Suggested title for paper:

Country differences in transmissibility, age distribution and case-fatality of SARS-CoV-2: a global ecological analysis

# Methods

## Analysis outcomes

### Reproduction number

We sourced time-varying reproduction numbers () by country, as estimated on a real-time basis by the Imperial College London MRC Centre for Global Infectious Disease Analysis [CITE]. A characteristic of these estimates is that they are informed by the evolution of observed deaths rather than cases. We averaged estimates over our analysis period.

### Age of observed cases and deaths

Age-specific data on cases were available from 61 countries and age-specific deaths from 39 (35 countries reported both). Eleven separate systems for age categorisation were used across countries, rendering direct comparisons difficult. Moreover, we wished to analyse differences in the average age of cases and deaths not merely explainable by countries’ demographic structure. Accordingly, for each country we present a “standardised median age” indicator , interpretable as the median age of cases or deaths if the country’s observed age-specific cumulative incidence or death rates were applied to the world’s population age structure. The quantity is computed as follows. Suppose a given country reports cumulative cases (or deaths) in each age category , where is censored at 100 years. We can calculate age-specific attack rates , where is the population within each age category. If the country had the same population and age structure as the world, its expected age-specific caseload would be , where denotes the world’s population. Accordingly, is the age at which the 50th percentile of the cumulative sum of occurs. Since age increments in our dataset were not annual, we linearly interpolated over the range before computing the median. We sourced projected age-specific population for each country and the world in 2020 from the United Nations World Population Prospects CITE.

### Case-fatality ratio

International comparisons of the lethality of SARS-CoV-2 are notoriously confounded by varying testing coverage, age distribution of cases, prevalence of co-morbidities and other factors [CITE. After omitting countries with < 50 total observed cases, we computed a crude case-fatality ratio (CFR) for each country by dividing observed deaths by cases. For countries with available age-specific data we also computed (i) an ‘age-standardised’ CFR, derived as above by applying countries’ age-specific crude CFRs to the world’s population structure; and (ii) an ‘incidence-standardised’ CFR derived by applying each country’s age-specific CFRs to the observed age-specific caseload in South Korea, selected as a reference due to this country’s reportedly high coverage of case detection (i.e. relatively low selection bias affecting the profile of observed cases) and standard of care [CITE]. Neither standardisation method fully removes confounding; (i) accounts for age differences in infection-fatality ratios (IFR) and (ii) reduces bias due to incomplete testing, but neither accounts for the effect of age on incidence.

## Analysis

We present three alternative approaches for exploring the effect of hypothesised exposures on each of the above outcomes, while adjusting for potential confounders and accounting for plausible effect modifications. For and the crude CFR, we did a global as well as an AFRO region-specific analysis (the latter was not possible due to data sparsity for the other outcomes).

### Linear regression

As each outcome was continuous and not structured hierarchically, we applied ordinary least-squares (OLS) fixed-effects models, guided by our a priori causal framework. For each outcome, we first observed potential collinearity among independent variables through scatterplots and Pearson correlation coefficients. We screened out potential confounding variables through visual analysis and a p-value < 0.20 threshold. We then fit models through backward variable selection, retaining variables that improved goodness of fit (adjusted ) or influenced the effect of putative exposures on the outcomes, while also re-introducing previously screened-out variables, trying alternative collinear variables, testing plausible two-way interactions, observing collinearity through variance inflation factors and verifying model assumptions including normality and homoscedasticity of residuals. While the model’s intent was explanatory rather than predictive, we used leave-one-out cross-validation (LOOCV) to assess potential overfitting. We present models both with all exposures forced into regression, and only significant and/or model-influential exposures retained.

### Principal component analysis

[Maybe – not sure I can pull this off, or that it has an advantage over random forests; PCA is useful in cases with lots of collinear variables, which is kind of our case…]

### Random forest regression

The random forest (RF) method of machine-learning imposes minimal statistical assumptions on data, and copes well with collinearity [CITE]. It consists of generating a large number of regression trees (i.e. partitions of the independent variables, with each variable generating a node or ‘split’) and averaging over these based on their accuracy in predicting the outcomes. We implemented two RF options for each outcome, using the randomForest R package: (i) using non-missing data only; (ii) imputing missing data through the rfImpute proximity method [CITE].

# Results

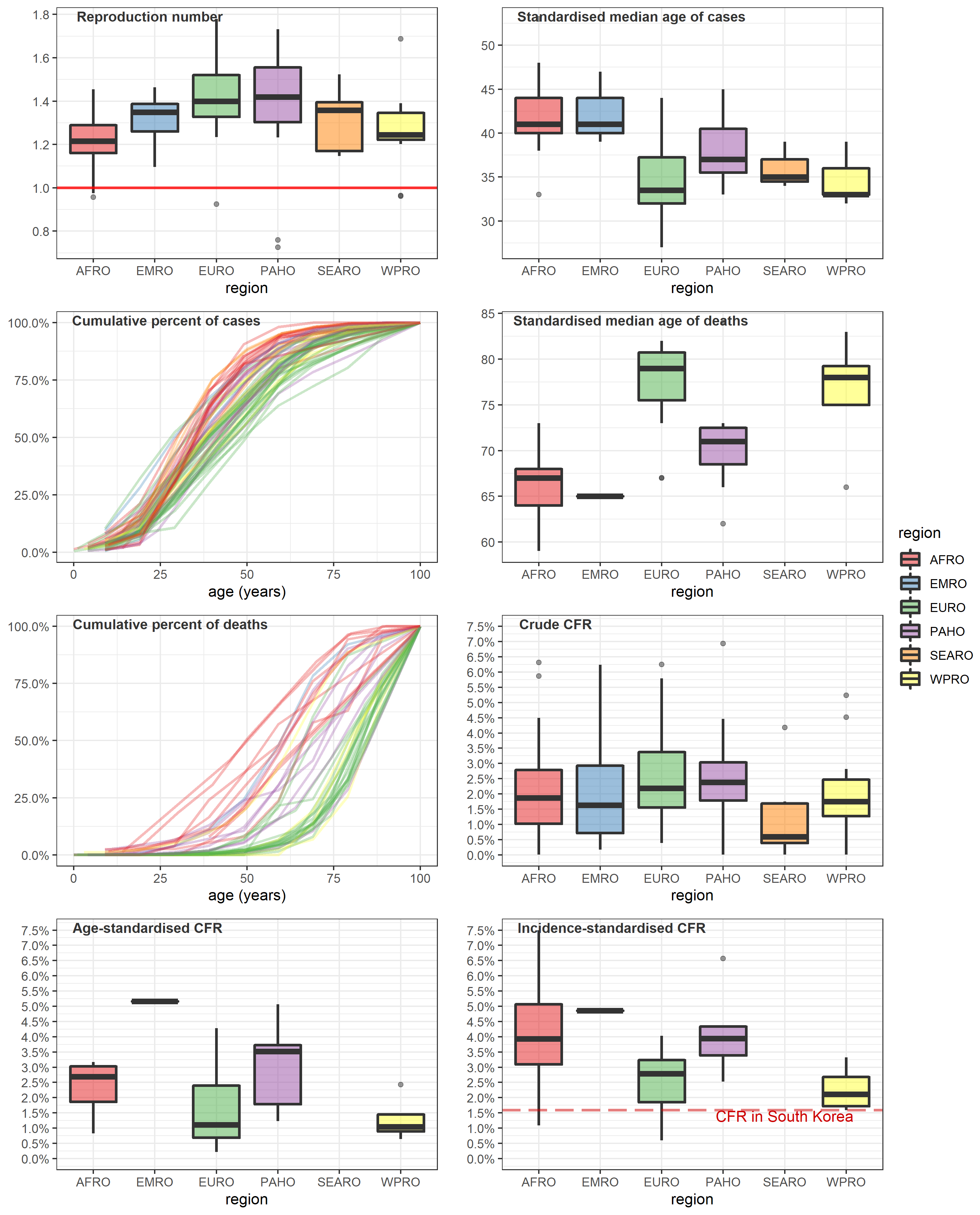


Figure . Analysis outcomes, by WHO region. All boxplots indicate the median and inter-quartile range (boxes), 95% percentile intervals (whiskers) and outliers (dots). CFR = case-fatality ratio.

