

1 **Asymptomatic SARS-CoV-2 infections tend to occur less frequently in developed**
2 **nations**

3 Shreya Chowdhury^{1,†}, Akshay Tiwari^{1,†}, Ananthu James^{1,§}, Budhaditya Chatterjee^{2,#,*},
4 Narendra M. Dixit^{1,2,*}

5 ¹Department of Chemical Engineering, Indian Institute of Science, Bangalore, India
6 560012

7 ²Department of Bioengineering, Indian Institute of Science, Bangalore, India 560012

8 [§]Current address: Department of Public Health, Erasmus MC, University Medical Center
9 Rotterdam, Rotterdam, The Netherlands

10 [#]Current address: Department of Microbiology and Immunology, School of Medicine,
11 University of Maryland Baltimore, Baltimore, MD, USA and NIAID, NIH, Bethesda, MD,
12 USA

13 †These authors contributed equally

14 *Correspondence: budhadityac@iisc.ac.in, narendra@iisc.ac.in

15 Manuscript details:

16 Abstract: 150 words; Text: ~4000 words (excluding methods, figure legends and
17 references)

18 Display items: 4 (3 figures, 1 table)

19 References: 74

20 Supplementary items: Texts – 3, Tables – 9, Figures – 8, References – 44.

21 **ABSTRACT**

22 Unlike severe infections, asymptomatic infections occur independently of healthcare
23 access and reflect the natural immunity to SARS-CoV-2. What determines their
24 prevalence, ψ , and its variation across nations is unknown. We conducted a systematic
25 review of serosurveys performed on samples representative of national populations
26 before vaccination and the emergence of variants. The studies that met our selection
27 criteria together sampled 4,58,489 individuals and yielded estimates of ψ in 33 nations.
28 Using random-effects modeling, we found the pooled global ψ to be 45.3% (95% CI:
29 33.6%-57.5%). ψ varied widely across nations (range: 6%-96%; $I^2=99.7\%$), highlighting
30 the enormous underlying variation in the natural immunity to SARS-CoV-2. Performing
31 meta-regression with national-level metrics, we found that the human development index
32 (HDI) was negatively correlated with ψ ($p=10^{-13}$; $R^2=65.5\%$). More developed nations
33 thus experienced less frequent asymptomatic infections on average. These findings have
34 implications for unraveling the origins of asymptomatic infections and for future pandemic
35 preparedness.

36 INTRODUCTION

37 Asymptomatic infections have been an enigmatic feature of the COVID-19 pandemic¹.
38 They stand at the favorable end of the spectrum of outcomes following SARS-CoV-2
39 infection, with the other end being severe respiratory distress and mortality². Their role in
40 driving the spread of the infection was recognized early in the pandemic^{3,4}. Significant
41 efforts have since been made to estimate their prevalence^{3,5-7}. Yet, unlike severe
42 infections and their associated healthcare burden which were found to vary substantially
43 across nations^{8,9}, how the prevalence of asymptomatic infections varies across nations
44 remains unknown. Understanding this variation is important. It has implications for
45 forecasting the course of the pandemic and evaluating intervention strategies^{10,11}. It also
46 offers insights into the variation in the natural immunity (as opposed to vaccine-induced
47 immunity) to SARS-CoV-2, a globally new pathogen, across populations: The prevalence
48 of asymptomatic infections, unlike other metrics of disease burden such as the infection
49 fatality ratio⁸, is expected to be robust to differences in health infrastructure and non-
50 pharmaceutical interventions. It is therefore a more reliable indicator of the natural ability
51 of populations to fight the infection, differences in which may arise, for instance, from
52 genetic polymorphisms or differential immunological memory from exposure to related
53 pathogens^{12,13}.

54 Understanding the variation in the prevalence of asymptomatic infections has been
55 challenging for two key reasons. First, epidemiological studies on asymptomatic
56 infections have often resorted to sampling from convenient sources, such as healthcare
57 workers, which introduces biases in the estimates of prevalence^{14,15}. For instance, of the
58 568 positive cases in the New York city jail system, only 2.6% were asymptomatic¹⁶,
59 whereas of the 61 individuals who tested positive in two large academic health systems
60 in Wisconsin, 88.5% remained asymptomatic¹⁷. The overall health of individuals in jails is
61 expected to differ vastly from those in academic health systems, potentially resulting in
62 the extreme variation in the prevalence observed across the two samples. Neither of the
63 samples, however, would be representative of the general populations of the two states
64 or of the USA.

65 Second, factors that could explain the observed variation have not been forthcoming.
66 Meta-analyses have identified factors associated with asymptomatic infections, like the
67 absence of comorbidities, but have not explained the large variation observed^{3,5-7}. A
68 recent study discovered a significantly higher frequency of the allele *HLA-B*15:01* in
69 asymptotically infected individuals than in those displaying symptoms, unravelling
70 potential genetic underpinnings of asymptomatic infections¹³. The allele, however, was
71 present in only ~11% of the asymptotically infected individuals (and in ~5% of the
72 symptomatic cases)¹³, leaving a vast majority of the asymptomatic cases unexplained.

73 Here, we conducted a systematic review of randomized serosurveys, following PRISMA
74 guidelines¹⁸, to identify studies that reported on asymptomatic infections in samples
75 representative of the general populations of nations. We restricted our review to the early
76 phase of the pandemic when infections were predominantly due to the ancestral SARS-

77 CoV-2 strain and vaccination programs were yet to be initiated. From the studies
78 identified, which spanned nations across continents and socio-economic developmental
79 status, we estimated the extent of variation in the prevalence of asymptomatic infections
80 across nations. We then performed meta-regression to identify key predictors of the
81 prevalence.

