group on immunization (NITAG)?	n/a	**−0.563 (0.118; <0.0001)**
Is Measles 2nd dose vaccine integrated into the routine immunization systems?*	n/a	**−0.996 (0.353; 0.006)**
Number of adverse events following immunization (AEFI) reported?	n/a	−0.066 (0.034; 0.055)
Births attended by skilled health staff	**0.190 (0.034; <0.0001)**	**0.156 (0.044; 0.001)**
Current health expenditure	−0.092 (0.060; 0.134)	**−0.155 (0.064; 0.021)**
GDP per capita	**0.281 (0.050; <0.0001)**	**0.166 (0.055; 0.004)**
Hospital beds	0.040 (0.052; 0.449)	−0.063 (0.049; 0.204)
Nurses and midwives	**−0.108 (0.051; 0.038)**	n/a
Physicians	0.051 (0.057; 0.377)	**−0.104 (0.043; 0.020)**
Political stability and absence of violence/terrorism	**0.168 (0.055; 0.005)**	0.010 (0.060; 0.870)
Population	0.068 (0.044; 0.131)	0.053 (0.041; 0.193)
Population ages 65 and above	−0.027 (0.061; 0.655)	0.049 (0.052; 0.344)
Surface area	**−0.142 (0.039; 0.0004)**	0.058 (0.032; 0.075)
Urban population	**0.223 (0.072; 0.003)**	**0.190 (0.071; 0.012)**
Individualism	0.152 (0.106; 0.162)	n/a
Indulgence	−0.144 (0.094; 0.134)	**−0.343 (0.076; <0.0001)**
Long-term orientation	**−0.135 (0.059; 0.027)**	**−0.151 (0.055; 0.008)**
Masculinity	−0.094 (0.071; 0.197)	0.040 (0.053; 0.455)
Power distance	**0.402 (0.072; <0.0001)**	n/a
Uncertainty avoidance	0.020 (0.099; 0.839)	−0.004 (0.049; 0.940)

*Covariate only included in the MCV1 analysis.

Discussion

With the present study, we attempted to uncover national-level determinants of vaccination coverage among a large number of factors related to immunization and healthcare system organization, economics, demographics, politics, culture, finances, and societal values in European states and OECD member countries. This study provides a marked expansion of most previous research, as we considered a large array of data for over 60 countries simultaneously, over the span of 30 years. We applied rigorous statistical methods to select variables, correct for missing data, and, finally, estimate longitudinal effects. Our results point to a number of interesting findings.

To begin with, one of the most stable effects observed in our analysis was that of GDP per capita. Often used as a proxy for development and wealth, this is the most widely studied and established positive determinant of general population health.[42] It is reasonable to expect nations experiencing economic prosperity to be more successful in the organization and financing of their healthcare systems, and, thus, immunization activities, by extension. However, when GDP per capita is already high, the marginal gains of rising GDP as a contributor to basic healthcare and health correlate are expected to decrease. Yet, we find that even among relatively wealthy countries with already relatively high vaccination coverage, GDP per capita still emerges, on average, as an

important determinant for vaccination coverage differences between countries and over time. Political stability and armed conflict would clearly be factors that undermine immunization efforts, as vaccination logistics then become more difficult to maintain and deploy. Indeed, the political stability and absence of violence/terrorism covariate in our analysis showed an important and persistent positive effect on DTP3 vaccination coverage – a finding also supported by other research.[43,44]

Another pair of determinants with anticipated observed effects in our study is that of surface area and urbanization. A larger territory (as in Canada, Russia, the USA, and Australia) means that remote areas are harder to reach and therefore likely underserved by routine immunization programs, and that healthcare and other policies may be more diverse and fragmented within the country compared to smaller-area countries. The negative effect of surface area in our analysis was statistically significant only for DTP3, perhaps due to the difference in age at vaccination relating to access to vaccination and ease of travel (typically, DTP3 is administered at around 16 weeks, and MCV1 at around 12–15 months of age). In accord, a number of individual-level studies in high-income countries have reported parents stating practical barriers as reasons for not vaccinating their children.[18] On the other hand, high levels of urbanization affect vaccination coverage rates in a positive way. The relationship is apparent in countries like Belgium, Israel, Japan, and Luxembourg, for example, where more than 90% of the population live in urban areas and where coverage rates have been very high in recent years (evidence from other, single-country, studies speaks of the same phenomenon).[15] Naturally, immunization activities are easier to organize in urban areas, they reach a much larger proportion of the population more efficiently, and provide for greater overall accessibility to vaccination.