# Annex: Additional tables and figures

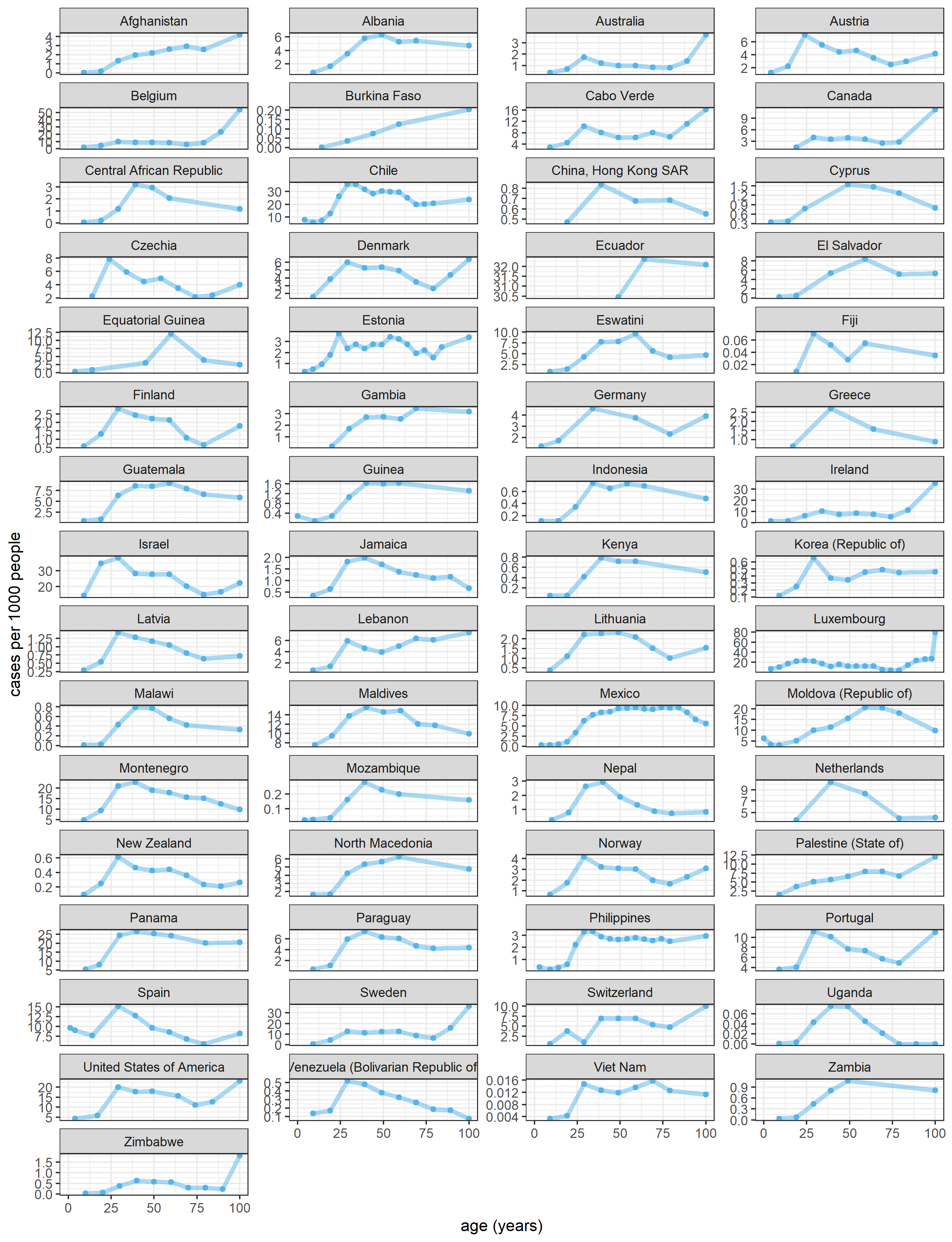


Figure S. Age-specific cumulative incidence rate, by country or territory (observed COVID-19 cases per 1000 population).

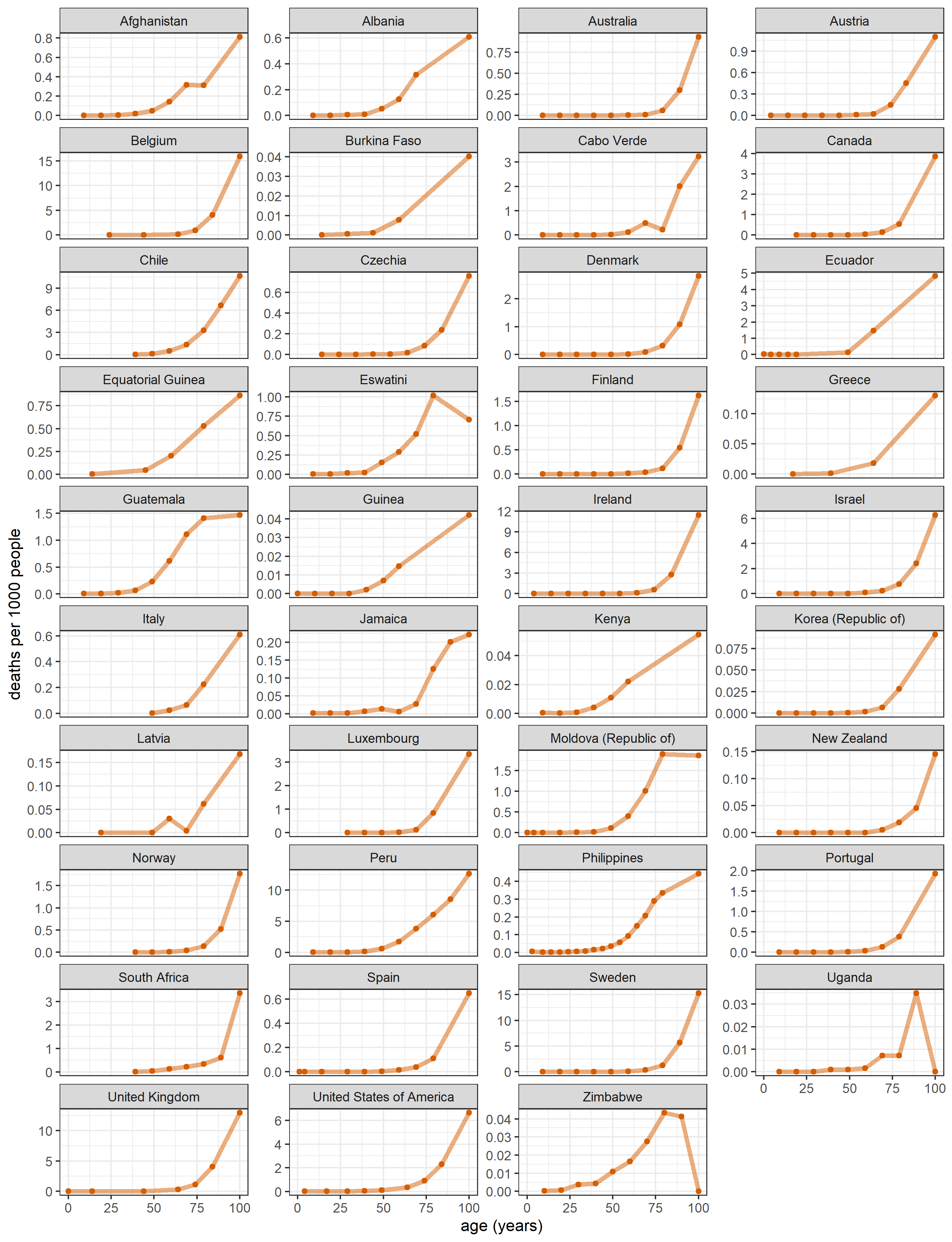


Figure S. Age-specific cumulative death rate, by country or territory (observed deaths due to COVID-19 per 1000 population).