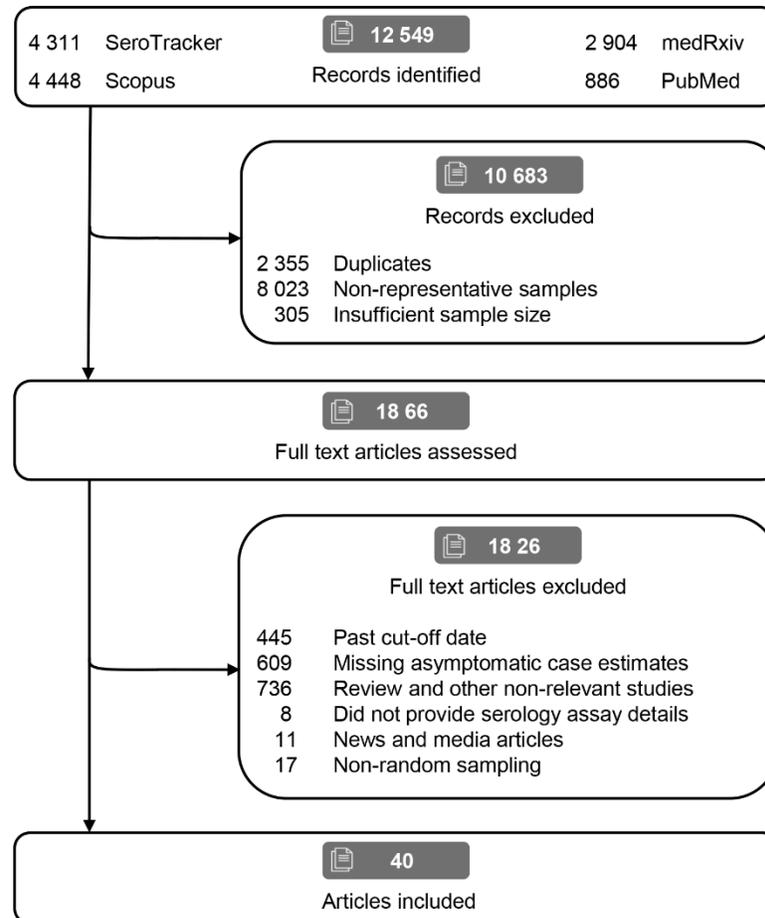
82 RESULTS

83 Systematic review and the global prevalence of asymptomatic infections

84 We conducted a systematic literature search and identified 12,549 records from four
85 databases that matched our search criteria (4311 from SeroTracker, 4448 from Scopus,
86 2904 from MedRxiv, and 886 from PubMed, Fig. 1). Applying our inclusion criteria to these
87 records (Fig. 1, Text S1, Methods), we identified 40 articles that offered estimates of the
88 percentage of asymptomatic infections (ψ) from serosurveys on samples representative
89 of the general populations of nations. For identifying asymptomatic infections,
90 serosurveys have an advantage over studies using nucleic acid tests because antibodies
91 last many months after the infection is cleared, allowing a larger time window for detection
92 of infected individuals. They are also, therefore, less likely to encounter a bias from
93 presymptomatic individuals¹⁴. Among the 40 studies, we found 7 instances where more
94 than one study was from the same nation (Table S1). Interestingly, in all but one of these
95 instances, the different studies from the same nation reported consistent estimates of ψ
96 (Table S1), giving us confidence in the estimates. For instance, the two studies from
97 Austria^{19,20} reported ψ of 18.0% and 19.7%. This similarity was despite the vastly different
98 seroprevalence—45% and 8.2%, respectively—in the studies, reiterating the notion that ψ
99 is an intrinsic property of the populations and is robust to variations in seroprevalence,
100 the latter typically dependent on the state of the pandemic and on non-pharmacological
101 interventions⁷. Given this consistency, we chose one study for each nation—the one with
102 the largest sample size, for its better representativeness of the national population—for
103 further analysis.

104 Thus, we had estimates of ψ from 33 studies, one for each of the 33 nations (Table 1;
105 Text S2 provides a brief description of each study). The nations were spread across
106 continents – Asia, the Americas, Europe, and Africa – and across socio-economic
107 developmental status²¹. Of the studies, 21 (63.6%) were nation-wide serosurveys,
108 whereas 6 (18.2%) were conducted at subnational levels and 6 (18.2%) were local
109 serosurveys but using samples representative of the general populations of the respective
110 nations (Table 1). Using a grading scale we developed by adapting the Joanna Briggs
111 Institute (JBI) checklist²² (Methods; Text S3), we assessed the 33 studies and found 14
112 (42.4%) to be of high quality, 14 (42.4%) of moderate quality, and 5 (15.2%) of low quality
113 (Table 1, Table S2). The studies together examined a total of 4,58,489 individuals, of
114 whom 30,086 individuals were seropositive. Of the latter, 15,615 individuals were
115 asymptomatic. To obtain a global average, we estimated the pooled ψ across the studies

116 by random effects modeling (Fig. 2, Table S3, Methods). The pooled ψ was 45.3% (95%
117 CI: 33.6%-57.5%).



118 **Figure 1: PRISMA flow diagram.** The search results and the selection steps leading to
119 the studies included in our meta-analysis.

120 Variation of the prevalence of asymptomatic infections across nations

121 Interestingly, the estimates of ψ revealed an enormous variation across nations (Table 1,
122 Fig. 2). The Netherlands had the lowest ψ , 6%, and India the highest, 96%. 7 nations had
123 ψ between 75% and 100%, 4 between 50% and 75%, 16 between 25% and 50%, and 6
124 below 25%. We estimated confidence intervals on ψ using the Wilson score interval
125 (Methods). The intervals closely matched estimates where available (Table S4). Small
126 sample sizes often led to large confidence intervals, as with Georgia (Fig. 2).
127 Nonetheless, the data yielded an inconsistency index (I^2) of 99.7% (Methods), implying
128 extremely high inter-nation heterogeneity, not attributable to sampling variations.

129 **Table 1: Summary of the studies included in systematic review.** ψ is percentage of
 130 asymptomatic infections; Q is study quality, with high (H), medium (M), and low (L)
 131 categories; G is the geographic scope of the study, with national (N), sub-national (S),
 132 and local (L) categories. In cases where the study months but not exact dates were
 133 mentioned, the dates have been set to the 15th of the respective months.

Nation	Study period	Cohort size	Seropositive	Asymptomatic	ψ (95% CI)	Q	G
Austria ¹⁹	21-Apr-20 to 27-Apr-20	1259	566	102	18.0 (15.1- 21.4)	M	L
Brazil ²³	21-Jun-20 to 24-Jun-20	31869	849	103	12.1 (10.1-14.5)	H	N
Canada ²⁴	01-Dec-20 to 15-Jan-21	6955	444	94	21.2 (17.7-25.2)	H	N
Chile ²⁵	25-Sept-20 to 25-Nov-20	2493	242	89	36.7 (31.0-43.0)	H	S
China ²⁶	10-Apr-20 to 18-Apr-20	34857	427	326	76.3 (72.0-80.0)	H	N
Colombia ²⁷	15-Jul-20 to 15-Nov-20	2564	1045	377	36.1 (33.2-39.1)	H	S
Denmark ²⁸	15-Aug-20 to 11-Dec-20	13095	369	112	30.4 (25.9-35.3)	H	N
Ethiopia ²⁹	24-Jun-20 to 08-Jul-20	16932	314	292	93.3 (90.0-95.6)	M	N
France ³⁰	26-Jun-20 to 24-Jul-20	2006	43	7	16.3 (8.1-30.0)	M	L
Georgia ³¹	18-May-20 to 27-May-20	1068	9	5	55.6 (26.7-81.2)	L	L
Germany ³²	20-May-20 to 09-Jun-20	2203	249	61	24.5 (19.6-30.2)	M	L
Ghana ³³	27-Jul-20 to 14-Sep-20	1305	252	199	79.1 (73.7-83.7)	L	N
Hungary ³⁴	01-May-20 to 16-May-20	10474	70	32	47.1 (35.9-58.6)	M	N
India ³⁵	18-Aug-20 to 20-Sep-20	29082	3135	3029	96.6 (95.9-97.2)	M	N
Iran ³⁶	17-Apr-20 to 02-Jun-20	3530	1164	416	35.7 (33.0-38.5)	H	N
Ireland ³⁷	22-Jun-20 to 16-Jul-20	1733	33	9	27.3 (15.1-44.2)	M	S
Italy ³⁸	25-May-20 to 15-Jul-20	64660	1617	441	27.3 (25.2-29.5)	H	N
Jordan ³⁹	27-Dec-21 to 06-Jan-21	5044	1723	1071	62.2 (59.9-64.5)	L	N
Lithuania ⁴⁰	10-Aug-20 to 10-Sep-20	3087	43	30	69.0 (54.1-80.8)	H	N
Luxembourg ⁴¹	15-Apr-20 to 05-May-20	1807	35	10	28.6 (16.3-45.0)	H	N
Mexico ⁴²	15-Aug-20 to 15-Nov-20	9640	2400	1615	67.3 (65.4-69.1)	H	N
Netherlands ⁴³	31-Mar-20 to 11-May-20	3147	74	5	6.8 (2.9-14.9)	M	N
Oman ⁴⁴	12-Jul-20 to 08-Dec-20	17457	3841	3562	92.7 (91.8-93.4)	M	N
Pakistan ⁴⁵	15-Apr-20 to 22-Aug-20	3005	364	333	91.5 (88.2-94.0)	M	L
Peru ⁴⁶	24-Jun-20 to 10-Jul-20	2010	595	151	25.4 (0.22-0.29)	M	L
Portugal ⁴⁷	21-May-20 to 08-Jul-20	2301	67	29	44.0 (32.8-55.9)	H	N
Senegal ⁴⁸	25-Oct-20 to 26-Nov-20	1422	398	129	32.4 (28.0-37.1)	M	N
Spain ⁴⁹	27-Apr-20 to 22-Jun-20	61092	2669	766	28.7 (26.1-31.4)	M	N
Switzerland ⁵⁰	06-Apr-20 to 30-Jun-20	8344	590	77	13.1 (10.6-16.1)	H	S
UK ⁵¹	20-Jun-20 to 13-Jul-20	105651	5544	1785	32.2 (31.0-33.4)	H	N
USA ⁵²	15-May-20 to 15-Jan-21	4510	161	45	28.0 (21.6-35.4)	L	N
Yemen ⁵³	15-Nov-20 to 15-Dec-21	2001	549	157	28.6 (0.25-0.33)	L	S
Zambia ⁵⁴	04-Jul-20 to 27-Jul-20	1886	205	156	76.2 (69.9-81.5)	M	S