A determinant deserving special mention in our discussion is that of under-5 mortality, considered indicative of the performance of healthcare services in general.[45] Its negative relationship with vaccination coverage suggests that, on average, the impact of lower vaccination coverage causing higher disease burden dominates upward pressure on vaccination coverage through an expected higher willingness to accept vaccination when the disease burden it would prevent is higher (referring to the so-called prevalence elasticity of vaccine demand).[46] That is, causality from lower vaccination coverage to increased mortality under 5 seems to dominate potential reverse causality of higher mortality leading to higher vaccination coverage. The observed effect in our analysis was relatively large, for both vaccines. When under-5 mortality was taken out of the equation in the secondary branch of our analysis, the proportion of births attended by skilled health staff came to the forefront as a statistically significant positive determinant. This indicator can be seen as a proxy to overall maternal and postnatal care organization, indicating that a lower threshold – both in access to information and logistics to facilitate actual access to vaccination – is important in order to improve vaccination coverage.[47]

With regard to immunization and healthcare system indicators other than births attended by skilled health staff, our analysis yielded more puzzling results, at least at first glance. Current health expenditure, density of physicians, density of

nurses and midwives, whether the country had a multi-year plan (MYP) or a standing technical advisory group (NITAG) for/on immunization, or if the second dose of the Measles vaccine was integrated into the routine immunization system, all showed statistically significant negative effects, mostly for MCV1 and mostly in the absence of under-5 mortality (i.e., in the secondary branch of the analysis). One possible explanation for that could be that structures like MYP and NITAG were put in place when and where vaccination coverage rates were low, and MCV2 was relatively more included in routine immunization (both as a booster dose and as a "catch-up" opportunity) in countries where, or in periods when, relatively more children missed MCV1. Furthermore, if we consider factors like health expenditure and medical personnel availability to be indicative of well-organized and functioning healthcare systems with a focus on cure rather than prevention, and we know these are more characteristic of high-income countries, we come to suspect vaccine hesitancy to be the underlying factor behind some of our observed effects. The point that demand-related issues are at the core of low coverage rates in "developed" countries (as opposed to availability issues in "developing" nations) has already been put forward in the literature.[18,48] With vaccine-preventable disease burden less visible, the subsequent undervaluation of prevention, the existing safety net of good healthcare, and the easier spread of misinformation (the most well-known and cited example of MMR vaccine scare),[49] it should not be surprising that strong hesitancy attitudes have emerged in recent decades. For instance, Larson et al.[50] suggested links between higher GDP, total health spending, births attended by skilled health staff, access to healthcare, and mean years of schooling on the one hand, and lower positive vaccine sentiments on the other. If we accept that these individual-level sentiments translate directly into lower national vaccination rates, our findings provide congruent complementary information to these observations.

Lastly, based on our results, we would like to draw attention to the concept of national culture, as a possible important player when it comes to childhood vaccination. The six Hofstede indices of national culture that we considered were consistently identified among the most important covariates during the variable-selection phase of our analysis, across the multiple RF models we ran. Power distance, or a population's tendency to accept the existing uneven distribution of power, showed a relatively strong statistically significant positive effect in both branches of the analysis for DTP3. These results may be attributed to a general acceptance of and trust in the immunization practices authorities have implemented. At the same time, indulgence, or the tendency to "allow relatively free gratification of basic and natural human drives related to enjoying life and having fun," showed a negative effect of similar size for MCV1, perhaps alluding to a desire to avoid non-pleasurable activities without immediate gratification, such as vaccination. A more curious observation though was the negative effect of long-term orientation, statistically significant for both vaccines and in both branches of the analysis. Higher long-term orientation stands for easier adaptation to changing conditions, a strong inclination to save and invest, and perseverance in achieving results. That long-term orientation would negatively affect vaccination coverage may again be pointing in the direction of vaccine

hesitancy and possible vaccine safety concerns. While we did not account for vaccine hesitancy explicitly, we did include a number of country-level attitudinal covariates relating to trust, perceived fairness and democracy, and information connectedness, for example, which concepts may be assumed linked to hesitancy to some degree. None of those indices survived the variable-selection phase of the analysis, though, while, again, all cultural factors did. To our knowledge, no other study to date has investigated national culture characteristics in the context of vaccination, unlike individual-level attitudes and hesitancy toward vaccines, that have received considerable attention. It is conceivable that different cultures are susceptible to vaccine hesitancy to a different degree, and therefore this may be a research route worth exploring in the future. Some approaches may have different effects in one cultural context, as opposed to another. Another topic of further study may be the interplay of different cultural traits and the communication and implementation strategies adopted by vaccinating authorities.