Table S. Median and inter-quartile range of the age of cases and deaths, standardised by applying each country’s observed age-specific cumulative incidence and death rates to the world’s population age structure.

| Country or territory | Standardised median age of cases (inter-quartile range) | Standardised median age of deaths (inter-quartile range) |
| --- | --- | --- |
| Afghanistan | 47 (34 to 60) | 65 (55 to 76) |
| Albania | 41 (29 to 54) | 67 (58 to 82) |
| Australia | 32 (22 to 48) | 83 (75 to 89) |
| Austria | 30 (19 to 46) | 78 (70 to 87) |
| Belgium | 37 (24 to 54) | 81 (72 to 90) |
| Burkina Faso | 53 (38 to 73) | 73 (60 to 87) |
| Cabo Verde | 33 (22 to 51) | 69 (62 to 84) |
| Canada | 37 (24 to 53) | 84 (74 to 92) |
| Central African Republic | 40 (32 to 50) | n/a |
| Chile | 36 (25 to 51) | 72 (62 to 81) |
| Cyprus | 39 (26 to 53) | n/a |
| Czechia | 29 (18 to 44) | 78 (69 to 87) |
| Denmark | 33 (21 to 48) | 80 (71 to 87) |
| Ecuador | 49 (49 to 57) | 71 (57 to 85) |
| El Salvador | 44 (31 to 55) | n/a |
| Equatorial Guinea | 48 (32 to 56) | 63 (50 to 74) |
| Estonia | 37 (23 to 52) | n/a |
| Fiji | 34 (24 to 52) | n/a |
| Finland | 33 (22 to 47) | 80 (68 to 87) |
| Gambia | 45 (33 to 59) | n/a |
| Germany | 32 (20 to 48) | n/a |
| Greece | 32 (22 to 46) | 74 (59 to 87) |
| Guatemala | 42 (30 to 55) | 62 (53 to 72) |
| Guinea | 42 (31 to 55) | 71 (56 to 85) |
| Hong Kong | 33 (20 to 50) | n/a |
| Indonesia | 39 (27 to 53) | n/a |
| Ireland | 37 (26 to 54) | 82 (74 to 90) |
| Israel | 28 (17 to 45) | 78 (68 to 86) |
| Italy | n/a | 75 (66 to 86) |
| Jamaica | 35 (24 to 49) | 71 (50 to 79) |
| Kenya | 41 (31 to 54) | 67 (52 to 83) |
| Korea (Republic of) | 33 (22 to 53) | 78 (70 to 89) |
| Latvia | 34 (23 to 48) | 73 (56 to 85) |
| Lebanon | 39 (26 to 57) | n/a |
| Lithuania | 36 (24 to 49) | n/a |
| Luxembourg | 27 (16 to 42) | 82 (73 to 91) |
| Malawi | 40 (31 to 51) | n/a |
| Maldives | 35 (20 to 50) | n/a |
| Mexico | 45 (33 to 58) | n/a |
| Moldova (Republic of) | 44 (28 to 57) | 67 (59 to 75) |
| Montenegro | 35 (24 to 50) | n/a |
| Mozambique | 38 (29 to 52) | n/a |
| Nepal | 34 (25 to 44) | n/a |
| Netherlands | 33 (22 to 47) | n/a |
| New Zealand | 33 (23 to 49) | 78 (71 to 86) |
| North Macedonia | 39 (26 to 54) | n/a |
| Norway | 33 (22 to 48) | 80 (70 to 88) |
| Palestine (State of) | 41 (25 to 56) | n/a |
| Panama | 37 (25 to 52) | n/a |
| Paraguay | 38 (28 to 52) | n/a |
| Peru | n/a | 66 (57 to 76) |
| Philippines | 39 (28 to 54) | 66 (56 to 75) |
| Portugal | 33 (22 to 48) | 80 (69 to 90) |
| South Africa | n/a | 67 (55 to 80) |
| Spain | 28 (16 to 42) | 81 (68 to 90) |
| Swaziland | 42 (31 to 54) | 64 (53 to 73) |
| Sweden | 38 (26 to 53) | 81 (71 to 87) |
| Switzerland | 42 (31 to 55) | n/a |
| Uganda | 38 (30 to 47) | 68 (59 to 81) |
| United Kingdom | n/a | 77 (67 to 87) |
| United States of America | 35 (23 to 50) | 73 (60 to 83) |
| Venezuela (Bolivarian Republic of) | 33 (22 to 46) | n/a |
| Vietnam | 38 (25 to 55) | n/a |
| Zambia | 44 (33 to 65) | n/a |
| Zimbabwe | 41 (31 to 53) | 59 (46 to 70) |

Table . Crude, age-standardised and incidence-standardised case-fatality ratios, by country or territory. The last column is the ratio of incidence-standardised CFR for the country, relative to that of South Korea.