135 This wide variation in ψ was intriguing. It suggested that substantial differences exist in
 136 the natural immunity of the populations in different nations to the infection. To elucidate
 137 plausible origins of this variation, we next performed meta-regression.

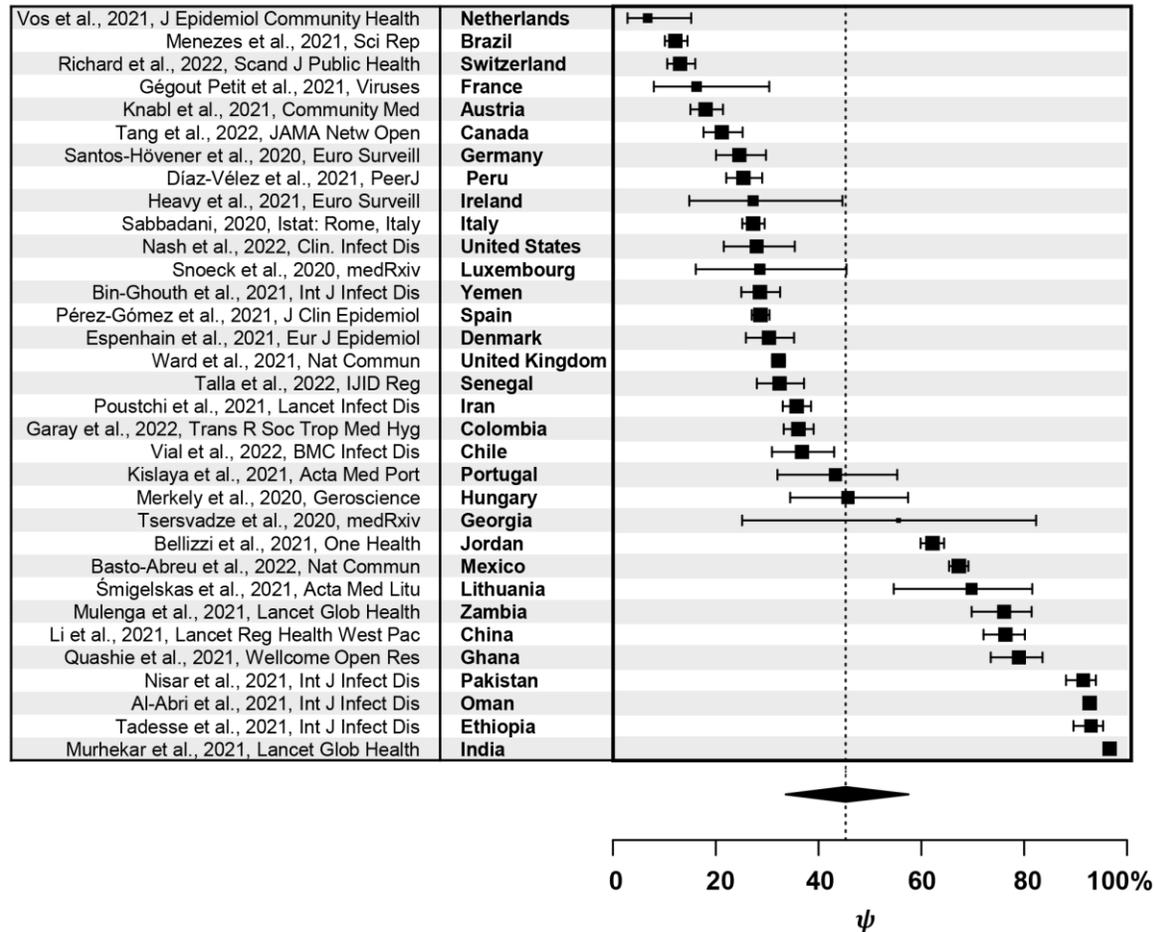


Figure 2: Estimates of ψ and its variation across nations. Black rectangles represent ψ for the 33 nations studied (see Table 1). Error bars are 95% confidence intervals. The sizes of the rectangles are proportional to the weights assigned to the studies for the random-effects modeling (Methods, Table S3). The vertical dashed line is the pooled estimate of ψ (45.3%) with its 95% confidence interval (33.6%-57.5%) shown by the diamond below.