Before concluding, we need to acknowledge a few statistical and conceptual limitations to our study. Firstly, although we used only widely accepted authoritative international databases for both the determinants and the vaccination coverage data and many variables are measured in the same manner for all countries using international standards, there is always the possibility that some countries' and/or time period data are less reliable than others. It is impossible for us to determine whether this occurs, and if so, how this would influence our findings since there is no telling which and how many of the variables are potentially impacted and whether this affects some countries more than others. If data would be unreliable due to fraud, one can speculate that this occurs more for country-reported data in countries where there is more corruption, and then it could have the effect that the influence of variables expressing corruption is underestimated, especially if vaccination coverage outcome data are more subject to this type of data unreliability than the other variables we used. Secondly, with regard to the presence of missing data in our dataset, while we would have liked to estimate effects for all of the 50 covariates we initially had, multiple imputation, as the "state-of-the-art" technique to apply, was just technically impossible to implement with that many variables relative to the 61 clusters, or countries, that we had. However, the consistency we saw in the results of the different runs of the RF procedure, regardless of the different model specifications, gave reassurance that, indeed, the most important covariates were identified. Next, we were, unfortunately, unable to provide direct interpretation of the size of the effects we found, due to the necessary data transformations. Conversely, as variables were now on the same scale, we were able to compare effects and speak of stronger or weaker associations in relative terms. Mild to moderately strong correlations were present in the dataset, as can be expected with so many related variables. Dealing with longitudinal data, we did not have the option of applying a well-established

dimension-reduction technique eliminating multicollinearity (such as principal components, for example), even if the subsequent decrease in interpretability would have been deemed acceptable. We, therefore, have to emphasize that results from this study should be read with due caution. Additionally, there are inevitable limitations of using national-level vaccination coverage rates by a certain age as an outcome of interest – these are not sensitive to vaccination timeliness[51] and ignore possible sub-national and regional coverage differences. Nevertheless, we believe that the macro-level insights obtained in this study constitute a significant contribution toward a greater understanding of the main factors influencing vaccination uptake. Understanding the mechanisms and determinants of vaccine uptake remains important, given the indisputable public health impact of higher vaccination coverage, which also reverts back to some of the covariates that may explain it, like GDP per capita. What is striking in this analysis is the identification of a relatively small subset of determinants that is influential to explain between-country differences in childhood vaccine uptake. This may guide politicians and public health authorities to develop the circumstances in which these determinants can be most effective in achieving high vaccination coverage. Clearly, this task is easier to achieve for the more specific determinants than for the more contextual determinants our analysis identified as influential.

ORCID

Vladimira Varbanova (iD) http://orcid.org/0000-0002-0508-4018
Frederik Verelst (iD) http://orcid.org/0000-0001-8399-743X
Niel Hens (iD) http://orcid.org/0000-0003-1881-0637
Philippe Beutels (iD) http://orcid.org/0000-0001-5034-3595

