**Dataset User Guide: COVID-19 mortality among migrant health care workers**

**Grant reference: ES/W013193/1**

**Sarah Tipping and Nicola Yeates**

This work was supported by the UK Research and Innovation Fund [grant reference ES/W013193/1]. It does not necessarily reflect the opinions of the UKRI.

For more details on the methodological procedures and decisions pertaining to this dataset please see:

Sarah Tipping, Vicky Murphy, Nicola Yeates, Carlos Montoro, Gihan Ismail and Nashwa Ismail (2022) *Migrant health worker deaths during Covid-19: a methodological exploration and initial estimates.* 10th September. Milton Keynes, UK/Ferney Voltaire, France: The Open University/Public Services International. DOI: <https://doi.org/10.21954/ou.ro.00014b4b>

CC BY-NC-SA: This license allows re-users to distribute, remix, adapt, and build upon the material in any medium or format for non-commercial purposes only, and only so long as attribution is given to the creators. If you remix, adapt, or build upon the material, you must license the modified material under identical terms.

**Overview**

This dataset is the result of a review into the available data on Covid-19 and migrant health care workers. As part of the review, a methodology was developed to estimate how many migrant health workers died due to Covid-19 at a country-level. The information needed for estimates was drawn from a number of sources and compiled to create this dataset. It includes information on the number of foreign-born health workers, the number of Covid-19 deaths, and the overall population size.

Not all pieces of information were available for all countries. For some countries, data were almost entirely missing, while for others there were gaps in the information available, or it was only available through a mix of official and non-official sources. The dataset includes figures from international datasets only: country-specific figures were only sourced for the four trial countries included in the review, hence some gaps will remain where there were gaps in the original worldwide sources. A list of the information included in the dataset is given in Table 1 below.

The following sections outline the data sources used to create the dataset, including details on imputing missing information on the proportion of foreign-born health workers.

**Data on Covid-19 deaths**

The dataset includes data on Covid-19 deaths officially reported to the WHO by different countries and estimates of excess deaths produced by various academics and research organisations.

*Officially reported Covid-19 deaths*

The official Covid-19 mortality figures for each country were obtained from the WHO’s Covid-19 dashboard[[1]](#footnote-1). The dataset includes total counts recorded up until 31 December 2021.

However, the dataset does not include officially-reported Covid-19 deaths by sex and age. This information was not available from the primary WHO Covid-19 dashboard. A second source, the WHO Covid-19 Detailed Surveillance Data Dashboard[[2]](#footnote-2) was considered for the four trial countries included in the review. This dashboard combines data from case report forms, daily and weekly counts reported to the WHO, Our World In Data (OWID), the Foundation for Innovative New Diagnostics (FIND), and official public websites to produce a dataset that includes various breakdowns of the data, including age and sex.

Despite its increased detail compared to the WHO Covid-19 dashboard, the data contained some issues around data quality; for example, some records were incomplete and there were internal inconsistencies (such as the sum of male and female death counts not equalling the total for all sexes). As such, this dataset was considered for the four trial countries and used to estimate sex and age proportions of Covid-19-related deaths for two, Mexico and the UK, but has not been included in the dataset[[3]](#footnote-3). Future researchers are required to source this information independently.

*Excess death estimates*

The dataset includes three measures of excess mortality due to Covid-19 in the period up to 31 December 2021. 'Excess mortality' is defined as the difference between the total number of deaths from all causes and the number of deaths that would have been expected in the absence of the pandemic. Several groups have attempted to estimate the ‘true’ death count of Covid-19 using this approach. This task is not trivial and it should be noted that each of these estimates comes with a great deal of uncertainty, given the large amount of missing data and the quality issues around some of the available data. These issues include the low registration of deaths in some countries due to poor civil registration infrastructures and weak administrative capacity, and delays in reporting.

The estimates produced by the research team used the excess death figures produced by the WHO. The WHO had been tracking global excess mortality as the pandemic had evolved and published a set of figures showing excess mortality by country, and globally, for the period 1 January 2020 to 31 December 2021[[4]](#footnote-4). Unlike the other figures on excess deaths, the WHO released a set of preliminary figures broken down by age and sex. The research team built models using data from countries with national monthly data to predict all-cause mortality in countries where these data were unavailable[[5]](#footnote-5). Additional modelling was used in countries where only sub-national data were available to model regional values, based on the assumptions that the distribution of deaths across regions remained constant over time. The dataset includes total deaths for the period up to 31 December 2021 and breakdowns by age and sex.

It should be noted that the age/sex breakdowns in the dataset include some minus values. These correspond to age groups where the excess deaths were lower than expected. The impact of lockdowns and Covid-19 restrictions reduced the death rates from other causes for groups that were less likely to suffer high rates of Covid-19 mortality. In the estimation spreadsheets, the total number of excess deaths for these groups is set to zero in order to redistribute the total number of excess deaths by age and therefore calculate age-specific crude mortality rates.

**Information on health workers**

There is no single, global source of complete information on foreign-born health workers. Different data sources have been drawn on to provide estimates of foreign-born health workers.

A key data source was the National Healthcare Workforce Accounts (NHWA)[[6]](#footnote-6). This is an extensive source of information on health workers that contains information on the number of medical doctors, nursing professionals, midwifery professionals, dentists and pharmacists per country and includes breakdowns by age and by sex (but not by age within sex). The NHWA also includes the proportion of foreign-born and foreign-trained doctors, nurses and midwives. However, this information is not included for dentists and pharmacists.

Whilst the NHWA is a useful and valuable data source, the data are not complete for all countries. The proportion of countries with missing data increases when looking at breakdowns by age, and it rises further when looking at the proportion of foreign-born and foreign-trained health workers. Furthermore, the recency of the information available varies, with the most up-to-date information for some countries being from more than ten years ago. Different groups may also be updated at different times; hence for the same country, the date of the most recent nurse figures and the date of the most recent doctor figures may differ. The dataset includes the latest information for each country and the year the data was entered.