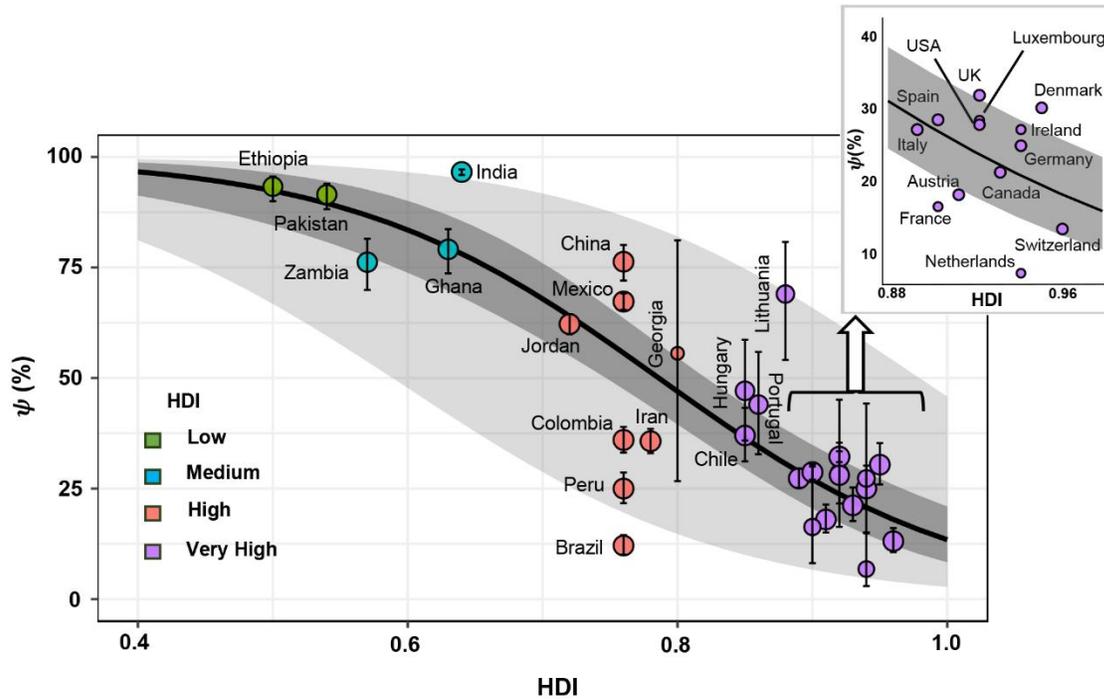
138 Meta-regression and the human development index

139 Key clinical and demographic correlates implicated in severe COVID-19 infections include
 140 age, comorbidities, and prior exposure to related pathogens⁵⁵. While similar factors are
 141 implicated in asymptomatic infections¹², population level predictors of ψ are yet to be

142 identified. Here, we therefore assessed whether the above predictors of severe infections
143 could also predict the variation in ψ . A population level metric of the prevalence of
144 comorbidities is lacking. The largest cause of mortality in adults across the globe is
145 cardiovascular disease (CVD)⁵⁶. We therefore used the prevalence, or rate, of CVD per
146 10^5 individuals (CVDR) as a proxy for the prevalence of comorbidities. While severe
147 infections tend to increase with age, asymptomatic infections within a nation seemed to
148 be much less sensitive to age, barring children and the elderly^{23,24,49,51} (Fig. S1). Besides,
149 intriguingly, some studies have observed a higher prevalence of asymptomatic infections
150 in the elderly^{24,51}. We therefore considered the demographic median age (DMA) of each
151 nation as a metric representative of the ages of the respective national populations. Prior
152 exposure to related pathogens is difficult to quantify at the population level. We reasoned,
153 following recent evidence, that the overall prevalence of infections may be a good
154 indicator of the exposure to related pathogens, including circulating coronaviruses, and
155 hence of the pre-existing immunity to SARS-CoV-2¹². The overall prevalence of infections
156 is often dependent on the development status of a nation, as observed, for instance, with
157 lower respiratory tract infections (LRIs) over the last several decades⁵⁷. LRIs were more
158 prevalent in less developed nations⁵⁷. The human development index (HDI) is a widely
159 used marker of the socio-economic developmental status of nations²¹. The higher the
160 HDI, the more developed is a nation. We therefore considered HDI as a proxy for prior
161 exposure to related pathogens. We collected national-level estimates of these predictors
162 for the nations included in our analysis (Table S5).

163 Before performing meta-regression, we assessed whether the predictors, CVDR, DMA,
164 and HDI, were correlated. We found that the predictors were all highly correlated (Fig.
165 S2). For regression, we therefore constructed models with the predictors chosen one at
166 a time (Methods). To identify influential observations, including which in the regression
167 would substantially adversely affect the model performance, we performed a detailed
168 influential case diagnosis for each model (Methods, Fig. S3). This involved performing
169 meta-regression with each model by leaving out one observation at a time and evaluating
170 the model performance. For the model with CVDR as predictor, we found that ψ for
171 Lithuania was an influential observation. For the model with DMA, ψ for Senegal and
172 Yemen were influential observations. For the model with HDI, the influential observations
173 were ψ for Senegal, Yemen, and Oman. For each model, we therefore excluded the
174 respective influential observations and performed meta-regression using logit-
175 transformed ψ .

176 The model with HDI as the predictor emerged as the best model (Table S6). The best-fit
177 showed that HDI was strongly negatively correlated with ψ ($p = 4 \times 10^{-13}$; Fig. 3).
178 Examining nations based on the United Nations guidelines for HDI-based categorization
179 of developmental status⁵⁸, we found that nations with low (<0.55) and medium (0.55–
180 0.699) HDI had a much higher ψ compared to their high (0.7–0.799) and very high (0.8–
181 1) HDI counterparts (Fig. 3). Remarkably, HDI explained nearly 65.5% of the variation in
182 ψ observed. Models with DMA or CVDR as predictors explained much less of the
183 variation, ~42.4% and ~24.8%, respectively (Table S6; Fig. S4).



184 **Figure 3: HDI is a robust predictor of ψ .** Best-fit meta-regression line (solid line)
185 showing the negative correlation between ψ and HDI, overlaid on the estimates of ψ
186 from 30 nations (symbols) (omitting the 3 influential observations). The dark ribbon is the
187 95% confidence interval and the light ribbon the 95% prediction interval. The symbol sizes
188 are proportional to the weights of the studies in the meta-regression (see Methods). The
189 symbols are color-coded according to the HDI categories indicated. Error bars are the
190 same as in Fig. 2. Regression coefficients and statistics are in Table S6.

191 We ruled out the influence of potential confounding factors. First, although the studies
192 used antibody tests with high specificities, they had varying sensitivities (Table S7). We
193 found that ψ was uncorrelated with the sensitivity of the tests (Fig. S5), ruling out any
194 confounding effect from the assay variations. The studies also used distinct symptom sets
195 in the questionnaires (Fig. S6). A study using fewer symptoms may over-estimate
196 asymptomatic cases. We found that ψ was uncorrelated with the number of symptoms
197 used (Fig. S7). We noticed that most nations used fever and cough as part of their
198 symptom set. It is rare that other symptoms are experienced without also experiencing
199 fever or cough⁵⁹. Consequently, the other symptoms may not significantly increase the
200 chances of detecting symptomatic cases, explaining the lack of a correlation between ψ
201 and the number of symptoms. Finally, we assessed whether distinct age distributions in
202 the nations included could confound our inferences. In particular, the presence of children
203 or the elderly could skew estimates of ψ . Although several studies included children,
204 children accounted for only a small fraction of the seropositive cases (Table S8). Their

205 presence would thus not affect estimates of ψ substantially. The greater prevalence of
206 the elderly in more developed nations could confound the association of ψ with HDI. We
207 therefore performed meta-regression with the percentage of the population aged >65
208 years in each nation as the predictor. The predictor explained only 37% of the variation
209 in ψ observed (Fig. S8, Table S9), rendering it a much poorer predictor than HDI (and
210 median age). HDI was thus a robust predictor of ψ .

