

BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

COVID-19 morbidity in Afghanistan: a nationwide, population-based seroepidemiological study

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2021-060739
Article Type:	Original research
Date Submitted by the Author:	21-Jan-2022
Complete List of Authors:	Saeedzai, Sayed Ataullah; Ministry of Public Health, Monitoring, Evaluation and Health Information System Sahak, Mohammad; World Health Organization, WHE Arifi, Fatima; WHO Aly, Eman; World Health Organisation Regional Office for the Eastern Mediterranean Van Gurp, Margo; KIT Royal Tropical Institute White, Lisa; University of Oxford, Nuffield Department of Medicine; University of Oxford Chen, Siyu; University of Oxford, Discovery, Nuffield Department of Medicine Barakat, Amal; World Health Organisation Regional Office for the Eastern Mediterranean Azim, Giti; Ministry of Public Health Rasoly, Bahara; Ministry of Public Health Safi, Soraya; Ministry of Public Health Flegg, Jennifer; University of Melbourne Ahmed, Nasar; Florida International University, Department of Epidemiology Ahadi, Mohmmad Jamaluddin; Ministry of Public Health Achakzai, Niaz; Ministry of Public Health AbouZeid, Alaa; Cairo University Kasr Alainy Faculty of Medicine
Keywords:	COVID-19, Epidemiology < INFECTIOUS DISEASES, EPIDEMIOLOGY

SCHOLARONE™
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

1 **COVID-19 morbidity in Afghanistan: a nationwide, population-based**
2 **seroepidemiological study**

3 **Authors & affiliations**

4 Sayed Ataullah Saeedzai¹, Mohammad Nadir Sahak², Fatima Arifi^{2*}, Eman Abdelkareem Aly³,
5 Margo Van Gurp⁴, Lisa White⁵, Siyu Chen⁵, Amal Barkat⁶, Giti Azim¹, Bahara Rasoly¹, Soraya
6 Safi¹, Jennifer Flegg⁷, Nasar U. Ahmed⁸, Mohammad Jamaluddin Ahadi¹, Niaz M. Achakzai⁹,
7 Alaa Abouzeid¹⁰

8 ¹General Directorate of Monitoring, Evaluation and Health Information System (M&E-HIS),
9 Ministry of Public Health, Kabul, Afghanistan

10 ²WHO Health Emergencies Programme, WHO, Kabul, Afghanistan

11 ³ Information Systems for Health Unit, WHO Eastern Mediterranean Regional Office, Cairo, Egypt

12 ⁴ KIT Royal Tropical Institute, Amsterdam, the Netherlands

13 ⁵ Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of
14 Medicine, University of Oxford, United Kingdom

15 ⁶ Infectious Hazard Preparedness Unit, WHO Eastern Mediterranean Regional Office, Cairo, Egypt

16 ⁷ School of Mathematics and Statistics, University of Melbourne, Melbourne, Australia

17 ⁸ Department of Epidemiology, Robert Stempel College of Public Health, Florida International
18 University, Miami, FL

19 ⁹ Department of Molecular Biology, Forensic Medicine Directorate. Ministry of Public Health, Kabul,
20 Afghanistan

21 ¹⁰ Department of Public Health, Faculty of Medicine, Cairo University

22 *Corresponding author:

23 Fatima Arifi (Fatima.arefi@gmail.com, arifif@who.int)

24 **Word count: 4403 words**

25 List of Abbreviations

26 CoMo Consortium – COVID-19 International Modelling Consortium

27 EA – enumeration area

28 ELISA – enzyme-linked immunosorbent assay

29 IFR – infection fatality ratio

30 IgG – immunoglobulin G

31 IgM – immunoglobulin M

32 MCMC – Markov chain Monte Carlo

33 MoPH – Ministry of Public Health

34 NPI – nonpharmaceutical intervention

35 NSIA – National Statistics and Information Authority

36 R0 – basic reproduction number

37 RDT – rapid diagnostic test

38 WHO – World Health Organization

39

40 **Abstract (257 words)**

41 Introduction

42 The ongoing COVID-19 pandemic continues to result in considerable morbidity and mortality around
43 the world. However, in many countries it is difficult to estimate the true burden of COVID-19
44 infection in a population due to gaps in surveillance coverage and limited testing capacity.

45 Methods

46 Here, we describe a population-based, cross-sectional, age-stratified sero-epidemiological study
47 conducted throughout Afghanistan during June/July 2020. Participants were interviewed to

1
2
3 48 complete a questionnaire, and rapid diagnostic tests were used to test for SARS-CoV-2 antibodies.
4
5 49 The primary objectives were 1) to determine the magnitude of COVID-19 infections in the general
6
7 50 population and age-specific cumulative incidence, as determined by seropositivity and clinical
8
9 51 symptoms of COVID-19; and 2) to determine the magnitude of asymptomatic or subclinical
10
11 52 infections. To adjust the seroprevalence for test sensitivity and specificity, as well as seroreversion,
12
13 53 Bernoulli model methodology was used to infer the population exposure in Afghanistan.
14
15
16
17

18 54 Results

19 55 The survey revealed that, to July 2020, around 10 million people in Afghanistan (31.5% of the
20
21 56 population) had either current or previous COVID-19 infection. This implies that the herd immunity
22
23 57 threshold had not been reached and most of the population of Afghanistan remained at risk of
24
25 58 infection. However, the herd immunity threshold may have been crossed in some localities, such as
26
27 59 Kabul province, where more than half of the population had been infected with COVID-19, which
28
29 60 exceeds the lowest reported herd immunity threshold.
30
31
32
33

34 61 Conclusion

35 62 As most of the population remained at risk of infection at the time of the study, any lifting of public
36
37 63 health and social measures needed to be considered gradually.
38
39
40
41

42 64 Article Summary: Strengths and Limitations

47 65 Strengths:

- 51 66 • This is the first nationwide, population-based, large sample size, seroepidemiological study
52 67 conducted in Afghanistan
 - 53 68 • The study provides evidence on the high burden of COVID-19 morbidity in resource limited
54 69 and conflict affected country with limited mitigation measures and limited evidence on
55 70 burden
 - 56 71 • The study highlights the limited surveillance capacity and under-reporting of COVID-19 cases
57 72 in Afghanistan
- 58
59
60

73 Limitations:

- 74 • Due to security concerns, not all areas could be surveyed; the inability to conduct proper
75 household listing and create maps for enumeration areas in those areas where the
76 government lacked control may have affected the findings
- 77 • The findings may not reflect the current situation with regards to the new SARS-CoV-2 delta
78 and omicron variants of concern, as the R0 for these variants is not well-established
- 79 • The data were entered in the DHIS2 database, which created many challenges for data
80 verification, household matching and the subsequent analysis.

81 Summary box

82 What is already known?

- 83 • To 5 November 2021, the COVID-19 pandemic has resulted in more than 248 million cases
84 and more than 5 million deaths worldwide.
- 85 • As in other countries, Afghanistan has introduced nonpharmaceutical interventions to
86 control the spread of COVID-19.
- 87 • Seroepidemiological surveys can provide useful data to help inform public health policies.

88 What are the new findings?

- 89 • This national survey of COVID-19 morbidity and mortality in Afghanistan revealed that
90 around 10 million people (31.5% of the population) had either current or previous COVID-19
91 infection.
- 92 • There was regional heterogeneity in the burden of COVID-19 disease, with urban areas such
93 as Kabul showing higher cumulative rates of COVID-19.