The NHWA also does not contain information on the age and sex distribution of foreign-born health workers, only for health workers overall. The following age bands are used by the NHWA data tables: less than 25 years, 25-34 years, 35-44 years, 45-54 years, 55-64 years, 65 years and over. When creating corresponding age breaks for other data, such as population figures and Covid-19 deaths, it was necessary to establish upper and lower age cut-offs. The middle age bands (25-34 years and 35-44 years) were also collapsed into a single, wider age band to reflect the fact that other data sources did not provide figures at the same level of granularity. The following age bands were therefore used in the dataset: 15-25 years, 25-44 years, 45-54 years, 55-64 years, 65-69 years. The figures on the NHWA were also not broken down by age within sex, hence an assumption was made in the estimation that the age distribution for each sex matched that of the overall distribution. The dataset contains the original age and sex proportions only.

The dataset also contains data from the International Labour Organization (ILO). This was used to supplement the NHWA data. The ILO holds a set of modelled estimates and actual figures on the number of workers in the human health and social work sector (ISIC Q[[7]](#footnote-7)). These show the total number of persons of working age whose main activity, either in paid work or self-employment, was based in the human health care and social work sector, thereby covering a wider range of workers than the NHWA. The modelled figures are available for most countries split by sex. The actual figures include more missing data but are available by age and sex.

**Imputing missing data on the proportion of foreign-born health workers**

The proportion of foreign-born doctors, nursing professionals, and midwives in the NHWA was missing for a number of countries. This piece of information was a key input when estimating the number of Covid-19 deaths amongst foreign-born health workers, hence an attempt was made to impute the missing data.

Three measures were imputed: the proportion of foreign-born doctors, the proportion of foreign-born nurses, and the proportion of foreign-born midwives, with separate models (or sets of models) used to impute each measure. The models summarise the relationships between a range of country characteristics[[8]](#footnote-8) and the three measures and used to predict the proportion of foreign-born doctors/nurses/midwives in countries where the information was missing. These imputed values are included in the dataset as experimental estimates. They can be used in the absence of a reported value from the NHWA or another robust, national alternative but it should be acknowledged that they are modelled estimates and therefore subject to some uncertainty. The dataset includes a combined variable that combined the NHWA data (where available) and the imputed measure, with a flag indicating the data source. A fuller explanation of the modelling, a full list of the measures included in the model, and the model outputs is given in Appendix B of this User Guide.

There were also no reported figures from the NHWA for the UK. Instead figures on the proportion of foreign-born health workers were taken from the Office for National Statistics (ONS)[[9]](#footnote-9). These figures came from a large-scale nationally-representative sample and were therefore sufficiently robust to use in place of the modelled estimates. These figures suggest 29.0% of doctors, 17.5% nurses, and 7.5% of midwives are foreign-born. The same source also provides figures for the overall human health and social work sector, suggesting 22.3% of health workers are foreign-born. These figures are included in the dataset in lieu of the NHWA figures and flagged accordingly.

In addition to the estimated proportion of foreign-born doctors, nurses, and midwives, an overall proportion of foreign-born health workers was generated that could be used for the sector-level estimates for workers in the human health and social work sector. This was based on the combined proportion of foreign-born doctors, nurses and midwives from the NHWA. This combined proportion is calculated in the estimation spreadsheet and not included in the database.

**Population estimates**

The final input for the estimates was information on the overall population of each country. A set of figures broken down by age and sex was also required for the age-sex standardised estimates. These figures are available from the World Bank[[10]](#footnote-10) and updated annually for most countries.

**Further country data**

The remaining information in the dataset is the information used to impute the proportion of foreign-born health workers. A range of data from the World Bank, ILO and UN was used as predictor variables in the imputation. These include the country’s GDP, the proportion of GDP spent on healthcare, population size, population growth, the proportion of the population aged 15 years and under, the proportion aged 65 years and over, the country’s urban population, the country’s value of the Human Development Index, net migration, proportion of the country’s population living abroad as migrants, the proportion of working age population who were migrants, proportion of the working population working in the health care sector, and information on the number of doctors and nurses per 1,000 people.