211 DISCUSSION

212 Remarkable differences have been observed in the burden of COVID-19 across
213 nations^{8,9}. The differences arise not only from differences in health infrastructure, non-
214 pharmaceutical interventions, and demographics across nations, but also from potential
215 natural variations in immunity. While the former factors have been widely studied^{8,9},
216 natural variations have remained elusive. Here, we recognized that the prevalence of
217 asymptomatic infections, ψ , serves as a reliable measure of the natural immunity of a
218 population to SARS-CoV-2 and carefully collated and assessed its variation across
219 nations. Our findings revealed the enormous variation in the natural immunity to SARS-
220 CoV-2 across nations.

221 An intriguing question that follows is what determines this variation. Studies have
222 attempted to assess whether factors responsible for severe infections^{8,9,60} also explain
223 asymptomatic infections⁶¹. Genetic factors too have been assessed¹³. However, no
224 studies so far have considered metrics such as ψ quantified across nations. The origins
225 of the variation at the population level have therefore remained largely unknown⁶¹. Here,
226 we found that HDI was the best of the predictors of ψ we studied, explaining over 65% of
227 its variation. We contrast this with the genetic polymorphisms identified, which leave over
228 90% of the asymptomatic cases unexplained¹³.

229 HDI is a marker of the overall development of a nation, determined as a composite of
230 three measures, namely, life-expectancy at birth, years of schooling, and per capita
231 income²¹. That ψ was negatively correlated with HDI implied that more developed nations
232 experienced fewer asymptomatic infections on average and hence, presumably, had
233 weaker natural immunity. Because HDI is a composite metric, a precise, causal
234 relationship between HDI and natural immunity is difficult to establish. In more developed
235 nations, due to long-term lifestyle changes, including the practice of hygiene and the use
236 of antibiotics, the prevalence of infections in general is lower^{57,62}. We speculate that this
237 may be responsible for the weaker immunity to SARS-CoV-2, possibly due to altered
238 microbiota⁶³ and/or weaker trained immunity⁶⁴. In addition, the associated lower prior
239 exposure to circulating coronaviruses may limit asymptomatic infections⁶⁵.

240 We do not rule out the role of other factors associated with severe infections, such as age
241 and comorbidities, in establishing asymptomatic infections. Indeed, in our meta-
242 regression, median age (DMA) and the prevalence of cardiovascular disease (CVDR)
243 explained ~42% and ~25% of the variation in ψ . However, because these predictors were

244 strongly correlated with HDI, their specific contributions, over those attributable to their
245 correlation with HDI, remain to be delineated.

246 Severe infections increase significantly with age^{8,9}. Asymptomatic infections may thus be
247 expected, conversely, to decrease with age. We found, surprisingly, from age-stratified
248 data from four nations that ψ within a nation was only weakly dependent on age (Fig. S1).
249 Furthermore, in the UK and Canada, ψ increased with age (Fig. S1). An explanation of
250 these counter-intuitive trends is lacking¹. Nonetheless, while age-stratified data for the
251 other nations we studied was not available, the data above suggests a limited role for age
252 compared to HDI in explaining the variation in ψ .

253 We recognize that deviations in ψ from the mean trends predicted by HDI exist. For
254 instance, China and Brazil had identical HDIs but widely different estimates of ψ (Fig. 3).
255 Factors currently unknown that could explain the remaining ~35% of the variation in ψ not
256 explained by HDI may help reconcile these deviations.

257 Our finding that HDI is the predominant factor predicting ψ may inform ongoing studies
258 aimed at elucidating the mechanistic origins of asymptomatic infections. We recall that a
259 key outcome of COVID-19 vaccination has been to render potentially symptomatic
260 infections asymptomatic, which most approved vaccines achieved with high
261 efficacies^{66,67}. Vaccination studies have offered correlates of protection, such as the
262 antibody titre, that may help predict the likelihood of an infection remaining
263 asymptomatic^{66,67}. Interestingly, the genetic polymorphisms identified recently too point
264 to pre-existing immunity, specifically T-cell immunity, possibly triggered by pre-pandemic
265 exposure to circulating coronaviruses, as associated with asymptomatic infections¹³.
266 Future studies may assess how differences in HDI may lead to differences in such
267 correlates of protection across nations, leading to a better understanding of the natural
268 immunity underlying asymptomatic infections and the variation in ψ .

269 Our study also has implications for understanding the spread of the pandemic and for
270 intervention strategies. Asymptomatic infections have been estimated to have been
271 responsible for nearly a quarter of all transmission events⁴. Transmission by
272 asymptomatic individuals is hard to contain because the individuals are not readily
273 identified. Modeling studies have therefore used estimates of ψ to forecast the course of
274 the pandemic and evaluate interventions^{10,11}. Our findings may enable such models to
275 make more accurate predictions. The larger ψ in less developed nations, for instance,
276 would imply that in these nations, especially in the more populous ones like India, a much
277 larger share of the transmission events may have been due to asymptomatic individuals,
278 exacerbating the difficulty in containing the pandemic. This may have led, compounded
279 by the lack of adequate healthcare access, to the increased mortality in these nations⁸.
280 Quantifying these differences, especially in the context of emerging variants, would be
281 important for future pandemic preparedness.

282 **ONLINE METHODS**

283 **Systematic Review**

284 ***Search Strategy***

285 We systematically reviewed studies on seroprevalence, focusing on the proportion of
286 asymptomatic SARS-CoV-2 infections (ψ). This review adhered to the Preferred
287 Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines¹⁸. We
288 followed a primary search strategy to extract serosurvey studies from the open access
289 SeroTracker⁶⁸ database platform. Our secondary search involved comprehensive
290 screening of COVID-19 and SARS-CoV-2 compendiums from PubMed, SCOPUS and
291 MedRxiv without language restriction. We employed the terms 'COVID-19,' 'SARS-Cov-
292 2,' 'seroprevalence,' 'serological test,' and 'asymptomatic' as search keywords when
293 querying the databases. We attempted to mitigate possible publication bias by assessing
294 both published articles and unpublished literature such as grey literature, preprints,
295 institutional reports. We omitted news and media reports. All but four source articles from
296 where we extracted the data for our meta-regression were published in peer-reviewed
297 journals (Table 1). Of the studies not peer-reviewed, three were preprints (the articles on
298 Georgia³¹, Ghana³³, and Luxembourg⁴¹), and one was a report from governmental
299 sources (Italy³⁸). We included studies conducted between March 11, 2020, when WHO
300 declared the COVID-19 pandemic, and December 8, 2020, when the first vaccine was
301 administered outside clinical trials⁶⁹. We additionally included the studies from USA⁵²,
302 Canada²⁴, and Jordan³⁹ because, although they exceeded the above timeframes, either
303 no vaccination had begun in these nations till the end of the study period or the vaccinated
304 individuals were omitted from the reported serosurvey data (see Text S2).