References

1. Dube E, Laberge C, Guay M, Bramadat P, Roy R, Bettinger J. Vaccine hesitancy an overview. Hum Vacc Immunother. 2013;9 (8):1763–73. doi:10.4161/hv.24657.
2. Jarrett C, Wilson R, O'Leary M, Eckersberger E, Larson HJ, Swgv H. Strategies for addressing vaccine hesitancy - a systematic review. Vaccine. 2015;33(34):4180–90. doi:10.1016/j.vaccine.2015.04.040.
3. Organization WH. Vaccines and immunization 2022. https://www.who.int/health-topics/vaccines-and-immunization#tab=tab_1.
4. Beutels P, Scuffham PA, MacIntyre CR. Funding of drugs: do vaccines warrant a different approach? Lancet Infect Dis. 2008;8 (11):727–33. doi:10.1016/S1473-3099(08)70258-5.
5. Bloom DE. The value of vaccination. Adv Exp Med Biol. 2011;697:1–8.
6. Bloom DE, Madhavan G. Vaccines: from valuation to resource allocation. Vaccine. 2015;33:B52–B4. doi:10.1016/j.vaccine.2015.02.071.
7. Luyten J, Beutels P. The social value of vaccination programs: beyond cost-effectiveness. Health Aff. 2016;35(2):212–18. doi:10.1377/hlthaff.2015.1088.
8. Kaufman J, Ryan R, Walsh L, Horey D, Leask J, Robinson P, Hill S. Face-to-face interventions for informing or educating parents about early childhood vaccination (Review). Cochrane Database Syst Rev. 2018;(5). https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD010038.pub3/full.
9. Dube E, Leask J, Wolff B, Hickler B, Balaban V, Hosein E, Habersaat K. The WHO Tailoring Immunization Programmes (TIP) approach: review of implementation to date. Vaccine. 2018;36(11):1509–15. doi:10.1016/j.vaccine.2017.12.012.
10. Organization WH. Measles in Europe: record number of both sick and immunized 2019. https://www.euro.who.int/en/media-centre/sections/press-releases/2019/measles-in-europe-record-number-of-both-sick-and-immunized.
11. Organization WH. Ten threats to global health in 2019. 2019. https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019.
12. Schuster M, Eskola J, Duclos P, Hesitancy SWGV. Review of vaccine hesitancy: rationale, remit and methods. Vaccine. 2015;33 (34):4157–60. doi:10.1016/j.vaccine.2015.04.035.
13. Favin M, Steinglass R, Fields R, Banerjee K, Sawhney M. Why children are not vaccinated: a review of the grey literature. Int Health. 2012;4(4):229–38. doi:10.1016/j.inhe.2012.07.004.
14. Gallagher KE, Kadokura E, Eckert LO, Miyake S, Mounier-Jack S, Aldea M, Ross DA, Watson-Jones D. Factors influencing completion of multi-dose vaccine schedules in adolescents: a systematic review. BMC Public Health. 2016;16(1). doi:10.1186/s12889-016-2845-z.
15. Glatman-Freedman A, Nichols KA. The effect of social determinants on immunization programs. Hum Vacc Immunother. 2012;8 (3). 293–301.
16. Phillips DE, Dieleman JL, Lim SS, Shearer J. Determinants of effective vaccine coverage in low and middle-income countries: a systematic review and interpretive synthesis. BMC Health Serv Res. 2017;17. doi:10.1186/s12913-017-2626-0.
17. Rainey JJ, Watkins M, Ryman TK, Sandhu P, Bo A, Banerjee K. Reasons related to non-vaccination and under-vaccination of children in low and middle income countries: findings from a systematic review of the published literature, 1999-2009. Vaccine. 2011;29(46):8215–21. doi:10.1016/j.vaccine.2011.08.096.
18. Smith LE, Amlot R, Weinman J, Yiend J, Rubin GJ. A systematic review of factors affecting vaccine uptake in young children. Vaccine. 2017;35(45):6059–69. doi:10.1016/j.vaccine.2017.09.046.
19. Tauil MD, Sato APS, Waldman EA. Factors associated with incomplete or delayed vaccination across countries: a systematic review. Vaccine. 2016;34(24):2635–43. doi:10.1016/j.vaccine.2016.04.016.