**Appendix A Full list of variables included in the dataset**

|  |  |  |  |
| --- | --- | --- | --- |
| **Position** | **Variable name** | **Variable label** | **Source** |
| 1 | country | Country name | ISO |
| 2 | iso3code | ISO country code (3 digits) | ILO |
| 3 | iso2code | ISO country code (2 digits) | ILO |
| 4 | isonumcode | ISO country code (numeric) | ILO |
| 5 | cases\_cumulative\_who | Cumulative Covid-19 cases (WHO) | WHO |
| 6 | deaths\_reported\_who | Cumulative Covid-19 deaths at end 31/12/21 (reported to WHO) | WHO |
| 7 | deaths\_jh | Cumulative Covid-19 deaths at end 31/12/21 (John Hopkins) | John Hopkins |
| 8 | deaths\_excess\_ihme | Excess deaths from Covid-19 at end 31/12/21 (IMHE) | IMHE |
| 9 | deaths\_excess\_elife | Excess deaths from Covid-19 at end 31/12/21 (ELife) | Elife |
| 10 | death\_excess\_who\_type | Type of calculation used for WHO excess deaths | WHO |
| 11 | deaths\_excess\_who | Excess deaths from Covid-19 at end 31/12/21 (WHO) | WHO |
| 12 | deaths\_ex\_who\_flt24 | Excess deaths end 31/12/21 - Female 0 to 24 (WHO) | WHO |
| 13 | deaths\_ex\_who\_f2539 | Excess deaths end 31/12/21 - Female 25 to 39 (WHO) | WHO |
| 14 | deaths\_ex\_who\_f4049 | Excess deaths end 31/12/21 - Female 40 to 49 (WHO) | WHO |
| 15 | deaths\_ex\_who\_f5059 | Excess deaths end 31/12/21 - Female 50 to 59 (WHO) | WHO |
| 16 | deaths\_ex\_who\_f6069 | Excess deaths end 31/12/21 - Female 60 to 69 (WHO) | WHO |
| 17 | deaths\_ex\_who\_f7079 | Excess deaths end 31/12/21 - Female 70 to 79 (WHO) | WHO |
| 18 | deaths\_ex\_who\_f80pl | Excess deaths end 31/12/21 - Female 80 plus (WHO) | WHO |
| 19 | deaths\_ex\_who\_mlt24 | Excess deaths end 31/12/21 - Male 0 to 24 (WHO) | WHO |
| 20 | deaths\_ex\_who\_m2539 | Excess deaths end 31/12/21 - Male 25 to 39 (WHO) | WHO |
| 21 | deaths\_ex\_who\_m4049 | Excess deaths end 31/12/21 - Male 40 to 49 (WHO) | WHO |
| 22 | deaths\_ex\_who\_m5059 | Excess deaths end 31/12/21 - Male 50 to 59 (WHO) | WHO |
| 23 | deaths\_ex\_who\_m6069 | Excess deaths end 31/12/21 - Male 60 to 69 (WHO) | WHO |
| 24 | deaths\_ex\_who\_m7079 | Excess deaths end 31/12/21 - Male 70 to 79 (WHO) | WHO |
| 25 | deaths\_ex\_who\_m80pl | Excess deaths end 31/12/21 - Male 80 plus (WHO) | WHO |
| 26 | pop\_0\_14\_f\_2020 | Population ages 0-14 female 2020 | World Bank |
| 27 | pop\_15\_24\_f\_2020 | Population ages 15-24 female 2020 | World Bank |
| 28 | pop\_25\_34\_f\_2020 | Population ages 25-34 female 2020 | World Bank |
| 29 | pop\_35\_44\_f\_2020 | Population ages 35-44 female 2020 | World Bank |
| 30 | pop\_45\_54\_f\_2020 | Population ages 45-54 female 2020 | World Bank |
| 31 | pop\_55\_64\_f\_2020 | Population ages 55-64 female 2020 | World Bank |
| 32 | pop\_65\_74\_f\_2020 | Population ages 65-74 female 2020 | World Bank |
| 33 | pop\_75pl\_f\_2020 | Population ages 75+ female 2020 | World Bank |
| 34 | pop\_0\_14\_m\_2020 | Population ages 0-14 male 2020 | World Bank |
| 35 | pop\_15\_24\_m\_2020 | Population ages 15-24 male 2020 | World Bank |
| 36 | pop\_25\_34\_m\_2020 | Population ages 25-34 male 2020 | World Bank |
| 37 | pop\_35\_44\_m\_2020 | Population ages 35-44 male 2020 | World Bank |
| 38 | pop\_45\_54\_m\_2020 | Population ages 45-54 male 2020 | World Bank |
| 39 | pop\_55\_64\_m\_2020 | Population ages 55-64 male 2020 | World Bank |
| 40 | pop\_65\_74\_m\_2020 | Population ages 65-74 male 2020 | World Bank |
| 41 | pop\_75pl\_m\_2020 | Population ages 75+ male 2020 | World Bank |
| 42 | pop\_15\_64\_f\_2020 | Population ages 15-64 female 2020 | World Bank |
| 43 | pop\_15\_64\_m\_2020 | Population ages 15-64 male 2020 | World Bank |
| 44 | pop\_65pl\_f\_2020 | Population ages 65+ female 2020 | World Bank |
| 45 | pop\_65pl\_m\_2020 | Population ages 65+ male 2020 | World Bank |
| 46 | doc\_nhwa\_yr | Medical doctors from NHWA – year of data entry | NHWA |
| 47 | doc\_nhwa | Medical doctors from NHWA (total) | NHWA |
| 48 | mid\_nhwa\_yr | Midwifery personnel from NHWA – year of data entry | NHWA |
| 49 | mid\_nhwa | Midwifery personnel from NHWA (total) | NHWA |
| 50 | nur\_nhwa\_yr | Nursing personnel from NHWA – year of data entry | NHWA |
| 51 | nur\_nhwa | Nursing personnel from NHWA (total) | NHWA |
| 52 | doc\_nhwa\_fb\_yr | Proportion of foreign-born medical doctors from NHWA – year of data entry | NHWA |
| 53 | doc\_nhwa\_fb | Proportion of foreign-born medical doctors from NHWA (total) | NHWA |
| 54 | mid\_nhwa\_fb\_yr | Proportion of foreign-born midwifery personnel from NHWA – year of data entry | NHWA |
| 55 | mid\_nhwa\_fb | Proportion of foreign-born midwifery personnel from NHWA (total) | NHWA |
| 56 | nur\_nhwa\_fb\_yr | Proportion of foreign-born nursing personnel from NHWA – year of data entry | NHWA |
| 57 | nur\_nhwa\_fb | Proportion of foreign-born nursing personnel from NHWA (total) | NHWA |
| 58 | docpredict | Medical doctors - modelled proportion of foreign born | OU modelled estimates |
| 59 | nurpredict | Nurses - modelled proportion of foreign born | OU modelled estimates |
| 60 | midpredict | Midwives - modelled proportion of foreign born | OU modelled estimates |
| 61 | doc\_pfb | Combined actual and modelled proportion of foreign-born doctors | OU modelled estimates |
| 62 | nur\_pfb | Combined actual and modelled proportion of foreign-born midwives | OU modelled estimates |