305 This review was not registered. Two researchers (AT and SC) reviewed the title,
306 abstracts, and full texts of the articles independently. Disagreements were resolved by
307 three other reviewers (AJ, BC, and NMD). The researchers conducted initial searches
308 from November 22, 2022, to December 13, 2022, and continued to perform monthly
309 searches thereafter until September 20, 2023. Data collection did not involve any
310 automated processes.

311 ***Selection Criteria***

312 Asymptomatic individuals are those who become infected but never develop or perceive
313 any symptoms during the entire course of infection¹⁴. We chose retrospective cross-
314 sectional serosurveys that determined the proportion of asymptomatic infection through
315 interviews and questionnaires, collecting data on symptoms reported either at the time of
316 blood sample collection or during a prior period. The study window chosen (see above),
317 which lasted under a year of the onset of the epidemic in the respective settings,
318 minimized the chances of seroreversion¹⁴.

319 While screening available studies, we identified those that provided accurate estimates
320 of asymptomatic cases and were conducted on participants representing the national

321 populace, by methods such as household or community sampling. We ensured a sample
322 size $n > 500$. When investigations of such scale were not available for a nation, we
323 included studies that reported the seroprevalence in provinces^{26,45}, counties³⁷, regions⁴²,
324 governates³⁹, major cities and towns^{20,29,32,36,40}, districts⁵⁴, or municipalities^{27,30,35}, with
325 populations that were representative of the general population of the nation.

326 Subsequently, the complete texts of the screened studies underwent evaluation (Fig. 1,
327 Text S1). Serosurveys conducted on vaccinated individuals or after the emergence of
328 variants were excluded. Studies that lacked reporting of the total number of individuals
329 tested, details on seroprevalence, or the number/percentage of asymptomatic individuals
330 among the seropositive participants were also excluded. Studies that lacked details of
331 serology tests were excluded. Further, we excluded review articles and findings sourced
332 from news and media outlets. Finally, studies that did not ensure random sampling of
333 their participants were excluded. Collated in this manner, we ensured inclusion of the
334 most reliable evidence coming from large-scale cross-sectional seroprevalence studies
335 with representative samples that included data from antibody-based testing and identified
336 population-level asymptomatic proportions in a nation.

337 ***Data Extraction and Quality Assessment***

338 From each study, data of nation, sample frame (group of interest), study duration
339 including sampling start and end date, geographical scope of the study (national, regional,
340 local, sub-local), age of participants, number of tested individuals, number of confirmed
341 COVID-19 infections, number of asymptomatic infections, comorbid conditions,
342 participation rate, type of serological assay, test kit name and details, sensitivity and
343 specificity of the serological assay, number of symptoms recorded in the questionnaire,
344 source type, first author, and date of publication were extracted. Two reviewers (AT and
345 SC) performed the quality assessment independently. Disagreements were resolved by
346 three other reviewers (AJ, BC, and NMD). Because the outcome of our interest was the
347 proportion of asymptomatic infections and not seroprevalence, we adapted the Joanna
348 Briggs Institute (JBI) checklist²² to develop our own quality assessment tool and grading
349 scale (Text S3). Our quality assessment tool assessed the domains of symptom
350 assessment, reliability of the serological assays and questionnaires, and the sampling
351 method to ensure representativeness at the national level (Table S2). The studies were
352 classified as low, moderate, or high quality following the grading scales developed a priori
353 (Text S3). Neither the data extraction nor the quality assessment involved any automated
354 process.

355 ***Data on Predictors***

356 We collated national-level data on demographic median age, population age distribution,
357 prevalence of cardiovascular disease, and human development index (Table S5). All the
358 data were for 2019.

359 **Statistical analysis and meta-regression**

360 We used ψ as the outcome measure. We calculated confidence intervals on ψ using the
361 Wilson score interval⁷⁰. For computing the pooled estimate of ψ , we logit transformed the
362 reported estimates, $\Psi = \log\left(\frac{\psi}{100-\psi}\right)$, and assigned individual weights to each study using
363 the inverse variance method⁷¹. We used the Paule and Mandel method⁷² implemented in
364 the Metafor⁷³ package in R for parameter estimation. To quantify the heterogeneity, we
365 evaluated the I^2 metric⁷⁴. We calculated the Spearman's correlation coefficient to
366 evaluate the correlations between the predictors, CVDR, DMA, and HDI. We performed
367 meta-regression using the predictors one at a time. To identify outliers and influential
368 observations, we estimated studentized residuals and performed meta-regression by
369 leaving out one observation at a time. An observation was classified as influential if the
370 absolute values of the studentized residuals for the observation were larger than 3 and/or
371 leaving it out resulted in an increase in R^2 greater than 5 percentage points. We compared
372 the resulting models using R^2 . We performed additional tests to ascertain the robustness
373 to potential confounding factors (see Results).

374 **ACKNOWLEDGEMENTS**

375 We thank Shreyas Joshi for insightful discussions. BC was supported by the C. V. Raman
376 postdoctoral fellowship of the Indian Institute of Science.

377 **REFERENCES**

- 378 1. Boyton, R.J. & Altmann, D.M. The immunology of asymptomatic SARS-CoV-2
379 infection: what are the key questions? *Nat Rev Immunol* **21**, 762-768 (2021).
- 380 2. Lamers, M.M. & Haagmans, B.L. SARS-CoV-2 pathogenesis. *Nat Rev Microbiol*
381 **20**, 270-284 (2022).
- 382 3. Buitrago-Garcia, D., *et al.* Occurrence and transmission potential of asymptomatic
383 and presymptomatic SARS-CoV-2 infections: a living systematic review and meta-
384 analysis. *PLoS Med* **17**, e1003346 (2020).
- 385 4. Johansson, M.A., *et al.* SARS-CoV-2 transmission from people without COVID-19
386 symptoms. *JAMA Netw Open* **4**, e2035057 (2021).
- 387 5. Oran, D.P. & Topol, E.J. The proportion of SARS-CoV-2 infections that are
388 asymptomatic: a systematic review. *Ann Intern Med* **174**, 655-662 (2021).
- 389 6. Ma, Q., *et al.* Global percentage of asymptomatic SARS-CoV-2 infections among
390 the tested population and individuals with confirmed COVID-19 diagnosis: a
391 systematic review and meta-analysis. *JAMA Netw Open* **4**, e2137257 (2021).
- 392 7. Sah, P., *et al.* Asymptomatic SARS-CoV-2 infection: a systematic review and
393 meta-analysis. *Proc Natl Acad Sci USA* **118**, e2109229118 (2021).
- 394 8. Levin, A.T., *et al.* Assessing the burden of COVID-19 in developing countries:
395 systematic review, meta-analysis and public policy implications. *BMJ Glob Health*
396 **7**, e008477 (2022).