20. Thomson A, Robinson K, Vallee-Tourangeau G. The 5as: a practical taxonomy for the determinants of vaccine uptake. Vaccine. 2016;34(8):1018–24. doi:10.1016/j.vaccine.2015.11.065.
21. Minkov M. Cross-cultural analysis: the science and art of comparing the world's modern societies and their cultures. 2013[accessed 2022 August 18]. Thousand Oaks CA: SAGE Publications, Inc. Available from: https://sk.sagepub.com/books/cross-cultural-analysis.
22. Minkov M, Hofstede G. Clustering of 316 European regions on measures of values: do Europe's countries have national cultures? Cross-Cultural Res. 2014;48(2):144–76. doi:10.1177/1069397113510866.
23. Serajuddin U, Hamadeh N. New World Bank country classifications by income level: 2020-2021 2020. https://blogs.worldbank.org/opendata/new-world-bank-country-classifications-income-level-2020-2021.
24. Babalola S. Determinants of the uptake of the full dose of diphtheria-pertussis-tetanus vaccines (DPT3) in Northern Nigeria: a multilevel analysis. Matern Child Health J. 2009;13 (4):550–58. doi:10.1007/s10995-008-0386-5.
25. Burton A, Monasch R, Lautenbach B, Gacic-Dobo M, Neill M, Karimov R, Wolfson L, Jones G, Birmingham M. WHO and UNICEF estimates of national infant immunization coverage: methods and processes. B World Health Organ. 2009;87:535–41.
26. Chard AN, Gacic-Dobo M, Diallo MS, Sodha SV, Wallace AS. Routine vaccination coverage - worldwide, 2019. MMWR-Morbid Mortal W. 2020;69(45):1706–10. doi:10.15585/mmwr.mm6945a7.
27. Lim SS, Stein DB, Charrow A, Murray CJL. Tracking progress towards universal childhood immunisation and the impact of global initiatives: a systematic analysis of three-dose diphtheria, tetanus, and pertussis immunisation coverage. Lancet. 2008;372(9655):2031–46. doi:10.1016/S0140-6736(08)61869-3.
28. Lopalco PL, de Carvalho HG, Kreidl P, Leitmeyer K, Giesecke J. Childhood vaccination schedules in Europe vary widely is this a problem? Bundesgesundheitsbla. 2009;52(11):1095–98. doi:10.1007/s00103-009-0960-z.
29. Muhoza P, Danovaro-Holliday MC, Diallo MS, Murphy P, Sodha SV, Requejo JH, Wallace AS. Routine vaccination coverage — worldwide, 2020. MMWR-Morbid Mortal W. 2021;70 (43):1495–500. doi:10.15585/mmwr.mm7043a1.
30. Organization WH. WHO and UNICEF warn of a decline in vaccinations during COVID-19 2020. https://www.who.int/news/item/15-07-2020-who-and-unicef-warn-of-a-decline-in-vaccinations-during-covid-19.
31. (UNICEF) WHOWaUNCsF. WHO/UNICE Estimates of National Immunization Coverage (WUENIC). 2021.
32. (UNICEF) WHOWaUNCsF. Immunization data. 2021.
33. Bank TW. World development indicators. 2021.
34. Bank TW. The Worldwide Governance Indicators (WGI) project. 2021.
35. United Nations Educational SaCOUIfSU. Education. 2021.
36. Association TWVS. World values survey. 2019.
37. Hofstede G. Culture's consequences: comparing values, behaviors, institutions and organizations across nations. 2nded. Thousand Oaks CA: SAGE Publications; 2001.
38. Insights H. The 6-D model of national culture. 2015.
39. Breiman L. Random forests. Mach Learn. 2001;45(1):5–32. doi:10.1023/A:1010933404324.
40. van Buuren S-OK. Mice: multivariate imputation by chained equations in R. J Stat Softw. 2011;45(3). doi:10.18637/jss.v045.i01.
41. Liang KY, Zeger SL. Longitudinal data-analysis using generalized linear-models. Biometrika. 1986;73(1):13–22. doi:10.1093/biomet/73.1.13.
42. Varbanova V, Beutels P. Recent quantitative research on determinants of health in high income countries: a scoping review. PLoS One. 2020;15(9):e0239031. doi:10.1371/journal.pone.0239031.
43. Akmatov MK, Kretzschmar M, Kramer A, Mikolajczyk RT. Timeliness of vaccination and its effects on fraction of vaccinated