| Country or territory | Crude CFR (%) | Age-standardised CFR (%) | Incidence-standardised CFR (%) | Ratio of incidence-standardised CFR to South Korea’s |
| --- | --- | --- | --- | --- |
| Aruba | n/a | n/a | n/a | n/a |
| Afghanistan | 3.7 | 5.2 | 4.9 | 3.07 |
| Angola | 4.5 | n/a | n/a | n/a |
| Albania | 2.9 | 2.1 | 3 | 1.88 |
| United Arab Emirates | 0.5 | n/a | n/a | n/a |
| Argentina | 2.1 | n/a | n/a | n/a |
| Armenia | 2.2 | n/a | n/a | n/a |
| Antigua and Barbuda | 3.2 | n/a | n/a | n/a |
| Australia | 2.8 | 1.1 | 1.8 | 1.11 |
| Austria | 2.0 | 0.8 | 2.4 | 1.51 |
| Azerbaijan | 1.5 | n/a | n/a | n/a |
| Burundi | 0.2 | n/a | n/a | n/a |
| Belgium | 10.1 | n/a | n/a | n/a |
| Benin | 1.8 | n/a | n/a | n/a |
| Burkina Faso | 4.1 | 9.7 | 7.5 | 4.73 |
| Bangladesh | 1.4 | n/a | n/a | n/a |
| Bulgaria | 4.0 | n/a | n/a | n/a |
| Bahrain | 0.4 | n/a | n/a | n/a |
| Bahamas | 2.2 | n/a | n/a | n/a |
| Bosnia and Herzegovina | 3.0 | n/a | n/a | n/a |
| Belarus | 1.0 | n/a | n/a | n/a |
| Belize | 1.3 | n/a | n/a | n/a |
| Bolivia (Plurinational State of) | 4.5 | n/a | n/a | n/a |
| Brazil | 3.1 | n/a | n/a | n/a |
| Barbados | n/a | n/a | n/a | n/a |
| Brunei Darussalam | 2.1 | n/a | n/a | n/a |
| Bhutan | 0.0 | n/a | n/a | n/a |
| Botswana | 0.2 | n/a | n/a | n/a |
| Central African Republic | 1.3 | n/a | n/a | n/a |
| Canada | 6.9 | 3.7 | 3.9 | 2.49 |
| Switzerland | 3.4 | n/a | n/a | n/a |
| Chile | 2.7 | n/a | n/a | n/a |
| China | 5.2 | n/a | n/a | n/a |
| Côte d'Ivoire | 0.6 | n/a | n/a | n/a |
| Cameroon | 2.2 | n/a | n/a | n/a |
| Democratic Republic of the Congo | 2.6 | n/a | n/a | n/a |
| Congo | 2.0 | n/a | n/a | n/a |
| Colombia | 3.2 | n/a | n/a | n/a |
| Comoros | 1.8 | n/a | n/a | n/a |
| Cabo Verde | 1.0 | 1.6 | 2.4 | 1.49 |
| Costa Rica | 1.0 | n/a | n/a | n/a |
| Cuba | 2.3 | n/a | n/a | n/a |
| Curaçao | n/a | n/a | n/a | n/a |
| Cyprus | n/a | n/a | n/a | n/a |
| Czechia | 1.0 | 0.5 | 1.6 | 1.02 |
| Germany | 3.4 | n/a | n/a | n/a |
| Djibouti | 1.1 | n/a | n/a | n/a |
| Denmark | 2.6 | 1.1 | 2.8 | 1.76 |
| Dominican Republic | 1.9 | n/a | n/a | n/a |
| Algeria | 3.5 | n/a | n/a | n/a |
| Ecuador | 9.3 | 3.5 | 4.3 | 2.73 |
| Egypt | 5.5 | n/a | n/a | n/a |
| Eritrea | 0.0 | n/a | n/a | n/a |
| Western Sahara | n/a | n/a | n/a | n/a |
| Spain | 4.1 | 0.2 | 0.6 | 0.38 |
| Estonia | 2.1 | n/a | n/a | n/a |
| Ethiopia | 1.6 | n/a | n/a | n/a |
| Finland | 3.6 | 0.6 | 2.1 | 1.33 |
| Fiji | n/a | n/a | n/a | n/a |
| France | 5.8 | n/a | n/a | n/a |
| Micronesia (Fed. States of) | n/a | n/a | n/a | n/a |
| Gabon | 0.6 | n/a | n/a | n/a |
| United Kingdom | 9.4 | n/a | n/a | n/a |
| Georgia | 0.6 | n/a | n/a | n/a |
| Ghana | 0.6 | n/a | n/a | n/a |
| Guinea | 0.6 | 0.8 | 1.1 | 0.68 |
| Guadeloupe | n/a | n/a | n/a | n/a |
| Gambia | 3.4 | n/a | n/a | n/a |
| Guinea-Bissau | 1.5 | n/a | n/a | n/a |
| Equatorial Guinea | 1.7 | 3.2 | 5.3 | 3.34 |
| Greece | 2.2 | 1.1 | 3.3 | 2.1 |
| Grenada | n/a | n/a | n/a | n/a |
| Guatemala | 3.6 | 5.1 | 6.6 | 4.14 |
| French Guiana | n/a | n/a | n/a | n/a |
| Guam | n/a | n/a | n/a | n/a |
| Guyana | 3.0 | n/a | n/a | n/a |
| China, Hong Kong SAR | 2.1 | n/a | n/a | n/a |
| Honduras | 3.0 | n/a | n/a | n/a |
| Croatia | 1.7 | n/a | n/a | n/a |
| Haiti | 32.6 | n/a | n/a | n/a |
| Hungary | 4.1 | n/a | n/a | n/a |
| Indonesia | 4.2 | n/a | n/a | n/a |
| India | 1.7 | n/a | n/a | n/a |
| Ireland | 4.7 | 3.5 | 3.9 | 2.48 |
| Iran (Islamic Republic of) | 0.6 | n/a | n/a | n/a |
| Iraq | 3.0 | n/a | n/a | n/a |
| Iceland | 0.4 | n/a | n/a | n/a |
| Israel | 0.6 | 0.4 | 1.5 | 0.95 |
| Italy | 11.6 | n/a | n/a | n/a |
| Jamaica | 1.1 | 1.2 | 2.5 | 1.6 |
| Jordan | 0.7 | n/a | n/a | n/a |
| Japan | 1.9 | n/a | n/a | n/a |
| Kazakhstan | 1.6 | n/a | n/a | n/a |
| Kenya | 1.7 | 2.8 | 3.8 | 2.4 |
| Kyrgyzstan | 2.3 | n/a | n/a | n/a |
| Cambodia | 0.0 | n/a | n/a | n/a |
| Kiribati | n/a | n/a | n/a | n/a |
| *Republic of Korea* | *1.6* | *1.0* | *1.6* | *1.0* |
| Kuwait | 0.6 | n/a | n/a | n/a |
| Lao People's Democratic Republic | n/a | n/a | n/a | n/a |
| Lebanon | 1.6 | n/a | n/a | n/a |
| Liberia | 6.3 | n/a | n/a | n/a |
| Libya | 1.7 | n/a | n/a | n/a |
| Saint Lucia | n/a | n/a | n/a | n/a |
| Sri Lanka | 0.4 | n/a | n/a | n/a |
| Lesotho | 2.9 | n/a | n/a | n/a |
| Lithuania | 2.0 | n/a | n/a | n/a |
| Luxembourg | 1.4 | n/a | n/a | n/a |
| Latvia | 2.1 | 4.3 | 3.8 | 2.43 |
| China, Macao SAR | n/a | n/a | n/a | n/a |
| Morocco | 1.9 | n/a | n/a | n/a |
| Republic of Moldova | 2.5 | 2.2 | 2.8 | 1.78 |
| Madagascar | 1.3 | n/a | n/a | n/a |
| Maldives | 0.4 | n/a | n/a | n/a |
| Mexico | 10.6 | n/a | n/a | n/a |
| North Macedonia | 4.1 | n/a | n/a | n/a |
| Mali | 2.8 | n/a | n/a | n/a |
| Malta | 1.1 | n/a | n/a | n/a |
| Myanmar | 0.5 | n/a | n/a | n/a |
| Montenegro | 1.6 | n/a | n/a | n/a |
| Mongolia | 4.5 | n/a | n/a | n/a |
| Mozambique | 0.6 | n/a | n/a | n/a |
| Mauritania | 2.3 | n/a | n/a | n/a |
| Martinique | n/a | n/a | n/a | n/a |
| Mauritius | 2.9 | n/a | n/a | n/a |
| Malawi | 3.2 | n/a | n/a | n/a |
| Malaysia | 1.4 | n/a | n/a | n/a |
| Mayotte | n/a | n/a | n/a | n/a |
| Namibia | 0.9 | n/a | n/a | n/a |
| New Caledonia | n/a | n/a | n/a | n/a |
| Niger | 5.9 | n/a | n/a | n/a |
| Nigeria | 1.9 | n/a | n/a | n/a |
| Nicaragua | 3.0 | n/a | n/a | n/a |
| Netherlands | 5.6 | n/a | n/a | n/a |
| Norway | 2.0 | 2.5 | 2.8 | 1.76 |
| Nepal | 0.6 | n/a | n/a | n/a |
| New Zealand | 1.2 | 0.6 | 2.5 | 1.56 |
| Oman | 0.8 | n/a | n/a | n/a |
| Pakistan | 2.1 | n/a | n/a | n/a |
| Panama | 2.1 | n/a | n/a | n/a |
| Peru | 17.8 | n/a | n/a | n/a |
| Philippines | 1.6 | 2.4 | 3.3 | 2.09 |
| Papua New Guinea | 1.1 | n/a | n/a | n/a |
| Poland | 2.8 | n/a | n/a | n/a |
| Puerto Rico | n/a | n/a | n/a | n/a |
| Dem. People's Republic of Korea | n/a | n/a | n/a | n/a |
| Portugal | 2.6 | 1.0 | 1.8 | 1.12 |
| Paraguay | 1.9 | n/a | n/a | n/a |
| State of Palestine | 0.7 | n/a | n/a | n/a |
| French Polynesia | n/a | n/a | n/a | n/a |
| Qatar | 0.2 | n/a | n/a | n/a |
| Réunion | n/a | n/a | n/a | n/a |
| Romania | 3.8 | n/a | n/a | n/a |
| Russian Federation | 1.8 | n/a | n/a | n/a |
| Rwanda | 0.4 | n/a | n/a | n/a |
| Saudi Arabia | 1.2 | n/a | n/a | n/a |
| Sudan | 6.2 | n/a | n/a | n/a |
| Senegal | 2.1 | n/a | n/a | n/a |
| Singapore | 0.0 | n/a | n/a | n/a |
| Solomon Islands | n/a | n/a | n/a | n/a |
| Sierra Leone | 3.4 | n/a | n/a | n/a |
| El Salvador | 2.9 | n/a | n/a | n/a |
| Somalia | 2.9 | n/a | n/a | n/a |
| Serbia | 2.2 | n/a | n/a | n/a |
| South Sudan | 1.9 | n/a | n/a | n/a |
| Sao Tome and Principe | 1.7 | n/a | n/a | n/a |
| Suriname | n/a | n/a | n/a | n/a |
| Slovakia | 0.5 | n/a | n/a | n/a |
| Slovenia | 2.2 | n/a | n/a | n/a |
| Sweden | 6.3 | 2.7 | 4.0 | 2.54 |
| Eswatini | 2.0 | 3.1 | 4.8 | 3.05 |
| Seychelles | 0.0 | n/a | n/a | n/a |
| Syrian Arab Republic | 4.1 | n/a | n/a | n/a |
| Chad | 7.7 | n/a | n/a | n/a |
| Togo | 2.1 | n/a | n/a | n/a |
| Thailand | 1.7 | n/a | n/a | n/a |
| Tajikistan | 0.8 | n/a | n/a | n/a |
| Turkmenistan | n/a | n/a | n/a | n/a |
| Timor-Leste | n/a | n/a | n/a | n/a |
| Tonga | n/a | n/a | n/a | n/a |
| Trinidad and Tobago | 1.7 | n/a | n/a | n/a |
| Tunisia | 1.9 | n/a | n/a | n/a |
| Turkey | 2.6 | n/a | n/a | n/a |
| China, Taiwan Province of China | n/a | n/a | n/a | n/a |
| United Republic of Tanzania | n/a | n/a | n/a | n/a |
| Uganda | 1.0 | n/a | n/a | n/a |
| Ukraine | 1.8 | n/a | n/a | n/a |
| Uruguay | 2.4 | n/a | n/a | n/a |
| United States of America | 3.0 | 1.8 | 3.4 | 2.13 |
| Uzbekistan | 0.8 | n/a | n/a | n/a |
| Saint Vincent and the Grenadines | 0.0 | n/a | n/a | n/a |
| Venezuela (Bolivarian Republic of) | 0.8 | n/a | n/a | n/a |
| United States Virgin Islands | 1.5 | n/a | n/a | n/a |
| Viet N/am | 2.6 | n/a | n/a | n/a |
| Vanuatu | n/a | n/a | n/a | n/a |
| Samoa | n/a | n/a | n/a | n/a |
| Yemen | 28.9 | n/a | n/a | n/a |
| South Africa | 2.2 | n/a | n/a | n/a |
| Zambia | 2.4 | n/a | n/a | n/a |
| Zimbabwe | 3.0 | 2.6 | 3.9 | 2.48 |