94 What do the new findings imply?

- 95 • The cumulative number of COVID-19 cases across the country means Afghanistan had yet to
96 reach the herd immunity threshold at the time of study, which is reported to range from
97 43% to 85%.
- 98 • The cumulative number of COVID-19 cases in Kabul (53%) suggests that this region appeared
99 to might have reached the herd immunity threshold, if lower estimates for the herd
100 immunity threshold (43%) are used.
- 101 • Seroprevalence represents a low estimate of herd immunity, while predicted exposure
102 represents a higher estimate.

104 INTRODUCTION

105 The COVID-19 pandemic has resulted in more than 248 million confirmed cases and in excess of 5
106 million deaths globally to November 2021.¹ Many countries are continuing to experience epidemic
107 waves of COVID-19, including Brazil, India and Nepal.²⁻⁴ The first reported case of COVID-19 in
108 Afghanistan was in Herat province on 24 February 2020; as of 20 July 2021, Afghanistan has reported
109 156 363 confirmed cases of COVID-19 and 7284 deaths from the disease.⁵

110 When the COVID-19 pandemic began, there were no vaccines or specific treatments available for
111 COVID-19, so nonpharmaceutical interventions (NPIs) were recommended, including social
112 distancing, home quarantine, closure of schools and universities, and bans on public gatherings.
113 Afghanistan introduced NPIs as soon as the first case of COVID-19 was detected in the country. Case
114 detection and isolation were seen as key features in helping to reduce the spread of COVID-19. With
115 the recent political transition in the country and disruption of the health system, public health and
116 social measures to tackle COVID-19 have been completely neglected, which may pose a major risk
117 for increasing the spread of COVID-19 in Afghanistan.

118 The initial focus of the Afghanistan Ministry of Public Health (MoPH) was on patients with severe
119 COVID-19 disease and ways to decrease mortality associated with the disease. Serological testing of
120 patients can be used to provide useful information about an individual's status in terms of a current
121 or previous COVID-19 infection. Immunoglobulin M (IgM) and G (IgG) antibodies arise at around the
122 same time, between 1 to 3 weeks after infection; however, IgM antibodies decay more rapidly than
123 IgG antibodies.⁶ Therefore, for public health studies, IgM is used as a marker of current infection
124 while IgG is used as a marker of previous infection, i.e. within the previous few months. There are
125 various rapid diagnostic tests (RDTs) available that can be used to simultaneously test blood samples
126 for IgM/IgG antibodies against COVID-19.

1
2
3 127 Due to the limited testing and surveillance capacity in Afghanistan, it seemed likely there was
4
5 128 considerable under-reporting of cases and deaths; therefore, robust scientific studies are required to
6
7 129 determine the actual burden of COVID-19 in the country. Serological studies can be used to estimate
8
9 130 levels of past exposure and thus position a population in their epidemic timeline. However, serology
10
11 131 results might underestimate the total exposure in a population⁷ because of decaying antibody titres
12
13 132 over time.⁸⁻¹⁰ Here, we describe a national seroepidemiological survey initiated by the MoPH and
14
15 133 conducted throughout Afghanistan between June and July 2020, involving a questionnaire survey
16
17 134 and antibody testing of participants for COVID-19 infection using RDTs. The primary objectives of the
18
19 135 study were: 1) to determine the magnitude of COVID-19 infection in the general population and age-
20
21 136 specific cumulative incidence, as determined by seropositivity and clinical symptoms of COVID-19;
22
23 137 and 2) to determine the magnitude of asymptomatic or subclinical infections. The World Health
24
25 138 Organization (WHO) protocol for population-based age-stratified seroepidemiological investigations
26
27 139 for COVID-19 infection was adapted for the Afghanistan context to obtain seroprevalence
28
29 140 estimates.¹¹ To adjust the seroprevalence for test sensitivity and specificity, as well as seroreversion,
30
31 141 we further adapted a methodology¹² that was originally developed for the England setting and used
32
33 142 this to infer the population exposure and undocumented mortality associated with COVID-19 in
34
35 143 Afghanistan.
36
37
38
39
40
41
42

143 **METHODS**

144 **Patient and Public Involvement statement**

145
146 As the study was not clinical trial and it did not involve patients, the public and patients were not
147
148 involved directly during study design or dissemination. The study results were disseminated through
149
150 public workshops in universities, seminars and workshops, and media for the general public
information. Their consent was obtained for being included in the study and any personal identifier
information was excluded during study design and results.

151 Ethical considerations

152 Ethical and technical clearance to conduct the survey was obtained from the Institutional Review
153 Board of the Afghanistan MoPH with **Reference #: A.0321.0278**. Informed consent was obtained
154 from participants aged ≥ 18 years, and assent from family members was obtained for those aged 5–
155 17 years. Individuals who did not provide consent were excluded. Survey team members provided
156 advice about home isolation to participants who tested IgM-positive for COVID-19 during the survey.

157 Study design

158 This was a population-based, cross-sectional, age-stratified seroepidemiological study. Participants
159 were interviewed to complete a questionnaire, and RDTs were used to test for SARS-CoV-2
160 antibodies. The survey was conducted during June and July 2020.

161 Population and sampling

162 This was a national study conducted in the eight regions of Afghanistan plus Kabul province, which
163 was considered as a separate region, making nine regions in total (**online supplemental figure S1**).
164 The total sample size was 9514 and the number of participants required in each region was
165 estimated proportionate to the population size of each region (**online supplemental Table S1**). Two-
166 stage cluster sampling was used. In the first stage, an updated list of enumeration areas (EAs) was
167 used as the study sampling frame, with 31 to 44 EAs (clusters) randomly selected per region,
168 resulting in a total of 360 clusters. Due to time constraints and to ensure data validity, insecure or
169 inaccessible EAs were excluded from the study.
170 In the second stage, all households in an EA were listed and 16 households per EA were selected
171 using a random sampling table). For the age-stratification, two individuals from each household
172 were randomly selected for testing: one aged 5–17 years and one aged ≥ 18 years.

173 Serological testing

174 Finger-prick blood samples were collected from the randomly selected household members in each
175 age category. The antibody RDTs for COVID-19 were performed in the presence of the participant,
176 and the results were shared with them. The COVID-19 RDT used was the COVID-19 IgG/IgM Rapid
177 Test Cassette developed by Healgen Scientific LLC, USA. The RDT is US Food and Drug Administration
178 (FDA)-authorised, with IgM relative sensitivity and specificity of 95.7% and 97.3%, respectively; IgG
179 relative sensitivity and specificity of 91.8% and 96.4%, respectively; and both IgG-positive and/or
180 IgM-positive specificity of 97.5%.

181 Data collection and analysis

182 The survey used a validated questionnaire that was initially piloted in Kabul province. All participants
183 were interviewed by the survey team members, who completed a questionnaire that included
184 questions about the demographics of each participant and their household members, their history of
185 exposure to COVID-19, and deaths in the family during the 15-month period beginning in March
186 2019.

187 Data collection teams comprised two members, one male and one female; there were 191 teams in
188 total. Due to the need for blood-drawing for samples, the team members were either nurses,
189 midwives or laboratory technicians.