population. Vaccine. 2008;26(31):3805–11. doi:10.1016/j.vaccine.2008.05.031.

44. Feldstein LR, Mariat S, Gacic-Dobo M, Diallo MS, Conklin LM, Wallace AS. Global routine vaccination coverage, 2016. MMWR-Morbid Mortal W. 2017;66(45):1252–55. doi:10.15585/mmwr.mm6645a3.

45. Joshi M. Comprehensive peace agreement implementation and reduction in neonatal, infant and under-5 mortality rates in post-armed conflict states, 1989-2012. BMC Int Health Hum Rights. 2015;15:27. doi:10.1186/s12914-015-0066-7.

46. Philipson T. Private vaccination and public health: an empirical examination for US measles. J Hum Resour. 1996;31(3):611–30. doi:10.2307/146268.

47. Simone B, Carrillo-Santisteve P, Lopalco PL. Healthcare workers' role in keeping MMR vaccination uptake high in Europe: a review of evidence. Eurosurveillance. 2012;17(26):7–15. doi:10.2807/ese.17.26.20206-en.

48. Wariri O, Nkereuwem E, Erondu NA, Edem B, Nkereuwem OO, Idoko OT, Agogo E, Enegela JE, Sesay T, Conde IS, et al. A scorecard of progress towards measles elimination in 15 west African countries, 2001–19: a retrospective, multicountry analysis of national immunisation coverage and surveillance data. Lancet Glob Health. 2021;9(3):E280–E90. doi:10.1016/S2214-109X(20)30481-2.

49. Godlee F, Smith J, Marcovitch H. Wakefield's article linking MMR vaccine and autism was fraudulent. Br Med J. 2011:342. https://www.bmj.com/content/342/bmj.c7452).

50. Larson HJ, de Figueiredo A, Xiahong Z, Schulz WS, Verger P, Johnston IG, Cook AR, Jones NS. The state of vaccine confidence 2016: global Insights through a 67-country survey. Ebiomedicine. 2016;12:295–301. doi:10.1016/j.ebiom.2016.08.042.

51. Bailie RS, Si DM, Dowden MC, Selvey CE, Kennedy C, Cox R, O'Donoghue L, Liddle H, Connors CM, Thompson S, et al. A systems approach to improving timeliness of immunisation. Vaccine. 2009;27(27):3669–74. doi:10.1016/j.vaccine.2009.02.068.

52. Bundesamt S. 2017. Population in the former territories of the Federal Republic of Germany and the German Democratic Republic from 1950 to 2016.

53. Breiman LC. 2003. Manual – setting up, using, and understanding random forests V4.0.

54. Rubin DB. Multiple imputation for nonresponse in surveys. New York: Wiley; 1987.

55. Pepe MS, Anderson GL. A cautionary note on inference for marginal regression-models with longitudinal data and general correlated response data. Commun Stat Simulat. 1994;23:939–51.
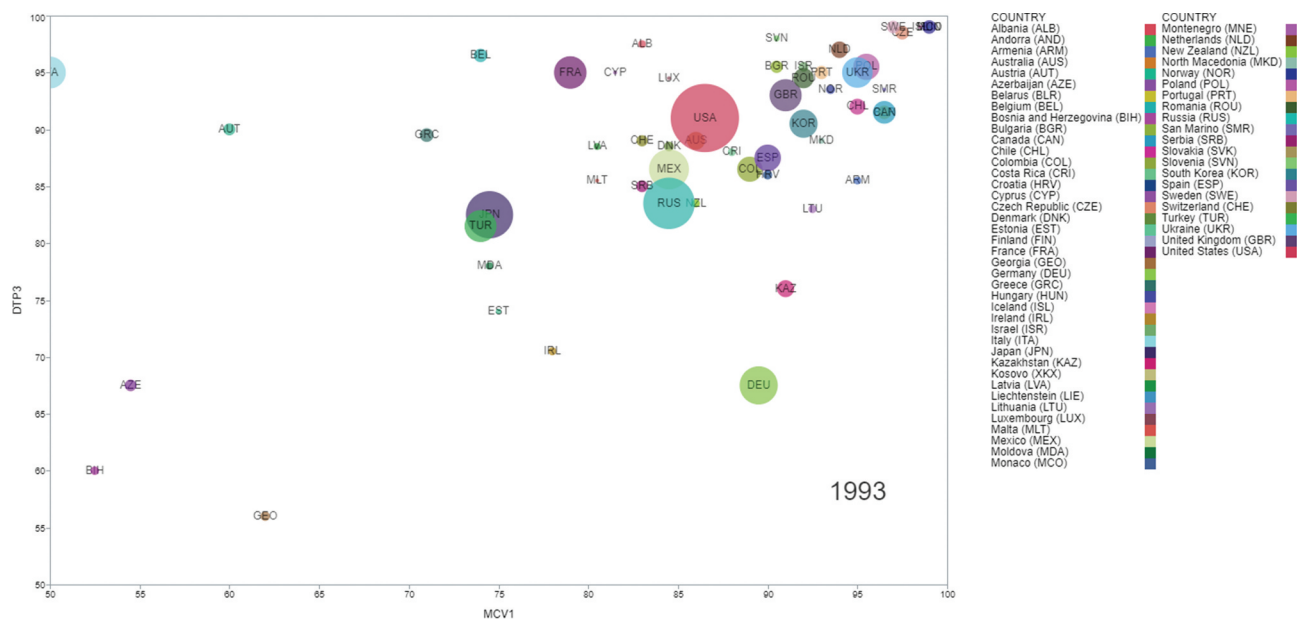
Appendix A

Table A1. Ranking of selected countries according to gross national income (GNI) per capita (in current USD) for 2019 (data not available for Andorra, Liechtenstein, Monaco, San Marino). The World Bank's classification by income level from the 1st of July 2020 shown is based on GNI per capita for 2019*.