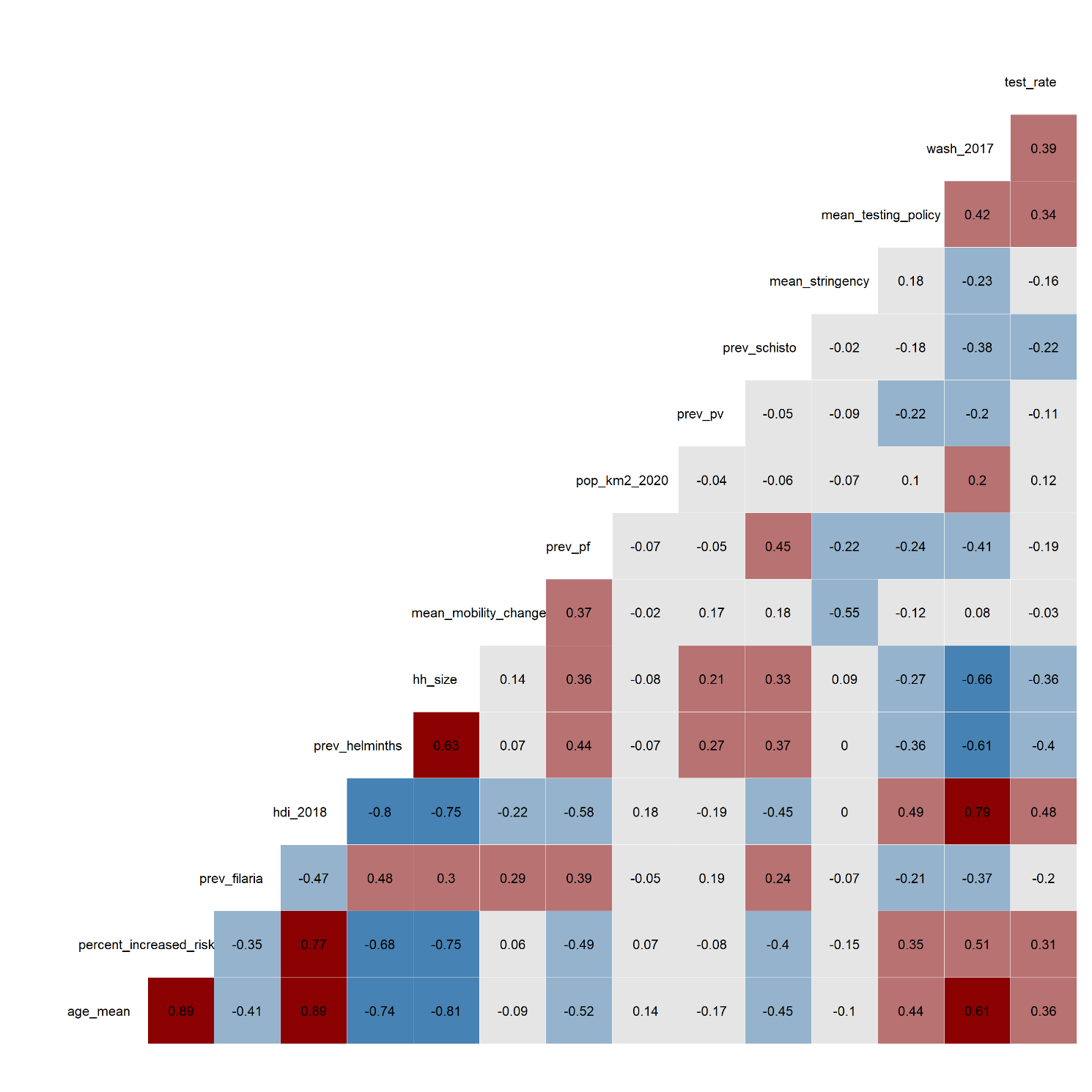


Figure S. Variable collinearity heatmap. Values in each cell are the Pearson correlation coefficient for a given pair of variables. Coefficients close to -1 or +1 indicate high correlation, with 0 indicating minimal correlation.