190 Regional COVID-19 data were entered into DHIS2 (District Health Information Software-2) by disease
191 surveillance officers in the provinces. DHIS2 is the national data warehouse for Afghanistan's health
192 information and includes data that inform the country's COVID-19 dashboard.⁵ Various steps were
193 taken for data quality assurance at both regional and central levels within the MoPH; data collection
194 teams were monitored by the master trainers in the regional capitals and by disease surveillance
195 staff in the provinces. Prior to being entered into the system, questionnaires were quality checked

1
2
3 196 and some participants whose phone numbers were available in the questionnaire were contacted at
4
5 197 random by phone call to confirm that their details were correct.
6
7

8 198 Data were imported into STATA version 15¹³ for the statistical analyses. To ensure a representative
9
10 199 sample and results, weighted analysis was applied to adjust for the complex survey design. Sample
11
12 200 weighting, non-response weighting and post-stratification weighting were performed. The
13
14 201 proportions of infections and 95% confidence intervals were calculated and adjusted to take the
15
16 202 survey design into account. To determine the overall levels of current and past infection of COVID-
17
18 203 19, individuals who tested positive for IgG, IgM or both were summed. To determine the incidence
19
20 204 of COVID-19 during the survey period, IgM positivity alone was used.
21
22
23
24

25 205 **Adjustment of seroprevalence and exposure inference**

26
27
28 206 We first used a simple Bernoulli model to estimate the regional SARS-CoV-2 (the virus that causes
29
30 207 COVID-19) seroprevalence, after adjusting the proportion of individuals in each region with current
31
32 208 or past COVID-19 infection according to the sensitivity and specificity of the serology test.¹⁴ (The
33
34 209 term 'seroprevalence' below denotes the serology positive ratio already adjusted by the test.)
35
36 210 Further details of the method used can be found in **online supplemental method, appendix 1**. We
37
38 211 revised the mathematical model¹² to estimate the total exposure in the population by region after
39
40 212 taking into account waning antibody levels. Further details of the method used can be found in
41
42 213 **online supplemental method, appendix 1**.
43
44
45
46
47

48 214 **RESULTS**

49 50 51 52 215 **Demographic details**

53
54
55 216 This seroepidemiological study has provided estimates of the prevalence of SARS-CoV-2 antibodies
56
57 217 across Afghanistan, in urban and rural areas, and in the nine regions of the country. Of the 360
58
59 218 clusters identified for participation in the study, 338 (94%) were included; the remainder were
60

1
2
3 219 excluded due to insecure or inaccessible EAs and time constraints. A total of 9514 household
4
5 220 members from these 338 clusters were interviewed and tested for COVID-19. The mean age of
6
7 221 respondents was 27 years, 53.9% were male and 46.1% were female, 73% were from rural areas
8
9
10 222 (online supplemental table S2), and most participants (79.2%) were married.
11
12
13 223

14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

224 COVID-19 infections in Afghanistan

225 The total proportion of COVID-19 infections (including all positive results, the average of both
226 current and past infection) for the whole of Afghanistan was 31.5%. By region, Kabul had the highest
227 proportion of COVID-19 infections (53%), while the Central highlands region had the lowest
228 proportion, at 21.1% (figure 1).

229 Based on further analysis, the adjusted seroprevalence by region was consistent with the serosurvey
230 results. Kabul still had the highest adjusted seroprevalence (51.8%) (table 1 and figure 2).

231 RDT results for participants aged 18 years or more

232 In total, 5618 participants aged ≥ 18 years were interviewed and tested for this survey. Among this
233 age group, 2056 (35.1%) of individuals tested positive for antibodies against SARS-CoV-2 (table 2).
234 There were 885 (37.2%) females and 1170 (33.9%) males who tested positive, and there was a
235 higher proportion of positive tests in individuals who lived in urban areas compared with the
236 proportion in people who lived in rural areas (773, 42.3% versus 1323, 31.7%, respectively) (table 2).
237 Kabul region had the highest proportion of participants aged ≥ 18 years who tested positive for
238 antibodies against SARS-CoV-2 (357, 56.8%) (table 2). The survey results revealed that 164 (2.6%) of
239 participants aged ≥ 18 years were IgM-positive for COVID-19, i.e. they had a current infection, with
240 the highest proportion of current infections in the South-east region (37, 7.0%) (table 1).

241 **Table 1. Seroprevalence of SARS-CoV-2 antibodies and proportion of IgM-seropositive in**
242 **participants aged ≥ 18 years by region, area of residence and sex**

	Number of positive COVID-19 tests [#]	Seroprevalence % [95% CI]	p- valu e	Adjusted seroprevalence [95% CI]	Number of IgM- positive COVID- 19 tests	IgM- seropositive % COVID-19 tests [95% CI]	p- value
National Region	2056	35.1 [31–39.5]		29.8 [28.8, 30.7]	164	2.6 [2.0–3.5]	
Central	254	45.5 [37.8–53.4]	***	34.6 [31.6, 37.6]	28	4.3 [2.4–7.6]	
Central highlands	105	24.9 [17.9–33.7]		19.0 [16.4, 21.8]	5	1.0 [0.4–2.3]	***
East	294	49.1 [41.5–56.8]		41.5 [38.6, 44.4]	16	2.5 [1.4–4.5]	
Kabul	357	56.8 [52.0–62.0]		51.8 [48.8, 54.8]	17	2.7 [1.4–5.0]	
North	212	35.3 [28.1–43.4]		28.9 [26.3, 31.8]	7	1.4 [0.6–3.4]	
North-east	263	39.3 [31.9–47.4]		30.7 [28.1, 33.3]	26	4.0 [2.1–7.8]	
South	115	26.6 [19.0–36.0]		23.9 [20.7, 27.1]	8	1.6 [0.7–3.4]	

	South-east	221	40.9 [34.4–47.9]	30.5 [27.4, 33.6]	37	7.0 [3.7–12.9]
	West	235	39.8 [34.8–45.1]	32.4 [29.7, 35.2]	20	3.4 [1.8–6.3]
	Area of residence					
	Rural		31.7 [26.5–37.4]		121	3.7 [1.7–7.9]
	Urban		42.3 [35.7–49.2]		43	2.3 [1.2–4.2]
	Sex					
	Male	1170	33.9 [29–39.2]		104	2.4 [1.4–4.0]
	Female	885	37.2 [32–42.6]		60	4.1 [1.8–9.2]
	Age (years)					
	18–39	1109	33.7 [28.5–39.2]		96	2.7 [1.9–3.9]
	40–59	657	36.5 [31.9–41.3]		50	2.4 [1.6–3.7]
	60+	290	40.0 [31.8–48.2]		18	2.1 [1.1–4.2]

#The total number of positive COVID-19 tests includes all positive results: both current and past infections i.e. IgG-positive, IgM-positive or both. *p<.05, **p<.01, *** p<.001

CI, confidence interval

RDT results for participants aged 5 to 17 years

There were 4346 participants aged 5–17 years interviewed and tested for this survey. Among this age group, a total of 850 (25.3%) individuals tested positive for antibodies against SARS-CoV-2 (table 2), 401 (27.8%) females and 446 (24.2%) males. Again, there was a higher proportion of positive tests in individuals who lived in urban areas compared with the proportion among people who lived in rural areas (322, 30.8% versus 528, 23.4%, respectively) (table 2). Kabul region had the highest proportion of participants aged 5–17 years who tested positive for antibodies against SARS-CoV-2 (177, 46.4%) (table 2). There were 89 (3.3%) participants aged 5–17 years who were IgM-positive for COVID-19, with the highest proportion of current infections in the South region (7, 4.7%) (table 2).