| 63 | mid\_pfb | Combined actual and modelled proportion of foreign-born nurses | OU modelled estimates |
| 64 | doc\_fb\_flag | Flag modelled proportion of foreign-born doctors | OU modelled estimates |
| 65 | nur\_fb\_flag | Flag modelled proportion of foreign-born nurses | OU modelled estimates |
| 66 | mid\_fb\_flag | Flag modelled proportion of foreign-born midwives | OU modelled estimates |
| 67 | doc\_nhwa\_age\_yr | Age of medical doctors from NHWA – year of data entry | NHWA |
| 68 | pr\_medic\_lt25 | Proportion of doctors aged less than 25 (NHWA) | NHWA |
| 69 | pr\_medic\_25\_44 | Proportion of doctors aged 25-44 (NHWA) | NHWA |
| 70 | pr\_medic\_45\_54 | Proportion of doctors aged 45-54 (NHWA) | NHWA |
| 71 | pr\_medic\_55\_64 | Proportion of doctors aged 55-64 (NHWA) | NHWA |
| 72 | pr\_medic\_65pl | Proportion of doctors aged more than 65 (NHWA) | NHWA |
| 73 | nur\_nhwa\_age\_yr | Age of nurses from NHWA – year of data entry | NHWA |
| 74 | pr\_nursi\_lt25 | Proportion of nurses aged less than 25 (NHWA) | NHWA |
| 75 | pr\_nursi\_25\_44 | Proportion of nurses aged 25-44 (NHWA) | NHWA |
| 76 | pr\_nursi\_45\_54 | Proportion of nurses aged 45-54 (NHWA) | NHWA |
| 77 | pr\_nursi\_55\_64 | Proportion of nurses aged 55-64 (NHWA) | NHWA |
| 78 | pr\_nursi\_65pl | Proportion of nurses aged more than 65 (NHWA) | NHWA |
| 79 | mid\_nhwa\_age\_yr | Age of midwives from NHWA – year of data entry | NHWA |
| 80 | pr\_midwi\_lt25 | Proportion of midwives aged less than 25 (NHWA) | NHWA |
| 81 | pr\_midwi\_25\_44 | Proportion of midwives aged 25-44 (NHWA) | NHWA |
| 82 | pr\_midwi\_45\_54 | Proportion of midwives aged 45-54 (NHWA) | NHWA |
| 83 | pr\_midwi\_55\_64 | Proportion of midwives aged 55-64 (NHWA) | NHWA |
| 84 | pr\_midwi\_65pl | Proportion of midwives aged more than 65 (NHWA) | NHWA |
| 85 | doc\_nhwa\_sex\_yr | Sex of medical doctors from NHWA – year of data entry | NHWA |
| 86 | pr\_medic\_f | Proportion of doctors female (NHWA) | NHWA |
| 87 | pr\_medic\_m | Proportion of doctors male (NHWA) | NHWA |
| 88 | nur\_nhwa\_sex\_yr | Sex of nurses from NHWA – year of data entry | NHWA |
| 89 | pr\_nursi\_f | Proportion of nurses female (NHWA) | NHWA |
| 90 | pr\_nursi\_m | Proportion of nurses male (NHWA) | NHWA |
| 91 | mid\_nhwa\_sex\_yr | Sex of midwives from NHWA – year of data entry | NHWA |
| 92 | pr\_midwi\_f | Proportion of midwives female (NHWA) | NHWA |
| 93 | pr\_midwi\_m | Proportion of midwives male (NHWA) | NHWA |
| 94 | ilo\_hwf\_f\_mdld | ILO modelled est 2019 - HCW (ISIC Q) - Female (thousands) | ILO |
| 95 | ilo\_hwf\_m\_mdld | ILO modelled est 2019 - HCW (ISIC Q) - Male (thousands) | ILO |
| 96 | ilo\_hwf\_t\_mdld | ILO modelled est 2019 - HCW (ISIC Q) - Total (thousands) | ILO |
| 97 | ilo\_workforce\_yr | Health WorkForce actual estimates from ILO - year of data entry | ILO |
| 98 | ilo\_hwf\_tot\_f\_act | ILO actual est – HCW (ISIC Q) - tot female (thousands) | ILO |
| 99 | ilo\_hwf\_tot\_m\_act | ILO actual est – HCW (ISIC Q) - tot male (thousands) | ILO |
| 100 | ilo\_hwf\_tot\_t\_act | ILO actual est – HCW (ISIC Q) - tot total (thousands) | ILO |
| 101 | ilo\_hwf\_15\_24\_f\_act | ILO actual est – HCW (ISIC Q) - 15\_24 female (thousands) | ILO |
| 102 | ilo\_hwf\_15\_24\_m\_act | ILO actual est – HCW (ISIC Q) - 15\_24 male (thousands) | ILO |
| 103 | ilo\_hwf\_15\_24\_t\_act | ILO actual est – HCW (ISIC Q) - 15\_24 total (thousands) | ILO |
| 104 | ilo\_hwf\_25\_54\_f\_act | ILO actual est – HCW (ISIC Q) - 25\_54 female (thousands) | ILO |
| 105 | ilo\_hwf\_25\_54\_m\_act | ILO actual est – HCW (ISIC Q) - 25\_54 male (thousands) | ILO |
| 106 | ilo\_hwf\_25\_54\_t\_act | ILO actual est – HCW (ISIC Q) - 25\_54 total (thousands) | ILO |
| 107 | ilo\_hwf\_55\_64\_f\_act | ILO actual est – HCW (ISIC Q) - 55\_64 female (thousands) | ILO |
| 108 | ilo\_hwf\_55\_64\_m\_act | ILO actual est – HCW (ISIC Q) - 55\_64 male (thousands) | ILO |
| 109 | ilo\_hwf\_55\_64\_t\_act | ILO actual est – HCW (ISIC Q) - 55\_64 total (thousands) | ILO |
| 110 | ilo\_hwf\_65pl\_f\_act | ILO actual est – HCW (ISIC Q) - 65pl female (thousands) | ILO |
| 111 | ilo\_hwf\_65pl\_m\_act | ILO actual est – HCW (ISIC Q) - 65pl male (thousands) | ILO |
| 112 | ilo\_hwf\_65pl\_t\_act | ILO actual est – HCW (ISIC Q) - 65pl total (thousands) | ILO |
| 113 | uncode | UN region/country code | UN |
| 114 | unregion | UN region name | UN |
| 115 | unsubregion | UN Sub-region name | UN |
| 116 | subreg | UN sub-region (numeric) | UN |
| 117 | ims\_prop\_male\_2554 | Proportion of men aged 25-54 in the population who are migrants (UN) | UN |
| 118 | ims\_fprop\_2554 | Female migrants as a % of migrant stock aged 25-54 years (UN) | UN |
| 119 | propsamereg | Proportion of in-migrants from the same sub-region as this country (UN) | UN |
| 120 | inoutmig | Ratio of nationals living abroad to foreign migrants in this country (UN) | UN |
| 121 | un\_prpop\_abroad\_m | Proportion of the overall male population living abroad (UN) | UN |
| 122 | pr\_total\_fortrain | Proportion of doctors + nurses + midwives foreign trained (NHWA) | NHWA |
| 123 | lf\_per\_f\_2020 | Labor force female (% of total labor force) 2020 (ILO) | ILO |
| 124 | unemp\_per\_tot\_2021 | Unemployment total (% of total labor force) 2021 (ILO) | ILO |
| 125 | unemp\_per\_f\_2021 | Unemployment female (% of female labor force) 2021 (ILO) | ILO |