Table S. Pairs of highly collinear variables (Pearson correlation coefficient > |0.7| ).

|  |  |
| --- | --- |
| Variable 1 | Variable 2 |
| Mean age of the country’s population | Percent of the population at increased risk of COVID-19 |
| Mean age of the country’s population | Human Development Index |
| Mean age of the country’s population | Prevalence of helminths |
| Mean age of the country’s population | Household size |
| Household size | Percent of the population at increased risk of COVID-19 |
| Human Development Index | Percent of the population at increased risk of COVID-19 |
| Human Development Index | Prevalence of helminths |
| Human Development Index | Household size |
| Human Development Index | Sanitation index |

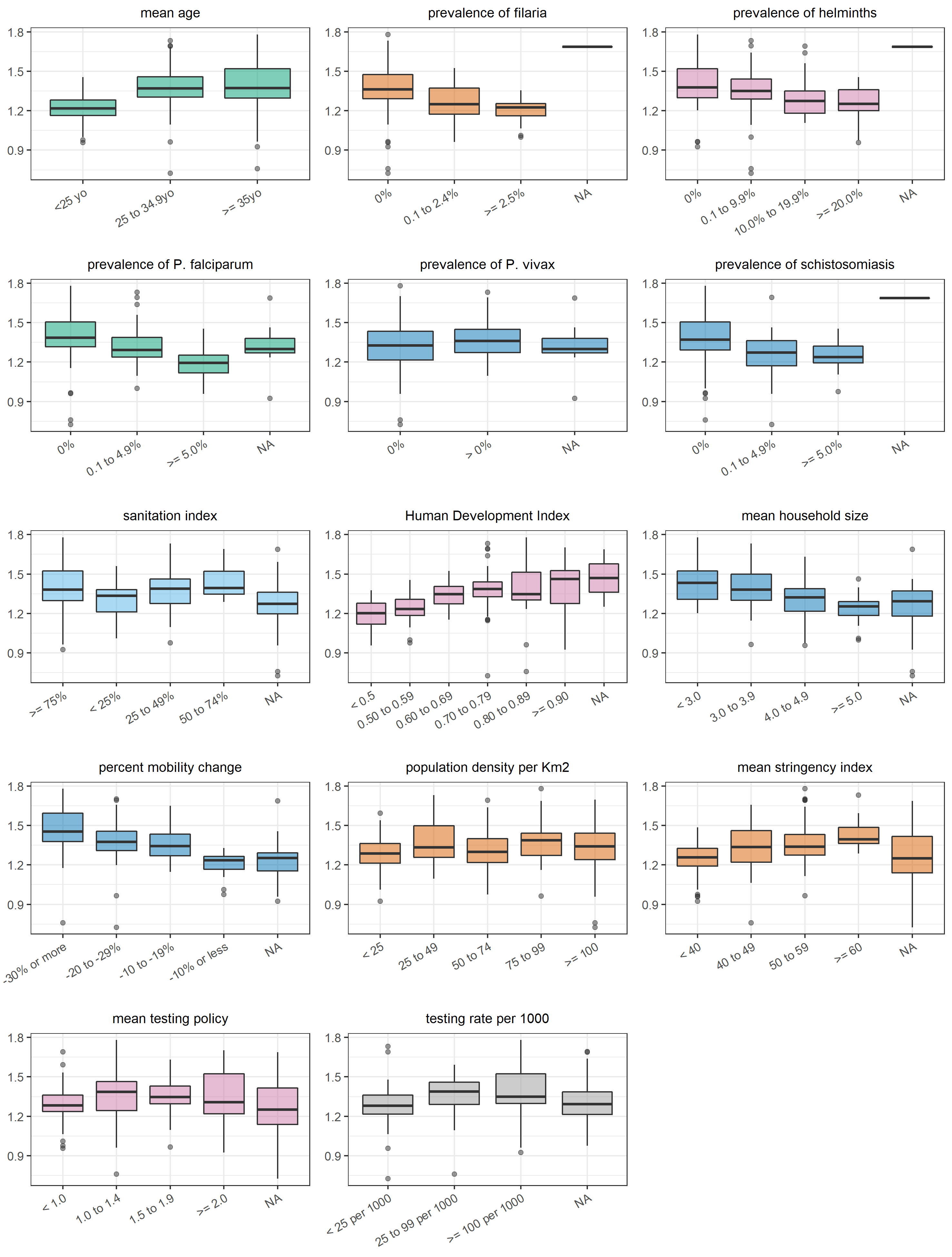


Figure S. Correlation between each independent variable (categorised) and the reproduction number.

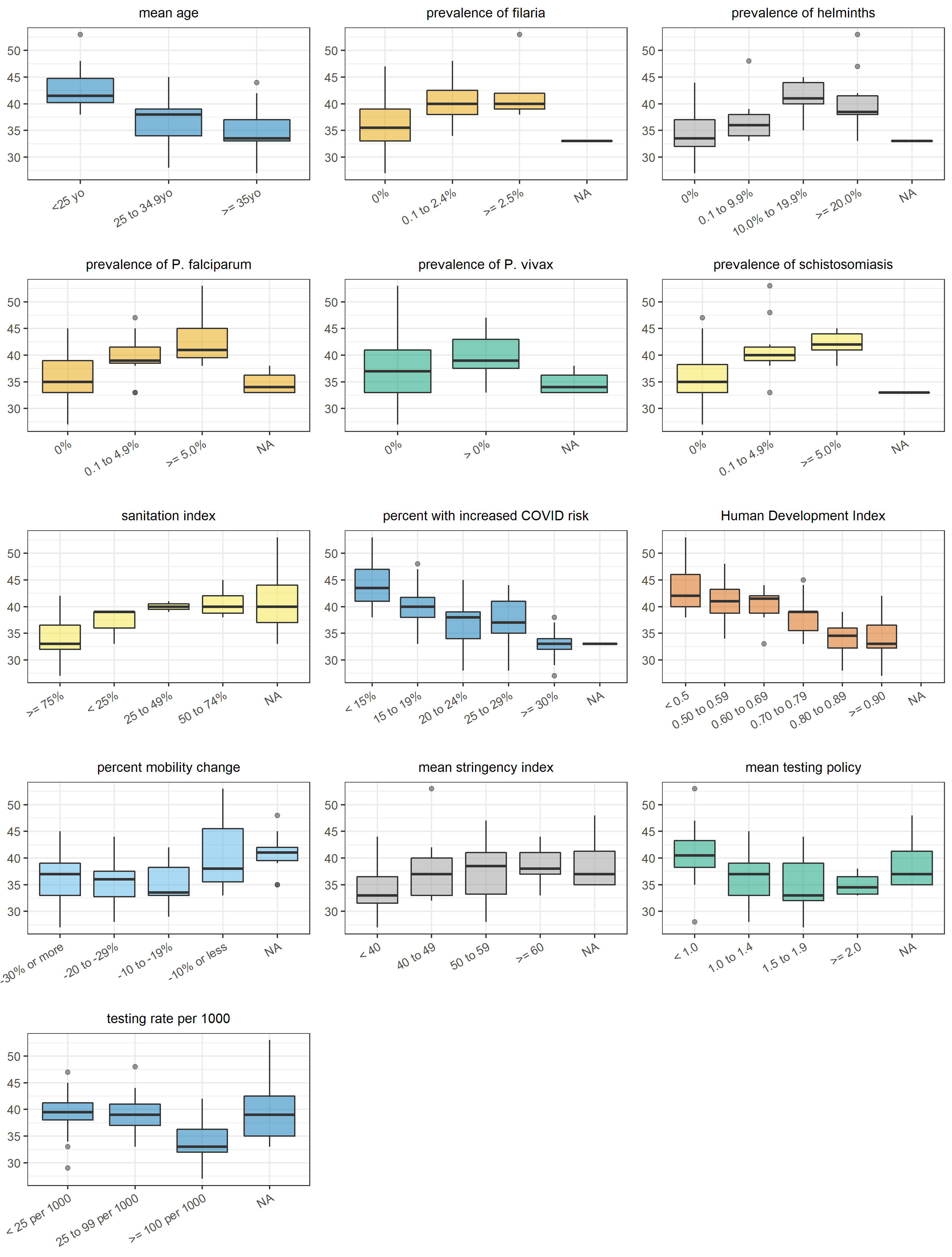


Figure S6. Correlation between each independent variable (categorised) and the age-standardised median age of cases.