- 397 9. Sorensen, R., *et al.* Variation in the COVID-19 infection-fatality ratio by age, time,
398 and geography during the pre-vaccine era: a systematic analysis. *Lancet* **399**,
399 1469-1488 (2022).
- 400 10. Bertozzi, A.L., Franco, E., Mohler, G., Short, M.B. & Sledge, D. The challenges of
401 modeling and forecasting the spread of COVID-19. *Proc Natl Acad Sci USA* **117**,
402 16732-16738 (2020).
- 403 11. Reiner, R., Collins, J.K., Team, C.-F. & Murray, C.J. Forecasting the trajectory of
404 the COVID-19 pandemic into 2023 under plausible variant and intervention
405 scenarios: a global modelling study. *medRxiv*, 2023.03 (2023).
- 406 12. Murray, S.M., *et al.* The impact of pre-existing cross-reactive immunity on SARS-
407 CoV-2 infection and vaccine responses. *Nat Rev Immunol* **23**, 304-316 (2023).
- 408 13. Augusto, D.G., *et al.* A common allele of HLA is associated with asymptomatic
409 SARS-CoV-2 infection. *Nature* **620**, 128–136 (2023).
- 410 14. Meyerowitz, E.A., Richterman, A., Bogoch, I.I., Low, N. & Cevik, M. Towards an
411 accurate and systematic characterisation of persistently asymptomatic infection
412 with SARS-CoV-2. *Lancet Infect Dis* **21**, e163-e169 (2021).
- 413 15. Accorsi, E.K., *et al.* How to detect and reduce potential sources of biases in studies
414 of SARS-CoV-2 and COVID-19. *Eur J Epidemiol* **36**, 179-196 (2021).
- 415 16. Chan, J., *et al.* COVID-19 in the New York city jail system: epidemiology and health
416 care response, March–April 2020. *Public Health Rep* **136**, 375-383 (2021).
- 417 17. Rivera, F., *et al.* Prevalence of severe acute respiratory syndrome coronavirus 2
418 (SARS-CoV-2) asymptomatic infections in 2 large academic health systems in
419 Wisconsin. *Clin Infect Dis* **73**, e3974-e3976 (2020).
- 420 18. Page, M.J., *et al.* The PRISMA 2020 statement: an updated guideline for reporting
421 systematic reviews. *BMJ* **372**, n71 (2021).
- 422 19. Knabl, L., *et al.* High SARS-CoV-2 seroprevalence in children and adults in the
423 Austrian ski resort of Ischgl. *Commun Med* **1**, 4 (2021).
- 424 20. Ladage, D., *et al.* SARS-CoV-2-specific antibody prevalence and symptoms in a
425 local Austrian population. *Front Med* **8**, 726 (2021).
- 426 21. Anand, S. & Sen, A. Human Development Index: methodology and measurement.
427 *Human Development Report Office Occasional Paper No. 12, UNDP, New York*
428 (1994).
- 429 22. Munn, Z., Moola, S., Lisy, K., Riitano, D. & Tufanaru, C. Methodological guidance
430 for systematic reviews of observational epidemiological studies reporting
431 prevalence and cumulative incidence data. *JBI Evid Implement* **13**, 147-153
432 (2015).
- 433 23. Menezes, A.M.B., *et al.* High prevalence of symptoms among Brazilian subjects
434 with antibodies against SARS-CoV-2. *Sci Rep* **11**, 13279 (2021).
- 435 24. Tang, X., *et al.* Assessment of SARS-CoV-2 seropositivity during the first and
436 second viral waves in 2020 and 2021 among Canadian adults. *JAMA Netw Open*
437 **5**, e2146798 (2022).
- 438 25. Vial, P., *et al.* Seroprevalence, spatial distribution, and social determinants of
439 SARS-CoV-2 in three urban centers of Chile. *BMC Infect Dis* **22**, 99 (2022).
- 440 26. Li, Z., *et al.* Antibody seroprevalence in the epicenter Wuhan, Hubei, and six
441 selected provinces after containment of the first epidemic wave of COVID-19 in
442 China. *Lancet Reg Health West Pac* **8**, 100094 (2021).

- 443 27. Garay, E., *et al.* SARS-CoV-2 in eight municipalities of the Colombian tropics: high
444 immunity, clinical and sociodemographic outcomes. *Trans R Soc Trop Med Hyg*
445 **116**, 139-147 (2022).
- 446 28. Espenhain, L., *et al.* Prevalence of SARS-CoV-2 antibodies in Denmark:
447 nationwide, population-based seroepidemiological study. *Eur J Epidemiol* **36**, 715-
448 725 (2021).
- 449 29. Tadesse, E.B., *et al.* Seroprevalence and risk factors for SARS-CoV-2 infection in
450 selected urban areas in Ethiopia: a cross-sectional evaluation during July 2020. *Int*
451 *J Infect Dis* **111**, 179-185 (2021).
- 452 30. Gégout Petit, A., *et al.* Seroprevalence of SARS-CoV-2, symptom profiles and
453 sero-neutralization in a suburban Area, France. *Viruses* **13**, 1076 (2021).
- 454 31. Tsertsvadze, T., *et al.* SARS-CoV-2 antibody seroprevalence in Tbilisi, the capital
455 city of country of Georgia. *medRxiv*, 2020.09 (2020).
- 456 32. Santos-Hövenner, C., *et al.* Serology-and PCR-based cumulative incidence of
457 SARS-CoV-2 infection in adults in a successfully contained early hotspot (CoMoLo
458 study), Germany, May to June 2020. *Euro Surveill* **25**, 2001752 (2020).
- 459 33. Quashie, P.K., *et al.* Trends of severe acute respiratory syndrome coronavirus 2
460 (SARS-CoV-2) antibody prevalence in selected regions across Ghana. *Wellcome*
461 *Open Res* **6**, 173 (2021).
- 462 34. Merkely, B., *et al.* Novel coronavirus epidemic in the Hungarian population, a
463 cross-sectional nationwide survey to support the exit policy in Hungary.
464 *Geroscience* **42**, 1063-1074 (2020).
- 465 35. Murhekar, M.V., *et al.* SARS-CoV-2 antibody seroprevalence in India, August-
466 September, 2020: findings from the second nationwide household serosurvey.
467 *Lancet Glob Health* **9**, e257-e266 (2021).
- 468 36. Poustchi, H., *et al.* SARS-CoV-2 antibody seroprevalence in the general population
469 and high-risk occupational groups across 18 cities in Iran: a population-based
470 cross-sectional study. *Lancet Infect Dis* **21**, 473-481 (2021).
- 471 37. Heavey, L., *et al.* The study to investigate COVID-19 infection in people living in
472 Ireland (SCOPI): a seroprevalence study, June to July 2020. *Euro Surveill* **26**,
473 2001741 (2021).
- 474 38. Sabbadini, L.L. & Istat, D.C. Primi risultati dell'indagine di sieroprevalenza SARS-
475 CoV-2. (2020).
- 476 39. Bellizzi, S., *et al.* A three-phase population based sero-epidemiological study:
477 assessing the trend in prevalence of SARS-CoV-2 during COVID-19 pandemic in
478 Jordan. *One Health* **13**, 100292 (2021).
- 479 40. Šmigelskas, K., *et al.* SARS-CoV-2 seroprevalence in Lithuania: results of national
480 population survey. *Acta Med Litu* **28**, 48 (2021).
- 481 41. Snoeck, C.J., *et al.* Prevalence of SARS-CoV-2 infection in the Luxembourgish
482 population—the CON-VINCE study. *medRxiv*, 2020.05 (2020).
- 483 42. Basto-Abreu, A., *et al.* Nationally representative SARS-CoV-2 antibody prevalence
484 estimates after the first epidemic wave in Mexico. *Nat Commun* **13**, 589 (2022).
- 485 43. Vos, E.R., *et al.* Nationwide seroprevalence of SARS-CoV-2 and identification of
486 risk factors in the general population of the Netherlands during the first epidemic
487 wave. *J Epidemiol Community Health* **75**, 489-495 (2021).