| 126 | unemp\_per\_m\_2021 | Unemployment male (% of male labor force) 2021 (ILO) | ILO |
| 127 | ilo\_t\_forborn | Proportion of migrant workers in overall workforce (ILO) | ILO |
| 128 | pr\_ilo\_hcwf\_f | Proportion of human health and social work sector employees who are female (ILO) | ILO |
| 129 | pr\_ilo\_emp\_hcwf | Proportion of employees working in human health and social work sector (ILO) | ILO |
| 130 | docper | Number of doctors per 1000 population (NHWA) | NHWA |
| 131 | nurper | Number of nurses per 1000 population (NHWA) | NHWA |
| 132 | hdi2019 | Human Development Index - HDI (2019) | HDI |
| 133 | che\_perc\_gdp | Current Health Expenditure (CHE) as % Gross Domestic Product (GDP) 2019 | World Bank |
| 134 | dom\_perc\_che | Domestic Health Expenditure (DOM) as % of Current Health Expenditure (CHE) 2019 | World Bank |
| 135 | cge\_perc\_gdp | General Government Expenditure (GGE) as % Gross Domestic Product (GDP) 2019 | World Bank |
| 136 | urbpop | WB - Urban population (% of total population) 2020 | World Bank |
| 137 | lifeexp | WB - Life expectancy at birth, total (years) 2019 | World Bank |
| 138 | lifeexpchange10 | WB - change in life exp - 2019/2009 | World Bank |
| 139 | lifeexpchange20 | WB - change in life exp - 2019/1999 | World Bank |
| 140 | lifeexpchange30 | WB - change in life exp - 2019/1989 | World Bank |
| 141 | agricu\_2019 | Agriculture forestry and fishing value added (% of GDP) 2019 | World Bank |
| 142 | attbir\_2019 | Births attended by skilled health staff (% of total) 2019 | World Bank |
| 143 | contra\_2019 | Contraceptive prevalence any method (% of married women 15-49) 2019 | World Bank |
| 144 | elepow\_2019 | Electric power consumption (kWh per capita) 2019 | World Bank |
| 145 | export\_2020 | Exports of goods and services (% of GDP) 2020 | World Bank |
| 146 | extdebt\_2020 | External debt stocks total (DOD current US$) 2020 | World Bank |
| 147 | fdinvst\_2019 | Foreign direct investment net (BoP current US$) 2019 | World Bank |
| 148 | fdinvstin\_2019 | Foreign direct investment net inflows (BoP current US$) 2019 | World Bank |
| 149 | forest\_2019 | Forest area (sq. km) 2019 | World Bank |
| 150 | gdp\_2019 | GDP (current US$) 2019 | World Bank |
| 151 | gdpgrowth\_2019 | GDP growth (annual %) 2019 | World Bank |
| 152 | gdpcap\_2019 | GDP per capita (current US$) 2019 | World Bank |
| 153 | gnicap\_2019 | GNI per capita Atlas method (current US$) 2019 | World Bank |
| 154 | gnicapppp\_2019 | GNI per capita PPP (current international $) 2019 | World Bank |
| 155 | gniatlas\_2019 | GNI Atlas method (current US$) 2019 | World Bank |
| 156 | gnippp\_2019 | GNI PPP (current international $) 2019 | World Bank |
| 157 | grosscap\_2019 | Gross capital formation (% of GDP) 2019 | World Bank |
| 158 | hitecexp\_2019 | High-technology exports (% of manufactured exports) 2019 | World Bank |
| 159 | immeas\_2019 | Immunization measles (% of children ages 12-23 months) 2019 | World Bank |
| 160 | import\_2020 | Imports of goods and services (% of GDP) 2020 | World Bank |
| 161 | indva\_2019 | Industry (including construction) value added (% of GDP) 2019 | World Bank |
| 162 | infgdpd\_2019 | Inflation GDP deflator (annual %) 2019 | World Bank |
| 163 | infcons\_2019 | Inflation consumer prices (annual %) 2019 | World Bank |
| 164 | lebirth\_2019 | Life expectancy at birth total (years) 2019 | World Bank |
| 165 | mcldc\_2020 | Market capitalization of listed domestic companies (% of GDP) 2020 | World Bank |
| 166 | merctrad\_2020 | Merchandise trade (% of GDP) 2020 | World Bank |
| 167 | militexp\_2020 | Military expenditure (% of GDP) 2020 | World Bank |
| 168 | mobsub\_2020 | Mobile cellular subscriptions (per 100 people) 2020 | World Bank |
| 169 | mortrate\_2020 | Mortality rate under-5 (per 1000 live births) 2020 | World Bank |
| 170 | netbarter\_2019 | Net barter terms of trade index (2000 = 100) 2019 | World Bank |
| 171 | premit\_2019 | Personal remittances paid (current US$) 2019 | World Bank |
| 172 | popgrow\_2019 | Population growth (annual %) 2019 | World Bank |
| 173 | poptotal\_2019 | Population total 2019 (World Bank) | World Bank |
| 174 | poptotal\_2020 | Population total 2020 (World Bank) | World Bank |
| 175 | povhc\_2019 | Poverty headcount ratio at national poverty lines (% of population) 2019 | World Bank |
| 176 | prevhiv\_2019 | Prevalence of HIV total (% of population ages 15-49) 2019 | World Bank |
| 177 | prevundw\_2019 | Prevalence of underweight weight for age (% of children under 5) 2019 | World Bank |
| 178 | primcomp\_2019 | Primary completion rate total (% of relevant age group) 2019 | World Bank |
| 179 | revexgr\_2019 | Revenue excluding grants (% of GDP) 2019 | World Bank |
| 180 | schenrolgend\_2019 | School enrollment primary and secondary (gross) gender parity index (GPI) 2019 | World Bank |
| 181 | schenrolsec\_2019 | School enrollment secondary (% gross) 2019 | World Bank |
| 182 | startup\_2019 | Start-up procedures to register a business (number) 2019 | World Bank |
| 183 | totdebtser\_2020 | Total debt service (% of GNI) 2020 | World Bank |
| 184 | expimp2010to2019 | Exports to imports 2010-19 (World Bank) | World Bank |
| 185 | prtotpop65pl | Proportion of population aged 65+ yrs (World Bank 2020) | World Bank |
| 186 | prtotpoplt15 | Proportion of population aged LT 15 yrs (World Bank 2020) | World Bank |