Table 2. Seroprevalence of SARS-CoV-2 antibodies and proportion of IgM-seropositive results in participants aged 5–17 years by region, area of residence and sex.

	Number of positive COVID-19 tests [#]	Seroprevalence % [95% CI]	p-value	Number of IgM-positive COVID-19 tests	IgM-seropositive % COVID-19 tests [95% CI]	p-value
National	850	25.3 [20.5–30.8]		89	3.3 [1.8–6.3]	
Region						
Central	79	21.0 [14.5–29.3]*	*	10	2.8 [1.2–6.3]	**
Central highlands	42	14.6 [8.6–23.8]		3	1.6 [0.4–6.6]	
East	172	32.4 [26.8–38.6]		10	1.4 [0.7–3.1]	
Kabul	177	46.4 [40.8–52.1]		14	3.5 [1.6–7.3]	
North	96	23.0 [16.8–30.8]		6	1.2 [0.4–3.7]	
North-east	108	20.9 [15.1–28.2]		18	2.8 [1.0–7.6]	
South	55	24.4 [14.5–38.0]		7	4.7 [1.6–13.1]	
South-east	42	17.6 [10.6–27.6]		9	2.4 [0.8–6.8]	

West	79	24.5 [18.4–31.8]		12	3.2 [1.7–6.0]	
Area of residence						
Rural	528	23.4 [17.5–30.6]		60	3.7 [1.7–7.9]	
Urban	322	30.8 [24.8–37.5]		29	2.3 [1.2–4.2]	
Sex						
Male	446	24.2 [18.5–31]		47	2.4 [1.4–4.0]	
Female	401	27.8 [21.3–33]		42	4.1 [1.8–9.2]	
Age (years)						
5–9	175	[13.4–26.2]	**	20	3.3 [1.1–9.5]	**
10–14	365	[20.8–33.8]		40	3.7 [1.7–7.9]	
15–17	310	[23.5–35.6]		29	2.8 [1.5–5.2]	

#The total number of positive COVID-19 tests includes all positive results: both current and past infections i.e. IgG-positive, IgM-positive or both. *p<.05, **p<.01, *** p<.001

CI, confidence interval

Predictions for cumulative exposure in the population up to 21 July 2020 in the nine regions of Afghanistan are shown in figure 3. The method used for the modelling analysis, which was developed by the COVID-19 International Modelling Consortium (CoMo Consortium), is detailed in the [online supplemental method, appendix 1](#).

The solid orange circles and black error bars in the panel for each region represent the observed seroprevalence data and the associated credible interval (CrI) after adjusting for the sensitivity and specificity of the antibody test. The green and orange lines show the median predictions for exposure and seroprevalence, respectively, while the shaded areas correspond to 95% CrI. The median predicted exposure levels by region (expressed as the proportion of the population that has been infected) as of 21 July 2020 are shown on the map of Afghanistan.

DISCUSSION

This national survey of COVID-19 morbidity in Afghanistan, which was conducted during June and July 2021, revealed that around 10 million people (31.5% of the population) were seropositive for antibodies against SARS-CoV-2, reflecting either current or previous COVID-19 infection. The population of Afghanistan is estimated to comprise approximately 33.6 million people.¹⁵ This finding is reasonably consistent with the results of another telephone survey conducted before July 2020 with a randomly selected sample of 713 healthcare workers to estimate COVID-19 morbidity in the

1
2
3 281 country. The estimated proportion of individuals who had experienced COVID-19 signs and
4
5 282 symptoms was 49.6%, which is close to the value for total infections for most regions reported in the
6
7 283 present study, however, no laboratory testing was conducted for the phone survey, which only
8
9 284 collected clinical information about symptoms. There is a discrepancy between the serosurvey
10
11 285 results and the detected number of COVID-19 infections reported to the surveillance system in the
12
13 286 country (36 710 cases reported by the surveillance system as of 30 July 2020 and 156 363 cases as of
14
15 287 5 November 2021) in Afghanistan. The under-reporting of COVID-19 cases is a problem globally due
16
17 288 to limited testing availability, flawed test sensitivity, poor surveillance and the indeterminate
18
19 289 proportion of asymptomatic infections.¹⁶

20
21
22
23
24 290 A modelling exercise was performed using the CoMo model to estimate the peak incidence of
25
26 291 COVID-19 in Afghanistan. The CoMo model was developed by the CoMo Consortium.¹⁷ The CoMo
27
28 292 Consortium adopts a participatory modelling approach,¹⁸ which places in-country subject matter
29
30 293 experts at the forefront of model development to ensure that contextual considerations, such as
31
32 294 local infrastructure, human resources and sociocultural considerations, are fully taken into account.
33
34 295 The CoMo model was used to estimate the peak incidence of COVID-19 in Afghanistan under four
35
36 296 scenarios: good, bad, very bad and appropriate, depending on the coverage of and adherence to the
37
38 297 NPIs. If the use of NPIs (in a very bad scenario) is not considered, then the COVID-19 peak was
39
40 298 predicted to occur in June 2020, with an estimated 69.6% of the population infected and 20 509
41
42 299 deaths by the end of 2020.

43
44
45
46
47 300 In communicable disease epidemiology, one of the key parameters used in decision-making is the
48
49 301 estimate of herd immunity in a population. Herd immunity occurs when a certain proportion of the
50
51 302 population is immune to a given infectious disease, reducing the probability that the disease will be
52
53 303 transmitted from one individual to another, thus helping to protect the entire population from that
54
55 304 disease.¹⁹ Herd immunity can be achieved either through individuals being exposed or vaccinated.
56
57 305 Determining a country's herd immunity threshold to a given disease is directly related to estimates
58
59
60

1
2
3 306 of the basic reproductive number, R_0 , of that disease. R_0 indicates the average number of
4
5 307 individuals one infected individual can go on to infect in a fully susceptible population. Different herd
6
7 308 immunity thresholds in different contexts have been estimated for COVID-19, ranging from 43% to
8
9 309 85%.¹⁹⁻²⁴ For example, one study indicated that if $R_0=3$, i.e. one infected individual can infect up to
10
11 310 three others, meaning 67% of the population must be immune to achieve herd immunity.²⁰
12
13 311 Estimates by Johns Hopkins University suggest that 70% of the population must be immune to
14
15 312 achieve herd immunity and end restrictions on people's day to day lives¹⁹, while another study
16
17 313 suggested that R_0 values of 1–2, 2–4 and >4 would require herd immunity thresholds of 50%, 56.1–
18
19 314 74.8% and 77.9–85%, respectively.²¹ In addition to R_0 and the herd immunity threshold, the rate of
20
21 315 antibody decline post-infection must also be considered, with one study suggesting that antibodies
22
23 316 to COVID-19 decline within 94 days of infection.¹⁰
24
25
26
27
28 317 A study conducted by Eckerle and Meyer revealed that by mid-2020, an insufficient proportion
29
30 318 of the population had been infected globally to achieve herd immunity, and these findings were
31
32 319 confirmed by reports of low COVID-19 morbidity levels from countries such as Sweden, where
33
34 320 an infection rate of 7% was reported by the end of April despite no lockdown; the mentioned
35
36 321 study also states that obtaining herd immunity by exposing the population to the disease results
37
38 322 in the simultaneous infection of the majority of the population and paves the way for a second
39
40 323 wave of the disease.
41
42
43
44
45 324 These estimates of herd immunity thresholds suggest that the present survey findings, of a SARS-
46
47 325 CoV-2 antibody seroprevalence of approximately 32% among the population in Afghanistan, mean
48
49 326 that the herd immunity threshold had not been reached by the time of the study and most of the
50
51 327 country's population remained at risk of infection. However, the herd immunity threshold may have
52
53 328 been crossed in some local areas. In Kabul province, for example, more than half of the population
54
55 329 has been infected, which exceeds the lowest reported herd immunity threshold of 43%. However, as
56
57 330 the majority of the population remains at risk of infection, NPIs should be lifted gradually, as per
58
59
60

1
2
3 331 WHO guidelines.²⁵ It should also be noted that this survey was conducted at a time when the SARS-
4
5 332 CoV-2 alpha variant was the most prevalent variant in Afghanistan; it is unclear what effect the
6
7 333 arrival of new variants, such as the delta and omicron variants, will have on herd immunity
8
9
10 334 thresholds.