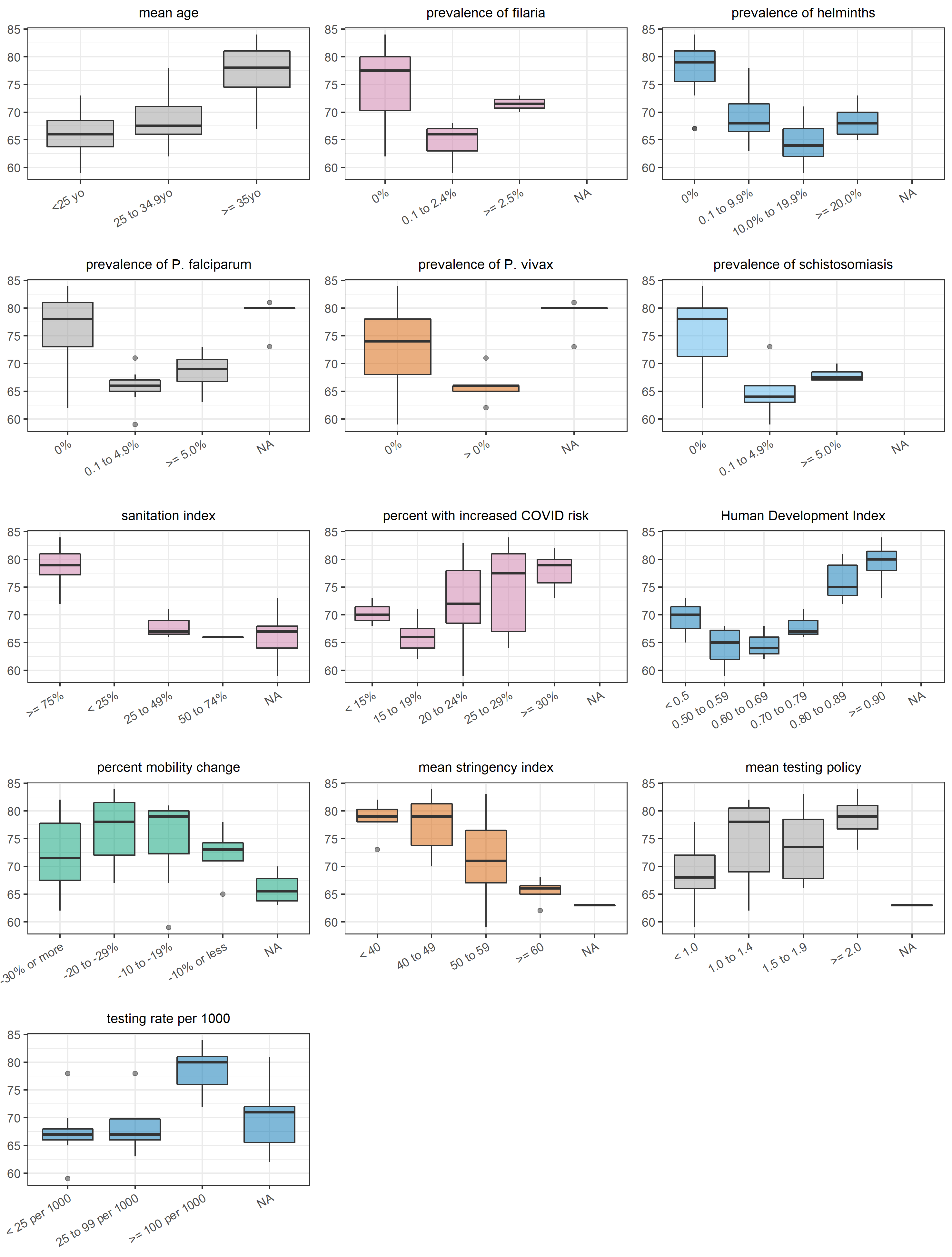


Figure S7. Correlation between each independent variable (categorised) and the age-standardised median age of deaths.

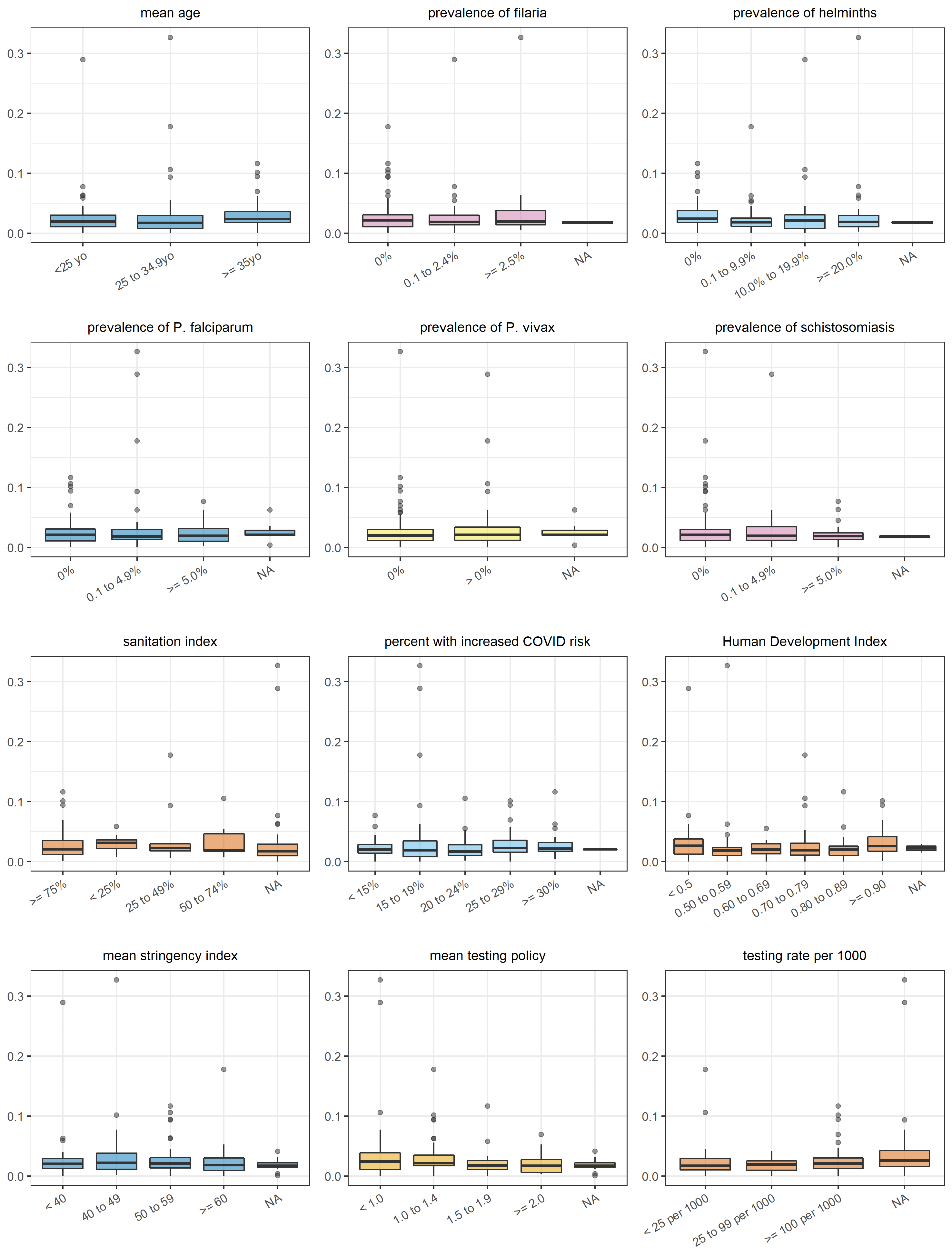


Figure S8. Correlation between each independent variable (categorised) and the crude case-fatality ratio.