**Appendix B Imputing missing information on the proportion of foreign-born health workers in a country**

The proportion of foreign-born doctors, nursing professionals, and midwives in the NHWA was missing for a number of countries (of 192 countries in the dataset, 62% were missing any information on the proportion of foreign-born doctors, 49% on nurses, and 72% on midwives).

An investigation was made into whether this missing information could be successfully imputed using a statistical model. A statistical model will create a relationship between various inputs and an outcome of interest. For example, in this case, the outcome of interest would be the proportion of foreign-born doctors, whereas inputs that might be connected to this number could be a country’s gross domestic product (GDP) or net migration figures. As data on the proportion of foreign-born doctors were available for 48% of countries, the model was trained using these data to assess the most accurate way to predict the proportion of foreign-born doctors. This relationship between inputs and the outcome of interest was then used to make predictions for countries where the data is missing. A generalized linear model with a logit link and the binomial family was used to predict the missing data. This type of model is appropriate when the outcome variable (the characteristic being predicted) is a proportion (in this case, the proportion of foreign-born workers). It ensures the predicted values fall between zero (no foreign-born workers) and one (100% foreign-born workers), unlike other approaches that may allow values to be predicted that fall outside this range.

Three outcome variables were modelled: the proportion of foreign-born doctors; the proportion of foreign-born nurses; and the proportion of foreign-born midwives, with separate models (or sets of models) used to estimate each outcome.

A range of data from the World Bank, ILO and UN was used as predictor variables in the models. These include the country’s GDP, the proportion of GDP spend on healthcare, population size, population growth, the proportion of the population aged 15 years and under, the proportion aged 65 years and over, the country’s urban population, the country’s value of the Human Development Index, net migration, proportion of the country’s population living abroad as migrants, the proportion of working age population who were migrants, proportion of the working population working in the health care sector, and information on the number of doctors and nurses per 1,000 people.

The models summarise the relationships between the various measures and the three outcome variables. This relationship is then used to predict the outcome for countries where the information about foreign-born health workers is missing. Whilst the primary aim of this analysis was to fill in information that was missing for the four selected countries, all 194 countries in the NHWA data were included in the modelling in order to build the best models and get the best understanding of the relationships between the different country measures and the outcome variables. The result is a dataset containing, for all countries, the reported proportions of foreign-born doctors, nurses and midwives, plus imputed values from the model.

In order to incorporate as much information as possible, more than one model was used to predict the proportion of foreign-born doctors and foreign-born nurses. The proportion of foreign-born doctors was estimated using two models. One model was run on those countries where there was information on foreign-trained doctors, since this measure was strongly predictive of the proportion of foreign-born doctors but was also missing for many countries. Hence where this information was available it was included in the model. A second model was run that excluded this measure, drawing on wider but less predictive information to estimate the proportion of foreign-born doctors. This approach meant the fullest, most predictive information was used where available. A similar approach was taken for nurses, resulting in two models for the estimate of the proportion of foreign-born nurses. The high degree proportion of missing data for midwives meant a single model was used here, there was no benefit to running additional models. The estimated and actual figures were checked using the root mean square error (RMSE) to ensure predicted values were within an acceptable range. The RMSE checks the difference between the predicted value of the model and the actual value for countries where that figure is available. The closer the RMSE is to zero, the closer the model’s predicted values are to the real figures of countries where that data is available. Given the large number of entry variables, the model outputs were checked for issues with collinearity and inflated covariates.

**Table B.1 Output from modelling the proportion of foreign-born doctors**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Model 1 | | | | Model 2 | | | |
|  | Coef. | Std Err. | Z-value | P-value | Coef. | Std Err. | Z-value | P-value |
|  |  |  |  |  |  |  |  |  |
| Proportion of migrant workers in overall workforce (ILO) | -3.71 | 0.46 | -8.16 | 0.00 | 3.50 | 1.24 | 2.83 | 0.01 |
| Proportion of Nurses+Docs+MW Foreign Trained | 5.42 | 0.53 | 10.16 | 0.00 |  |  |  |  |
| Proportion of men aged 25-54 in the population who are migrants | 0.11 | 0.01 | 8.98 | 0.00 | 0.08 | 0.02 | 4.50 | 0.00 |
| Proportion of the population aged less than 15 years old | 3.35 | 1.71 | 1.97 | 0.05 |  |  |  |  |
| Proportion of human health and social work sector employees who are female | 2.57 | 0.58 | 4.40 | 0.00 | -1.95 | 1.18 | -1.65 | 0.10 |
| Number of doctors per 1000 population | -0.09 | 0.06 | -1.55 | 0.12 |  |  |  |  |
| GNI per capita Atlas method (current US$) 2019 | 0.00 | 0.00 | -1.89 | 0.06 |  |  |  |  |
| Population growth (annual %) 2019 |  |  |  |  | -0.45 | 0.22 | -2.04 | 0.04 |
| Population size (log) |  |  |  |  | 0.35 | 0.15 | 2.25 | 0.02 |
| Urban population as a proportion of the total population 2020 |  |  |  |  | -0.03 | 0.01 | -2.17 | 0.03 |
| High Income Country |  |  |  |  | 5.44 | 1.73 | 3.15 | 0.00 |
| Upper Middle Income Country |  |  |  |  | 4.13 | 1.39 | 2.97 | 0.00 |
| Lower Middle Income Country |  |  |  |  | 1.48 | 0.87 | 1.71 | 0.09 |
| Country density of forest (log) |  |  |  |  | -0.38 | 0.11 | -3.52 | 0.00 |
| Immunization measles (% of children ages 12-23 months) 2019 |  |  |  |  | 0.04 | 0.02 | 1.86 | 0.06 |
| Ratio of nationals living abroad to number of foreign migrants in this couuntry |  |  |  |  | 0.22 | 0.04 | 4.99 | 0.00 |
| Female migrants as a % of migrant stock aged 25-54 years |  |  |  |  | -0.04 | 0.03 | -1.32 | 0.19 |
| Region: Europe+North America |  |  |  |  | -2.65 | 0.67 | -3.98 | 0.00 |
| Region: Latin America+Carib+Oceania |  |  |  |  | -0.36 | 0.45 | -0.81 | 0.42 |
| Region: Sub-Saharan Africa |  |  |  |  | 3.94 | 0.87 | 4.53 | 0.00 |
| Region: North Africa+Middle East |  |  |  |  | -1.79 | 0.51 | -3.53 | 0.00 |
| Region: Asia (baseline) |  |  |  |  | (baseline) |  |  |  |
| Missing ILO information on foreign born workers |  |  |  |  | 0.57 | 0.38 | 1.48 | 0.14 |
| Current Health Expenditure (CHE) as % Gross Domestic Product (GDP) 2019 |  |  |  |  | 0.21 | 0.04 | 4.96 | 0.00 |
| Agriculture forestry and fishing value added (% of GDP) 2019 |  |  |  |  | 0.13 | 0.06 | 2.35 | 0.02 |
| HDI2019 |  |  |  |  | 7.08 | 3.61 | 1.96 | 0.05 |
| Labor force female (% of total labor force) 2020 |  |  |  |  | 0.09 | 0.04 | 2.32 | 0.02 |
| UN Proportions of in-migrants from the same sub-region |  |  |  |  | 1.77 | 0.66 | 2.70 | 0.01 |
|  |  |  |  |  |  |  |  |  |
| Constant | -5.53 | 0.80 | -6.90 | 0.00 | -19.20 | 5.25 | -3.66 | 0.00 |
|  |  |  |  |  |  |  |  |  |
| AIC | 34.6 |  |  |  | 77.9 |  |  |  |
| BIC | 46.6 |  |  |  | 125.6 |  |  |  |
| P-value overall | 0 |  |  |  | 0 |  |  |  |
| Number of cases | 33 |  |  |  | 54 |  |  |  |