Ranking	Country	GNI
	high income(>$12,535)	
1	Switzerland	$84,260
2	Norway	$81,640
3	Luxembourg	$77,040
4	Iceland	$72,900
5	USA	$65,970
6	Denmark	$63,460
7	Ireland	$63,230
8	Sweden	$56,410
9	Australia	$54,910
10	Netherlands	$53,180
11	Austria	$50,960
12	Finland	$49,940
13	Germany	$49,140
14	Belgium	$48,010
15	Canada	$46,550
16	Israel	$43,540
17	UK	$43,460
18	New Zealand	$42,870
19	France	$42,510
20	Japan	$42,330
21	Italy	$34,870
22	South Korea	$33,860
23	Spain	$30,350
24	Cyprus	$28,600
25	Malta	$28,340
26	Slovenia	$25,950
27	Estonia	$23,250
28	Portugal	$23,170
29	Czech Republic	$22,110
30	Greece	$19,650
31	Slovakia	$19,230
32	Lithuania	$19,050
33	Latvia	$17,790
34	Hungary	$16,520
35	Poland	$15,360
36	Croatia	$15,320
37	Chile	$14,990
38	Romania	$12,620
	upper-middle income ($4,046 -$12,535)	
39	Costa Rica	$12,090
40	Russia	$11,250
41	Turkey	$9,690
42	Bulgaria	$9,500
43	Mexico	$9,470
44	Montenegro	$9,130
45	Kazakhstan	$8,820
46	Serbia	$7,040
47	Colombia	$6,570
48	Belarus	$6,370
49	Bosnia and Herzegovina	$6,180
50	North Macedonia	$5,890
51	Albania	$5,230
52	Georgia	$4,690
53	Armenia	$4,660
54	Kosovo	$4,640
55	Moldova	$4,580
56	Azerbaijan	$4,510
	lower-middle income ($1,036 -$4045)	
57	Ukraine	$3,310

*https://blogs.worldbank.org/opendata/new-world-bank-country-classifications-income-level-2020–2021.

Appendix B

Appendix C. Statistical analysis

- **Data preparation**

Based on data availability for different countries and time periods over the selection of data sources, we considered data for 61 European or Organisation for Economic Co-operation and Development (OECD) member countries for the years between 1990 and 2019. The main branch of our analysis included two outcome variables (DTP3 and MCV1 coverage rates) and 50 covariates in total, 48 for DTP3 and 49 for MCV1, as three covariates were only applicable to one vaccination rate or the other. As an additional, secondary, branch of the analysis, we also ran the whole procedure excluding life-expectancy at birth and under-5 mortality, because of a possible strong bi-directional causal relationship with the outcomes (thus having 46 remaining covariates for DTP3 and 47 for MCV1).

With the exception of the six Hofstede cultural indices that are considered time-invariant and thus consist of just one value per country (or region), all other data came in the form of annual entries, which gave our dataset a two-level structure. We considered only national values for the cultural indices and ignored any sub-country divisions. In the case of Germany, the World Values Survey (WVS) reported separate values for East and West Germany for all waves of the survey. In order to obtain a single national value, we used a weighted average based on the year-specific population size.[52] For this set of 12 WVS covariates, we disregarded the original wave-based structure of the obtained data and instead considered the year in which data for a particular country was collected, for the purposes of data integration across the different databases we utilized. This, in addition to the fact that some of the countries of interest had only participated in one or two waves of the survey, or never at all, resulted in pronounced data sparsity for the WVS indicators, which was treated as other missing data (see more below).