- 488 44. Al-Abri, S.S., *et al.* Seroprevalence of SARS-CoV-2 antibodies in the general
489 population of Oman: results from four successive nationwide sero-epidemiological
490 surveys. *Int J Infect Dis* **112**, 269-277 (2021).
- 491 45. Nisar, M.I., *et al.* Serial population-based serosurveys for COVID-19 in two
492 neighbourhoods of Karachi, Pakistan. *Int J Infect Dis* **106**, 176-182 (2021).
- 493 46. Diaz-Velez, C., *et al.* SARS-CoV-2 seroprevalence study in Lambayeque, Peru.
494 June-July 2020. *PeerJ* **9**, e11210 (2021).
- 495 47. Kislaya, I., *et al.* Seroprevalence of SARS-CoV-2 infection in Portugal in May-July
496 2020: results of the first national serological survey (ISNCOVID-19). *Acta Med Port*
497 **34**, 87-94 (2021).
- 498 48. Talla, C., *et al.* Seroprevalence of anti-SARS-CoV-2 antibodies in Senegal: a
499 national population-based cross-sectional survey, between October and
500 November 2020. *IJID Reg* **3**, 117-125 (2022).
- 501 49. Perez-Gomez, B., *et al.* ENE-COVID nationwide serosurvey served to characterize
502 asymptomatic infections and to develop a symptom-based risk score to predict
503 COVID-19. *J Clin Epidemiol* **139**, 240-254 (2021).
- 504 50. Richard, A., *et al.* Seroprevalence of anti-SARS-CoV-2 IgG antibodies, risk factors
505 for infection and associated symptoms in Geneva, Switzerland: a population-
506 based study. *Scand J Public Health* **50**, 124-135 (2022).
- 507 51. Ward, H., *et al.* SARS-CoV-2 antibody prevalence in England following the first
508 peak of the pandemic. *Nat Commun* **12**, 905 (2021).
- 509 52. Nash, D., *et al.* Severe acute respiratory syndrome coronavirus 2 incidence and
510 risk factors in a national, community-based prospective cohort of US adults. *Clin*
511 *Infect Dis* **76**, e375-e384 (2022).
- 512 53. Bin-Ghouth, A.S., *et al.* SARS-CoV-2 seroprevalence in Aden, Yemen: a
513 population-based study. *Int J Infect Dis* **115**, 239-244 (2022).
- 514 54. Mulenga, L.B., *et al.* Prevalence of SARS-CoV-2 in six districts in Zambia in July,
515 2020: a cross-sectional cluster sample survey. *Lancet Glob Health* **9**, e773-e781
516 (2021).
- 517 55. Brodin, P. Immune determinants of COVID-19 disease presentation and severity.
518 *Nat Med* **27**, 28-33 (2021).
- 519 56. Ritchie, H., Spooner, F. & Roser, M. Causes of death. *Our World In Data* (2018).
- 520 57. Troeger, C., *et al.* Estimates of the global, regional, and national morbidity,
521 mortality, and aetiologies of lower respiratory infections in 195 countries, 1990–
522 2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet*
523 *Infect Dis* **18**, 1191-1210 (2018).
- 524 58. United Nations. Human development index. (<https://hdr.undp.org/data-center/human-development-index#/indicies/HDI>).
- 525
- 526 59. Alimohamadi, Y., Sepandi, M., Taghdir, M. & Hosamirudsari, H. Determine the
527 most common clinical symptoms in COVID-19 patients: a systematic review and
528 meta-analysis. *J Prev Med Hyg* **61**, E304-E312 (2020).
- 529 60. Chatterjee, B., Singh Sandhu, H. & Dixit, N.M. Modeling recapitulates the
530 heterogeneous outcomes of SARS-CoV-2 infection and quantifies the differences
531 in the innate immune and CD8 T-cell responses between patients experiencing
532 mild and severe symptoms. *PLoS Pathog* **18**, e1010630 (2022).

- 533 61. Sette, A. & Crotty, S. Adaptive immunity to SARS-CoV-2 and COVID-19. *Cell* **184**,
534 861-880 (2021).
- 535 62. Sironi, M. & Clerici, M. The hygiene hypothesis: an evolutionary perspective.
536 *Microbes Infect* **12**, 421-427 (2010).
- 537 63. Finlay, B.B., *et al.* The hygiene hypothesis, the COVID pandemic, and
538 consequences for the human microbiome. *Proc Natl Acad Sci USA* **118**,
539 e2010217118 (2021).
- 540 64. Netea, M.G., *et al.* The role of trained immunity in COVID-19: lessons for the next
541 pandemic. *Cell Host Microbe* **31**, 890-901 (2023).
- 542 65. Sagar, M., *et al.* Recent endemic coronavirus infection is associated with less-
543 severe COVID-19. *J Clin Invest* **131**, e143380 (2021).
- 544 66. Houry, D.S., *et al.* Neutralizing antibody levels are highly predictive of immune
545 protection from symptomatic SARS-CoV-2 infection. *Nat Med* **27**, 1205-1211
546 (2021).
- 547 67. Padmanabhan, P., Desikan, R. & Dixit, N.M. Modeling how antibody responses
548 may determine the efficacy of COVID-19 vaccines. *Nat Comput Sci* **2**, 123-131
549 (2022).
- 550 68. Arora, R.K., *et al.* SeroTracker: a global SARS-CoV-2 seroprevalence dashboard.
551 *Lancet Infect Dis* **21**, e75-e76 (2021).
- 552 69. Carvalho, T., Krammer, F. & Iwasaki, A. The first 12 months of COVID-19: a
553 timeline of immunological insights. *Nat Rev Immunol* **21**, 245-256 (2021).
- 554 70. Wallis, S. Binomial confidence intervals and contingency tests: mathematical
555 fundamentals and the evaluation of alternative methods. *J Quant Linguist* **20**, 178-
556 208 (2013).
- 557 71. Veroniki, A.A., *et al.* Methods to estimate the between-study variance and its
558 uncertainty in meta-analysis. *Res Synth Methods* **7**, 55-79 (2016).
- 559 72. Paule, R.C. & Mandel, J. Consensus values and weighting factors. *J Res Natl Bur*
560 *Stand* **87**, 377-385 (1982).
- 561 73. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *J Stat*
562 *Softw* **36**, 1-48 (2010).
- 563 74. Higgins, J.P. & Thompson, S.G. Quantifying heterogeneity in a meta-analysis. *Stat*
564 *Med* **21**, 1539-1558 (2002).