11
12
13 335 Based on evidence from countries with a similar context to Afghanistan, and if we assume an R0 in
14
15 336 the country of 2–3, the herd immunity threshold would be between 56% and 75%. Kabul province,
16
17 337 with a SARS-CoV-2 seroprevalence of 53%, was within range of this threshold. The Eastern and
18
19 338 Central regions, with SARS-CoV-2 seroprevalences of 34% to 42%, were in a relatively good position,
20
21 339 but the remaining regions with SARS-CoV-2 seroprevalence of less than 35% were in a worse
22
23
24 340 position and not yet close to the herd immunity threshold at the time of the [study \(Online](#)
25
26 341 [supplementary Table S3\)](#). As in many low- and middle-income countries, COVID-19 vaccination rates
27
28 342 in Afghanistan are low, with just 12% of the population currently fully vaccinated.⁵ With the
29
30 343 disruptions to the health system as a result of the evolving political situation in the country, the
31
32
33 344 COVID-19 response may deteriorate if control measures are not implemented and vigilantly
34
35 345 maintained.

36
37
38 346 Based on the evidence outlined above, the NPIs currently in place in Afghanistan should not have
39
40 347 been lifted, as the herd immunity threshold for the nation has yet to be reached either through
41
42 348 natural infection or vaccination. If the NPIs are lifted, the rates of hospitalisation will increase, as will
43
44 349 the number of patients requiring ventilation; this will place the health system under considerable
45
46
47 350 pressure. However, after July 2021, the restrictions were reduced and since then the country has
48
49 351 only focused on school closures as a mitigation measure to balance the economy, social life and the
50
51 352 impact of COVID-19 on the health system. It is worth mentioning that with the recent transition of
52
53 353 government in Afghanistan and decreased funding for the country's health system, there are
54
55
56 354 evolving challenges that will ultimately lead to the increased spread of COVID-19 and other
57
58 355 infectious diseases. Greater levels of poverty, a displaced population and poor sanitation will further
59
60

1
2
3 356 exacerbate this problem. The influx of refugees from Afghanistan to other countries might also
4
5 357 facilitate the cross-border spread of disease. Particularly with the emergence of new variants and
6
7 358 low vaccination coverage, it is crucial to have continued public health and social measures to
8
9 359 mitigate the impact of COVID-19 in a conflict-affected and unstable country. For the continuation of
10
11 360 health services, functional hospitals, surveillance systems and laboratories, as well as a skilled
12
13 361 healthcare workforce, are needed to mitigate the spread of COVID-19 and other infections within
14
15 362 Afghanistan and prevent the regional and even global spread of disease.
16
17
18

19 363 This study had some limitations. First, the time available to conduct the survey was limited. Second,
20
21 364 security concerns meant that not all areas could be surveyed; the inability to conduct proper
22
23 365 household listing and create maps for enumeration areas in those areas where the government
24
25 366 lacked control may have affected the findings. Third, the findings may not reflect the current
26
27 367 situation with regards to the new SARS-CoV-2 delta and omicron variants of concern, as the R0 for
28
29 368 these variants is not well-established. Once a stable estimate of the R0 for these variants has been
30
31 369 established then our findings can be adjusted accordingly to assist with programme planning.
32
33 370 Fourth, the data were entered in the DHIS2 database, which created many challenges for data
34
35 371 verification, household matching and the subsequent analysis. In future surveys, it would be
36
37 372 preferable to collect data by entering them directly via a tablet or similar appropriate research data
38
39 373 entry tool to improve the data quality.
40
41
42
43
44
45
46
47

48 375 **CONCLUSION**

49
50
51
52 376 Although the immunity threshold may have been reached in some localities within Afghanistan,
53
54 377 specifically Kabul, this threshold has not yet been reached among the country's entire population. In
55
56 378 particular, the proportion of the population that is seropositive for antibodies against SARS-CoV-2 is
57
58 379 much lower in rural areas than urban areas. The seroprevalence represents a lower estimate of herd
59
60

1
2
3 380 immunity and the predicted exposure represents an upper limit. Given the large proportion of the
4
5 381 population that remains susceptible to COVID-19 infection, and limited COVID-19 vaccination
6
7 382 coverage, NPIs and vigilance should remain in place to protect the health system from an
8
9
10 383 unmanageable burden of hospitalisations. The link between the presence of antibodies and
11
12 384 immunity has yet to be established, as is the link between prior exposure and immunity. As antibody
13
14 385 levels wane, then seroprevalence may provide an underestimate of immunity but, conversely, if
15
16 386 immunity wanes, then prior exposure would provide a higher estimate.

387 **Data availability statement**

388 Survey serology data are stored in the Ministry of Public Health national database and are available
389 upon reasonable request. All data, code and materials used in the analyses can be accessed at:
390 <https://github.com/SiyuChenOxf/AfghanistanSerologyStudy/tree/master>. All parameter estimates
391 and figures 3 and 4 can be reproduced using the code provided. This work is licensed under a
392 Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use,
393 distribution and reproduction in any medium, provided the original work is properly cited.

394 **Acknowledgments**

395 This SARS-CoV-2 Seroprevalence in Afghanistan: A Population-based Study has been supported by
396 WHO Unity Studies, a global sero-epidemiological standardization initiative, with funding to WHO by
397 the COVID-19 Solidarity Response Fund and the German Federal Ministry of Health (BMG) COVID-19
398 Research and development Fund. We would like to acknowledge the Monitoring and Evaluation and
399 Health Information System General Directorate, which led the study in Afghanistan. The study was
400 funded by Ministry of Public Health. We would also like to acknowledge the regional master trainers
401 and the National Statistics and Information Authority (NSIA) team for their support in conducting the
402 survey. We would additionally like to thank the following for their efforts during the survey: the WHO

1
2
3 403 team, USAID, World Bank, CDC Atlanta, Performance Management Office (PMO), Grants and
4
5 404 Contracts Management Unit (GCMU), Afghanistan National Public Health Institute (ANPHI), all
6
7 405 Provincial Public Health Directorates (PPHDs), Health Promotion Department (HPD), all Basic Package
8
9 406 of Health Services (BPHS)- and Essential Package of Hospital Service (EPHS)-implementing NGOs, the
10
11 407 survey steering committee team, and the data management team. We acknowledge the national and
12
13 408 provincial surveillance teams for their extensive support in conducting the survey. We also appreciate
14
15 409 all those individuals who supported the data collection, data entry, and listing of households for the
16
17 410 survey frame. Medical writing assistance and editorial support was provided by Adam Bodley,
18
19 411 according to Good Publication Practice guidelines. The views expressed in this manuscript are those
20
21 412 of the authors and do not necessarily reflect the views of the authors' affiliated organisations or the
22
23 413 organisations listed here.
24
25
26
27
28

29 414 **Funding Statement:**

30
31
32
33 415 There are no funders to report for this submission.
34
35
36

37 416 **Authorship & Contributorship statement**

38
39
40
41 417 SA. Seedzai, planned the study, lead the study design and tools in the country, coordinated with
42
43 418 national authorities, provided ToT training, monitored process of data collection, analysed the data
44
45 419 and contributed to writing and finalizing the report and findings of the study and organized
46
47 420 dissemination workshops, seminars in the country.