**Table B.2 Output from modelling the proportion of foreign-born nurses**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Model 1 | | | | Model 2 | | | |
|  | Coef. | Std Err. | Z-value | P-value | Coef. | Std Err. | Z-value | P-value |
|  |  |  |  |  |  |  |  |  |
| Proportion of migrant workers in overall workforce (ILO) | -7.09 | 0.61 | -11.68 | 0.00 |  |  |  |  |
| Proportion of Nurses+Docs+MW Foreign Trained | 3.05 | 0.45 | 6.81 | 0.00 |  |  |  |  |
| Proportion of men aged 25-54 in the population who are migrants | 0.13 | 0.01 | 10.77 | 0.00 |  |  |  |  |
| Proportion of the population aged less than 15 years old | -4.88 | 2.37 | -2.06 | 0.04 |  |  |  |  |
| Proportion of human health and social work sector employees who are female | 5.50 | 0.59 | 9.26 | 0.00 |  |  |  |  |
| Number of doctors per 1000 population | 0.15 | 0.04 | 3.42 | 0.00 |  |  |  |  |
| Population growth (annual %) 2019 | 0.07 | 0.08 | 0.90 | 0.37 | -1.22 | 0.35 | -3.46 | 0.00 |
| Population size (log) | 0.44 | 0.03 | 13.18 | 0.00 | 0.41 | 0.15 | 2.70 | 0.01 |
| Urban population as a proportion of the total population 2020 | -0.07 | 0.01 | -14.50 | 0.00 | -0.03 | 0.01 | -2.88 | 0.00 |
| High Income Country | 1.77 | 0.95 | 1.85 | 0.06 | 4.18 | 1.68 | 2.49 | 0.01 |
| Upper Middle Income Country | 0.00 | 0.67 | 0.00 | 1.00 | 4.01 | 1.51 | 2.65 | 0.01 |
| Lower Middle Income Country | -1.19 | 0.45 | -2.67 | 0.01 | 0.75 | 1.10 | 0.69 | 0.49 |
| Low Income Country (baseline) | (baseline) |  |  |  | (baseline) |  |  |  |
| Country density of forest (log) | -0.43 | 0.02 | -20.75 | 0.00 | -0.20 | 0.10 | -1.99 | 0.05 |
| Life expectancy at birth (in years) 2019 | -0.27 | 0.03 | -8.68 | 0.00 |  |  |  |  |
| GNI per capita PPP (current international $) 2019 | 0.00 | 0.00 | -2.98 | 0.00 | 0.00 | 0.00 | -2.10 | 0.04 |
| Proportion of employees working in human health and social work sector (ILO) | 19.38 | 1.94 | 9.98 | 0.00 | 26.72 | 9.72 | 2.75 | 0.01 |
| Immunization measles (% of children ages 12-23 months) 2019 | -0.02 | 0.01 | -4.18 | 0.00 |  |  |  |  |
| Ratio of nationals living abroad to number of foreign migrants in this couuntry | -0.11 | 0.01 | -8.22 | 0.00 |  |  |  |  |
| Female migrants as a % of migrant stock aged 25-54 years | 0.06 | 0.01 | 5.55 | 0.00 | -0.06 | 0.02 | -2.57 | 0.01 |
| Unemployment total (% of total labor force) 2021 | 0.12 | 0.01 | 8.91 | 0.00 | -0.13 | 0.05 | -2.68 | 0.01 |
| Region: Europe+North America | -1.49 | 0.38 | -3.91 | 0.00 | -3.53 | 0.95 | -3.72 | 0.00 |
| Region: Latin America+Carib+Oceania | -0.26 | 0.38 | -0.70 | 0.49 | -1.69 | 0.65 | -2.59 | 0.01 |
| Region: Sub-Saharan Africa | -4.58 | 0.55 | -8.32 | 0.00 | 2.63 | 1.47 | 1.79 | 0.07 |
| Region: North Africa+Middle East | 0.87 | 0.48 | 1.83 | 0.07 | -0.74 | 0.63 | -1.18 | 0.24 |
| Region: Asia (baseline) | (baseline) |  |  |  | (baseline) |  |  |  |
| Agriculture forestry and fishing value added (% of GDP) 2019 |  |  |  |  | 0.13 | 0.02 | 7.89 | 0.00 |
| HDI2019 |  |  |  |  | 14.02 | 5.24 | 2.68 | 0.01 |
| Labor force female (% of total labor force) 2020 |  |  |  |  | 0.08 | 0.03 | 2.38 | 0.02 |
| Proportion of migrants in population (%) Female\_2554 |  |  |  |  | 0.12 | 0.02 | 5.35 | 0.00 |
| Proportion of the population aged over 65 years old |  |  |  |  | -11.58 | 6.59 | -1.76 | 0.08 |
| ExpImp2010to2019 |  |  |  |  | 0.02 | 0.01 | 2.65 | 0.01 |
| Number of nurses per 1000 population |  |  |  |  | -0.17 | 0.10 | -1.70 | 0.09 |
|  |  |  |  |  |  |  |  |  |
| Constant | 13.47 | 2.21 | 6.10 | 0.00 | -17.86 | 3.28 | -5.44 | 0.00 |
|  |  |  |  |  |  |  |  |  |
| AIC | 63.27 |  |  |  | 70.56 |  |  |  |
| BIC | 104.21 |  |  |  | 119.11 |  |  |  |
| P-value overall | 0 |  |  |  | 0 |  |  |  |
| Number of cases | 38 |  |  |  | 61 |  |  |  |