- **Random forests (RF)**

As a first step in both branches of the analysis – with and without life-expectancy and under-5 mortality, we employed the random forests statistical procedure as a variable-selection technique, using SAS. Random forests[39] is a well-known and widely used supervised machine-learning technique primarily used for predictive modeling. It builds an ensemble of decision trees introducing randomness at two different points in the process of developing each individual tree. On the one hand, tree training is done using only a fraction of the original data (in-bag fraction) selected via bootstrap sampling (without replacement, in our case), while the rest of the data is set aside for model validation. On the other hand, the split-variable for each split on each tree is selected from a random sample of the original covariates, thus reducing correlations between the trees. The procedure has great efficiency and is very robust. It is also well suited to handle missing data. In addition to all covariates, we included a country indicator in order to account for the clustering in the data. We ran multiple models for both outcomes in both branches of the analysis, varying three parameters in the model – the number of trees to grow (100 or 200), the number of variables to choose from for splitting (6 values at approximately equal intervals between 2 and the total number of covariates), and the in-bag fraction (0.4, 0.6, or 0.8). In order to determine the most important covariates of DTP3 and MCV1 vaccination coverage rates, we looked at the variable importance indicator calculated based on the loss reduction,[53] and took a count of the number of times each covariate appeared among the "top 20" variable importance scores. Top-scoring covariates were then retained for the next phase of the analysis.

- **Multiple imputation (MI)**

Initially, the dataset for the main branch of the analysis showed 51% overall missingness, while the dataset for the secondary branch showed 53% missingness. After variable-selection, these numbers were reduced to 19% and 24%, respectively. In order to address this remaining issue, we performed multilevel multiple imputation (in R) with the reduced dataset obtained in the previous analytical step. Multiple imputation[54] is the current "state-of-the-art" approach to dealing with missingness, as it accounts for the inherent uncertainty when imputing, which is ignored by single imputation techniques. Under the missing-at-random (MAR) assumption, we implemented the multivariate imputation by chained equations (MICE) algorithm,[40] also known as fully conditional specification (FCS). MICE does not assume a multivariate distribution for the data, but instead uses a set of conditional densities. Imputation is done on a variable-by-variable basis, iterating over a conditionally specified imputation model for each incomplete variable. We used the predictive mean matching (PMM) method, which entails first calculating predictions for each entry of the target variable, whether observed or missing in

reality. For each missing entry, then, a small number of "candidate donors" among the observed cases are selected, based on proximity between the predictions. At the end, one donor from the group is randomly selected, and its observed value is imputed for the missing entry. Some of the advantages of this method are as follows: it ensures that the imputed value is always within the plausible range as it is based on observed data; it has the ability to handle all types of variables (our dataset contained both continuous and categorical ones); and it is robust to transformations. As our outcome variables were proportional, the data ranged between 0 and 1. Roughly 75% of all vaccination coverage rates, in fact, lie between 0.90 and 0.99, which necessitated applying the variance-stabilizing arcsine transformation. In addition, we standardized all variables to have a mean of 0 and a standard deviation of 1. We implemented the simplest proper imputation model, where for the imputation of a target variable we used only variables that were going to be included together with it in the same substantive model (in the next step). In order to minimize convergence issues, we ran the MICE procedure multiple times (with different seeds) generating a smaller number of imputations at a time, for a total of 20 imputed datasets – a number considered sufficient with 18–19 covariates included in the model.

- **Generalized estimating equations (GEE)**

In the final stage of the analysis, we applied (again in SAS) the generalized estimating equations technique[41] to each imputed dataset obtained in the previous step individually, considering the covariates that were previously selected as most important via the random forests procedure (see above). GEE is a well-known semi-parametric approach to analyzing longitudinal data. It belongs to the marginal models family, which treats the within-subject (in our case, within-country) covariance structure as a nuisance and focuses on the mean, or population-averaged, response. The procedure has the advantages of being relatively simple computationally and avoiding distributional assumptions. It iteratively calculates re-weighted least squares using a pre-specified working correlation matrix for the weights. In the presence of time-dependent covariates, our only valid choice for this working correlation matrix was that of independence among the responses.[55] The normal distribution with the identity link function was used for our standardized arcsine-transformed vaccination coverage rates. All inferences were based on the robust, "sandwich estimator," standard errors. Lastly, for each of the four scenarios separately – DTP3 and MCV1 in the main branch, and DTP3 and MCV1 in the branch without life-expectancy and under-5 mortality, results from the 20 individual GEE models were pooled together according to Rubin's.[54]