48
49 421 MN. Sahak, supported in planning the study, supported in scientific study design, coordinated
50
51 422 internally for sharing scientific WHO tools and protocol, supported in data analysis, contributed to
52
53 423 writing and finalizing the study report.

54
55 424 F.Arifi, supported in the design and tools of the study, contributed in analysing and interpreting the
56
57 425 results, supported in drafting the report and communicated with all authors for compiling and
58
59 426 incorporating the feedback in the final version.

60
427 E. Aly, supported in the study design and tools, contributed to data cleaning, analysis &
428 interpretation, and writing the report and finalization of the study report.

429 M.V.Gurp, supported in study design, complex and multi-level analysis, reviewed and revised the
430 study findings and report.

- 1
2
3 431 L.J. White, supported in study method, analysis, interpretation of findings and modelling based on
4 432 the serosurvey data, reviewed and revised the final study report.
5
6 433 S. Chen, supported in study method, analysis, interpretation of findings and modelling based on the
7 434 serosurvey data, reviewed and revised the final study report.
8
9 435 A. Barakat, supported in proper laboratory testing of the study method, selection of testing kits,
10 436 contributed in interpretation of findings, reviewed the draft report and provided inputs to the
11 437 revision and finalization of the study report.
12
13 438 G. Azim, supported in planning the study, facilitated internal coordination within MoPH, provided
14 439 ToT training, supervised the data collection in a cluster, supported in drafting the study report and
15 440 contributed in finalizing the study findings.
16
17 441 B. Rasoly, supported in planning the study, provided ToT training, supervised the data collection in a
18 442 cluster, supported in data entry and cleaning, drafting the study report and contributed in finalizing
19 443 the study findings.
20
21 444 S. Safi, supported in planning the study, provided ToT training, supervised the data collection in a
22 445 clusters and collected data where needed, supported in data entry and cleaning, drafting the study
23 446 report and contributed in finalizing the study findings.
24
25 447 J. Flegg, contributed in study method, interpretation of serosurvey findings for the context, analysing
26 448 the data, reviewed the draft report and provided inputs to the finalization of the study report.
27
28 449 N. Ahmed, contributed in study method, interpretation of serosurvey findings for the context,
29 450 analysing the data, reviewed the draft report and provided inputs to the finalization of the study
30 451 report.
31
32 452 M.J. Ahadi, supported in planning the study, provided ToT training, supervised the data collection in
33 453 a clusters and collected data where needed, supported in data entry and cleaning, drafting the study
34 454 report and contributed in finalizing the study findings.
35
36 455 N.M, Achakzai, supported in proper laboratory testing of the study method, selection of testing kits,
37 456 contributed in interpretation of findings, reviewed the draft report and provided inputs to the
38 457 revision and finalization of the study report.
39
40 458 A. AbouZeid, supported in scientific study design, internal WHO coordination, supported in data
41 459 analysis, contributed to writing and finalizing the study report

460 Competing interests

461 No, there are no competing interests for any author

462 REFERENCES

463

- 464 1. COVID-19 dashboard. Johns Hopkins COVID-19 dashboard Johns Hopkins University; 2021
465 [Available from: <https://coronavirus.jhu.edu/map.html>.
466 2. Burki T. No end in sight for the Brazilian COVID-19 crisis. *Lancet Microbe* 2021;2(5):e180. doi:
467 10.1016/s2666-5247(21)00095-1 [published Online First: 2021/05/11]

- 1
2
3 468 3. Thiagarajan K. Why is India having a covid-19 surge? *BMJ* 2021;373:n1124. doi:
4 469 10.1136/bmj.n1124
- 5 470 4. ReliefWeb. Second wave of Covid-19 causes alarm: ReliefWeb; 2021 [Available from:
6 471 <https://reliefweb.int/report/nepal/second-wave-covid-19-causes-alarm>.
- 7 472 5. Ministry of Public Health. COVID-19 Dashboard 2021 [Available from: <http://covid.moph-dw.org/>
8 473 accessed 11/06/2021 2021.
- 9 474 6. CDC. Interim Guidelines for COVID-19 Antibody Testing Atlanta: Centers for Disease Control and
10 475 Prevention; 2021 [Available from: [https://www.cdc.gov/coronavirus/2019-
11 476 ncov/lab/resources/antibody-tests-guidelines.html2021](https://www.cdc.gov/coronavirus/2019-ncov/lab/resources/antibody-tests-guidelines.html2021).
- 12 477 7. Burgess S, Ponsford MJ, Gill D. Are we underestimating seroprevalence of SARS-CoV-2? *BMJ*
13 478 2020;370:m3364. doi: 10.1136/bmj.m3364
- 14 479 8. Ibarondo FJ, Fulcher JA, Goodman-Meza D, et al. Rapid Decay of Anti-SARS-CoV-2 Antibodies in
15 480 Persons with Mild Covid-19. *New England Journal of Medicine* 2020;383(11):1085-87. doi:
16 481 10.1056/NEJMc2025179
- 17 482 9. Harris RJ, Whitaker HJ, Andrews NJ, et al. Serological surveillance of SARS-CoV-2: Six-month trends
18 483 and antibody response in a cohort of public health workers. *The Journal of infection*
19 484 2021;82(5):162-69. doi: 10.1016/j.jinf.2021.03.015 [published Online First: 2021/03/22]
- 20 485 10. Seow J, Graham C, Merrick B, et al. Longitudinal observation and decline of neutralizing antibody
21 486 responses in the three months following SARS-CoV-2 infection in humans. *Nature*
22 487 *Microbiology* 2020;5(12):1598-607. doi: 10.1038/s41564-020-00813-8
- 23 488 11. WHO. Population-based age-stratified seroepidemiological investigation protocol for COVID-19
24 489 virus infection, 17 March 2020. Geneva: WHO, 2020.
- 25 490 12. Chen S, Flegg JA, White LJ, et al. Levels of SARS-CoV-2 population exposure are considerably
26 491 higher than suggested by seroprevalence surveys. *medRxiv* 2021:2021.01.08.21249432. doi:
27 492 10.1101/2021.01.08.21249432
- 28 493 13. StataCorp. Stata Statistical Software: Release 15 College Station, TX: StataCorp LLC; 2017
29 494 [Available from: <https://www.stata.com/stata15/>.
- 30 495 14. Healgen Scientific LLC. COVID-19 IgG/IgM Rapid Test Cassette (Whole Blood/Serum/Plasma)
31 496 Instruction for Use 2020 [Available from: <https://www.fda.gov/media/138438/download>.
- 32 497 15. NSIA. Estimated Population of Afghanistan 2021-22: NSIA; 2021 [Available from:
33 498 [https://www.nsia.gov.af:8080/wp-content/uploads/2021/06/Estimated-Population-2-of-
34 499 AFG-New-Design-14.pdf](https://www.nsia.gov.af:8080/wp-content/uploads/2021/06/Estimated-Population-2-of-AFG-New-Design-14.pdf).
- 35 500 16. Wu SL, Mertens AN, Crider YS, et al. Substantial underestimation of SARS-CoV-2 infection in the
36 501 United States. *Nature Communications* 2020;11(1):4507. doi: 10.1038/s41467-020-18272-4
- 37 502 17. CoMo Consortium. CoMo Consortium: BMJ; 2021 [Available from: <https://como.bmj.com/2021>.
- 38 503 18. Aguas R, White L, Hupert N, et al. Modelling the COVID-19 pandemic in context: an international
39 504 participatory approach. *BMJ Glob Health* 2020;5(12) doi: 10.1136/bmjgh-2020-003126
40 505 [published Online First: 2020/12/29]
- 41 506 19. D'Souza G, Dowdy D. What is Herd Immunity and How Can We Achieve It With COVID-19?
42 507 Maryland: Johns Hopkins Bloomberg School of Public Health; 2021 [Available from:
43 508 <https://www.jhsph.edu/covid-19/articles/achieving-herd-immunity-with-covid19.html2021>.
- 44 509 20. Randolph HE, Barreiro LB. Herd Immunity: Understanding COVID-19. *Immunity* 2020;52(5):737-
45 510 41. doi: 10.1016/j.immuni.2020.04.012 [published Online First: 2020/05/21]
- 46 511 21. Kwok KO, Lai F, Wei WI, et al. Herd immunity - estimating the level required to halt the COVID-19
47 512 epidemics in affected countries. *J Infect* 2020;80(6):e32-e33. doi: 10.1016/j.jinf.2020.03.027
48 513 [published Online First: 2020/03/27]
- 49 514 22. Gomes MGM, Corder RM, King JG, et al. Individual variation in susceptibility or exposure to SARS-
50 515 CoV-2 lowers the herd immunity threshold. *medRxiv* 2020:2020.04.27.20081893. doi:
51 516 10.1101/2020.04.27.20081893