**Table B.3 Output from modelling the proportion of foreign-born midwives**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Coef. | Std Err. | Z-value | P-value |
|  |  |  |  |  |
| Proportion of men aged 25-54 in the population who are migrants | -0.06 | 0.03 | -2.03 | 0.04 |
| Proportion of human health and social work sector employees who are female | -8.55 | 2.12 | -4.04 | 0.00 |
| Population size (log) | 0.19 | 0.07 | 2.74 | 0.01 |
| High Income Country | -5.23 | 1.71 | -3.06 | 0.00 |
| Upper Middle Income Country | -7.61 | 2.02 | -3.77 | 0.00 |
| Lower Middle Income Country | -9.96 | 2.63 | -3.78 | 0.00 |
| Life expectancy at birth (in years) 2019 | 0.23 | 0.08 | 3.08 | 0.00 |
| GNI per capita PPP (current international $) 2019 | 0.00 | 0.00 | 4.84 | 0.00 |
| Female migrants as a % of migrant stock aged 25-54 years | -0.12 | 0.03 | -3.79 | 0.00 |
| Unemployment total (% of total labor force) 2021 | 0.09 | 0.04 | 2.01 | 0.04 |
|  |  |  |  |  |
| Constant | -9.57 | 4.11 | -2.33 | 0.02 |
|  |  |  |  |  |
| AIC | 33.404 |  |  |  |
| BIC | 48.058 |  |  |  |
| P-value overall | 0 |  |  |  |
| Number of cases | 28 |  |  |  |

Notes: The coefficient (coef.) summarises the relationship between the proportion of foreign-born health workers in a country and the country characteristics. It shows the expected increase or decrease in the proportion of that country’s foreign-born health workers associated with a unit change in the characteristic. The value of this coefficient is tested in the model using a z-test. The resulting p-value indicates whether the characteristic is significantly related to changes in the proportion of foreign-born health workers. The difference is deemed to be significant at the 95% level if the p-value is smaller than 0.05. P-values are probabilities, a small p-value indicates that there is a very small probability that the differences we are testing occurred purely by chance.

1. WHO reported deaths from <https://covid19.who.int/data> . Downloaded on 19 May 2022 [↑](#footnote-ref-1)
2. [https://app.powerbi.com/view?r=eyJrIjoiYWRiZWVkNWUtNmM0Ni00MDAwLTljYWMtN2EwNTM3YjQzYmRmIiwidCI6ImY2MTBjMGI3LWJkMjQtNGIzOS04MTBiLTNkYzI4MGFmYjU5MCIsImMiOjh9](%09https://app.powerbi.com/view?r=eyJrIjoiYWRiZWVkNWUtNmM0Ni00MDAwLTljYWMtN2EwNTM3YjQzYmRmIiwidCI6ImY2MTBjMGI3LWJkMjQtNGIzOS04MTBiLTNkYzI4MGFmYjU5MCIsImMiOjh9) (accessed 31 July 2022) [↑](#footnote-ref-2)
3. For the remaining trial countries, country-specific sources were used to fill the information gaps. These resources ranged from official sources to data collected by academics and journalists. These figures are not included in the dataset. A similar exercise would be needed to supplement the data in countries not included in this trial. [↑](#footnote-ref-3)
4. <https://www.who.int/data/stories/global-excess-deaths-associated-with-covid-19-january-2020-december-2021> (accessed 23 June 2022) [↑](#footnote-ref-4)
5. More details on the modelling can be found here: <https://www.who.int/publications/m/item/methods-for-estimating-the-excess-mortality-associatedwith-the-covid-19-pandemic> (accessed 23 June 2022) [↑](#footnote-ref-5)
6. The National Health Worker Accounts data is collated by the WHO and contains processed data extracts of the national reporting in the National Health Workforce Accounts data platform. Complementing the national reporting, additional sources such as the National Census, Labour Force Surveys and key administrative national and regional sources are also employed. The data can be accessed at: [https://apps.who.int/nhwaportal/](https://apps.who.int/nhwaportal/c) [↑](#footnote-ref-6)
7. ISIC is the International Standard Industrial Classification of All Economic Activities. Sector Q covers the provision of health and social work activities. This includes a wide range of activities, starting from health care provided by trained medical professionals in hospitals and other facilities, over residential care activities that still involve a degree of health care activities to social work activities without any involvement of health care professionals. <https://unstats.un.org/unsd/publication/seriesm/seriesm_4rev4e.pdf> (accessed 22 June 2022) [↑](#footnote-ref-7)
8. Taken from the World Bank, ILO and the United Nations (UN). These include, but were not limited to: GDP; the GDP spend on healthcare; population size; the proportion of the population aged 15 years and under; the proportion aged 65 years and over; net migration; the proportion of working-age population who were migrants; the proportion of the working population working in the health care sector, and information on the number of doctors and nurses per 1,000 people. [↑](#footnote-ref-8)
9. Office for National Statistics – Annual Population Survey 3-year pooled dataset 2016-18. [↑](#footnote-ref-9)
10. <https://databank.worldbank.org/source/population-estimates-and-projections> (accessed 05 July 2022) [↑](#footnote-ref-10)