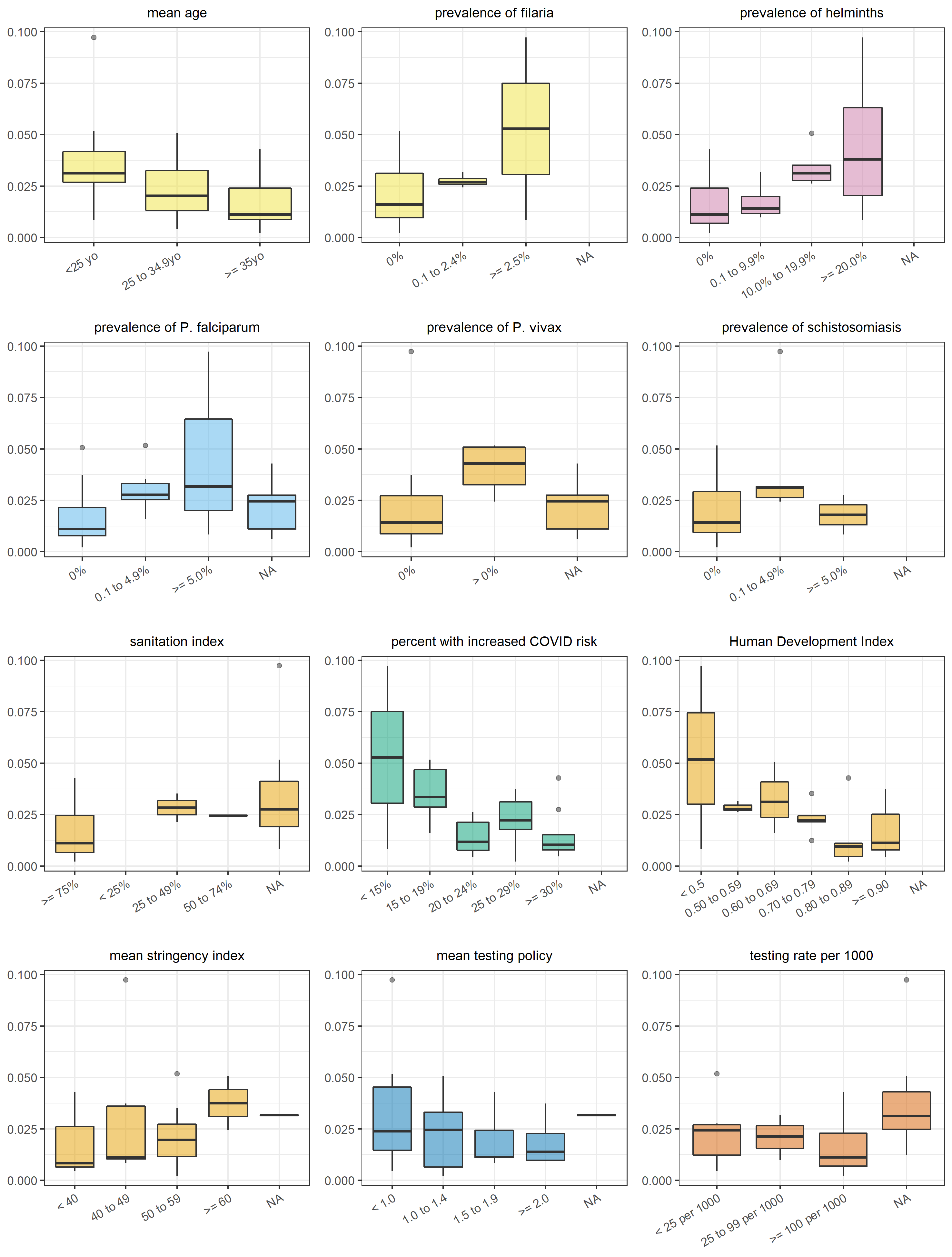


Figure S9. Correlation between each independent variable (categorised) and the age-standardised case-fatality ratio.

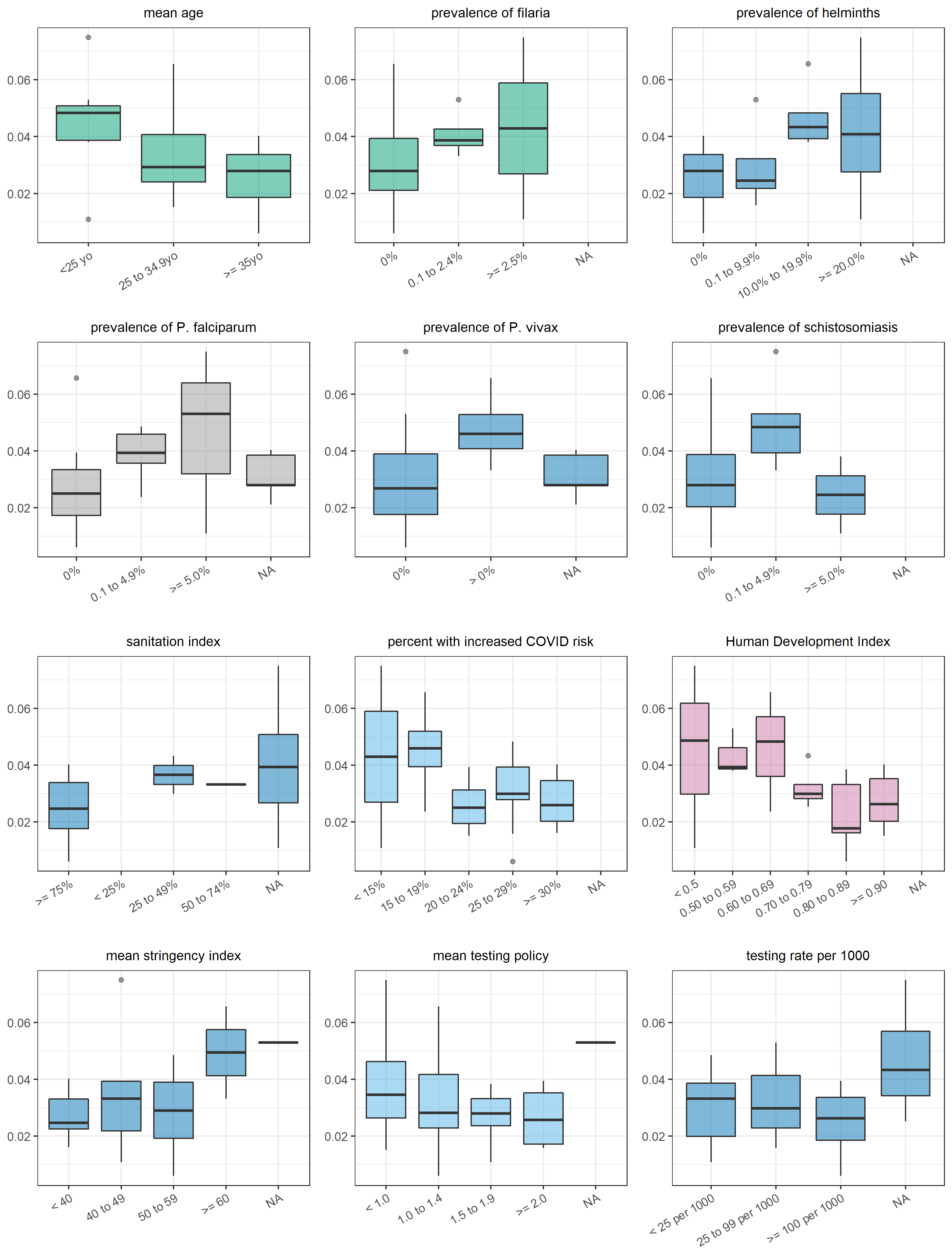


Figure S10. Correlation between each independent variable (categorised) and the incidence-standardised case-fatality ratio.