- 1
2
3 517 23. Altmann DM, Douek DC, Boyton RJ. What policy makers need to know about COVID-19
4 518 protective immunity. *Lancet* 2020;395(10236):1527-29. doi: 10.1016/s0140-6736(20)30985-
5 519 5 [published Online First: 2020/05/01]
6
7 520 24. Britton T, Ball F, Trapman P. The disease-induced herd immunity level for Covid-19 is
8 521 substantially lower than the classical herd immunity level. *arXivorg* 2020:
9 522 arXiv:2005.03085v1.
10 523 25. World Health Organization. Considerations for implementing and adjusting public health and
11 524 social measures in the context of COVID-19. Geneva: World Health Organization, 2021.
12
13 525
14
15 526
16
17 527
18
19 528
20
21 529

22 530 **Figure Caption/Legend**

25 531 Figure 1. Seroprevalence of SARS-CoV-2 antibodies (including all positive results: IgG-positive, IgM-
26
27 532 positive or both) among all age groups by region in Afghanistan.

30 533 Figure 2. Adjusted seroprevalence by region by the sensitivity and specificity of the serology test for
31
32 534 IgG-positive and/or IgM-positive.

34 535 Figure 3. Time course of the COVID-19 pandemic up to 21 July 2020 for the nine regions in
35
36 536 Afghanistan, for all age groups.
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

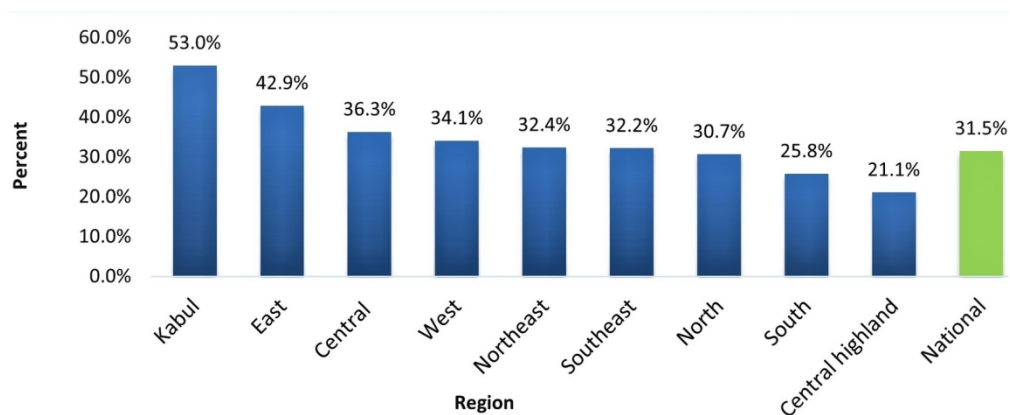


Figure 1. Seroprevalence of SARS-CoV-2 antibodies (including all positive results: IgG-positive, IgM-positive or both) among all age groups by region in Afghanistan.

149x63mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

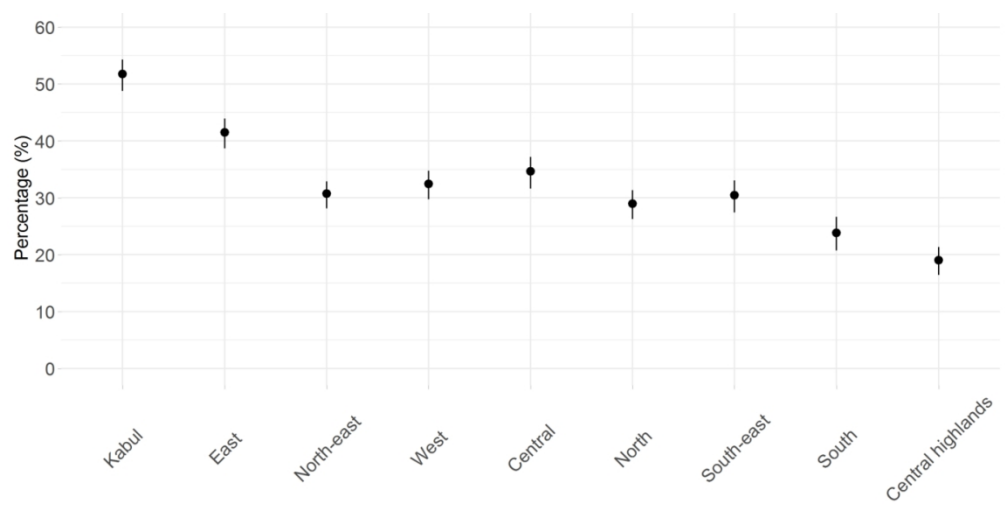


Figure 2. Adjusted seroprevalence by region by the sensitivity and specificity of the serology test for IgG-positive and/or IgM-positive.

158x82mm (220 x 220 DPI)

BMJ Open: first published as 10.1136/bmjopen-2021-060739 on 27 July 2022. Downloaded from <http://bmjopen.bmj.com/> on April 22, 2024 by guest. Protected by copyright.

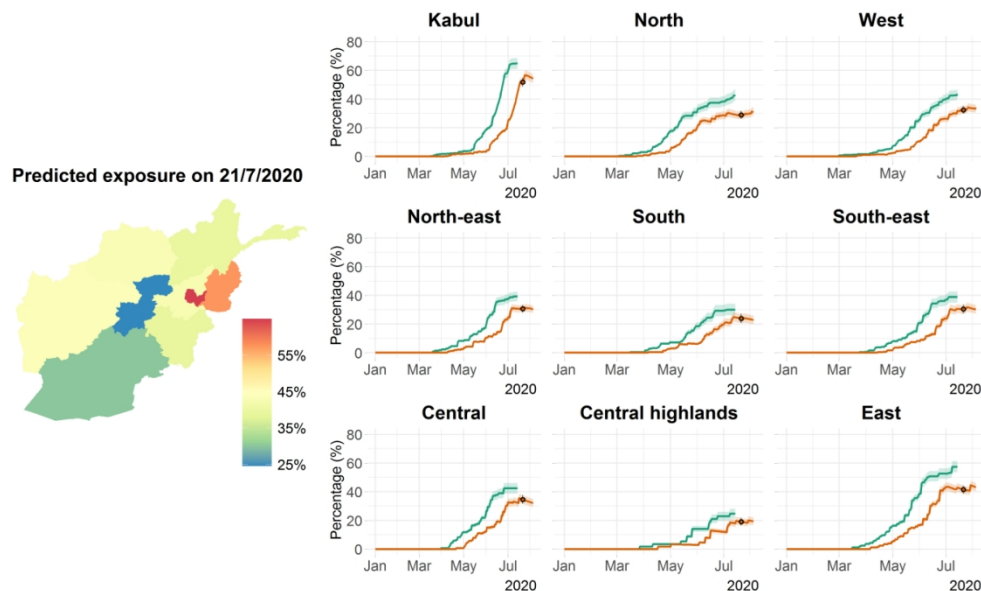


Figure 3. Time course of the COVID-19 pandemic up to 21 July 2020 for the nine regions in Afghanistan, for all age groups.

159x93mm (220 x 220 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Supplementary material

Supplementary appendix 1 – Methods

Data sources

Regional daily deaths

The documented daily mortality data associated with COVID-19, which might be subject to underreporting, for each of the nine Afghanistan regions (Kabul, East, West, North, South, North-east, South-east, Central, and Central highlands) from January 1 to August 4, 2020, were extracted from the Afghanistan Ministry of Health DHIS2 database.

Regional serology data

The proportion of individuals with current or past COVID-19 infection in each region were obtained from the seroepidemiological study data (table 3, main text). The serology survey provided a result for both IgM and IgG antibodies for each participant, using the COVID-19 IgG/IgM Rapid Test Cassette.¹ The dynamics of IgM and IgG antibodies within an infected individual are complicated.² Here, we take the simplified view that an individual who is either IgG positive and/or IgM positive has been exposed to COVID-19 (either past or current infection). Therefore, in the following modelling, the sensitivity and specificity provided by the manufacturer of the imperfect serology test for IgG+ and/or IgM+ was employed.

Adjustment of seroprevalence

We first used a simple Bernoulli model to estimate the regional seroprevalence, after adjusting the proportion of individuals in each region with current or past COVID-19 infection (table 3, main text) according to the sensitivity and specificity of the serology test.¹ (The term ‘seroprevalence’ below denotes the serology positive ratio already adjusted by the test used.) The Bayesian framework was as follows:

$$x_i(t_0) \sim \text{Beta}(1,1)$$

$$w_{ij} \sim \text{bernoulli}(k_{se} \times x_i(t_0) + (1 - k_{sp}) \times (1 - x_i(t_0)))$$

$$w_{ij} = \begin{cases} 0, \text{IgG} + \text{ or } \text{IgM} + \\ 1, \text{IgG} - \text{ and } \text{IgM} - \end{cases}, j = \{1, \dots, N_i\} \quad (1)$$

where in the first equation given above we have specified a uniform prior for $x_i(t_0)$, which is the proportion of the population in region i that is serology positive, either for IgM or IgG, on July 21,

2020; w_{ij} is the serology survey result for the j -th participant in the serology study from region i ; N_i is the total number of participants in the serology survey for region i (table S1); and k_{se} (k_{sp}) is the median of the serology test cassette sensitivity (specificity) reported by the manufacturer.¹ The posterior for seroprevalence on the date the serology survey was conducted, $t = t_0$, was estimated using a Markov chain Monte Carlo (MCMC) implemented in Rstan³ and denoted as $\tilde{x}_i(t_0)$. The code associated with this work is publicly available, see the Data availability statement of the paper.

Mechanistic model

We revised the mathematical model⁴ to account for the underreporting of mortality in the Afghanistan setting according to the varying serology status of the population, $X_i(t)$, of each regional population (for each region $i = 1, \dots, 9$, corresponding to Kabul, East, West, North, South, North-east, South-east, Central, and Central highlands, respectively). The population that has positive serology status increased with exposure of the population to COVID-19 and decreased due to the waning of antibodies.

Given that the constant age-averaged infection fatality rate by region is β_i , the documented mortality over time by region is $m_i(t)$, and the reporting rate of mortality associated with COVID-19 by region is q_i , which is assumed to be constant over time, then, at each time step, of the $\frac{1}{q_i\beta_i}m_i(t)$ individuals who were exposed, $\frac{1}{q_i}m_i(t)$ die and the remaining number of individuals, $\frac{1-\beta_i}{q_i\beta_i}m_i(t)$, seroconvert from negative to positive. Then, assuming that positive individuals convert to negative at a rate of α , the equation for the rate of change of the number of seropositive individuals is given by:

$$\frac{dX_i(t)}{dt} = \frac{(1-\beta_i)}{q_i\beta_i}m_i(t) - \alpha X_i(t) \quad (2)$$

Solving Equation (2), subject to the initial condition $X_i(t=0) = 0$ where $t = 0$ is time since January 1, 2020, gives:

$$X_i(t) = \frac{(1-\beta_i)e^{-\alpha t}}{q_i\beta_i} \int_0^t e^{\alpha r} m_i(r) dr \quad (3)$$

Discretising Equation **Error! Reference source not found.** with daily intervals ($\Delta r = 1$) gives:

$$X_i(t) = \frac{(1-\beta_i)e^{-\alpha t}}{q_i\beta_i} \sum_{r=0}^t e^{\alpha r} m_i(r) \quad (4)$$

Then, the proportion of the population that is serology positive over time, $x_i(t)$, is

$$x_i(t) = \frac{X_i(t)}{P_i - \frac{\sum_{r=0}^t m_i(r)}{q_i}} \quad (5)$$

Where P_i is the reported population in region i before the COVID-19 outbreak, and the total proportion of the population that has been exposed over time, $\varepsilon(t - \delta_\epsilon)$, is

$$\varepsilon_i(t - \delta_\epsilon) = \frac{\frac{1 - \beta_i}{q_i \beta_i} \sum_{r=0}^t m_i(r)}{P_i - \frac{\sum_{r=0}^t m_i(r)}{q_i}} \quad (6)$$

where δ_ϵ is the time lag between exposure and seroconversion and is fixed at 21 days.⁴

Exposure inference

We use the posterior samples of seroprevalence at t_0 , $\tilde{x}_i(t_0)$, from the MCMC and combine it with Equations (4) and (5) to calculate the posterior samples of reporting rate for mortality, \tilde{q}_i :

$$\tilde{q}_i = \frac{(1 - \beta_i) e^{-\alpha t_0} \sum_{t=0}^{t_0} e^{\alpha t} m_i(t)}{x_i(t_0) \beta_i P_i} + \frac{\sum_{t=0}^{t_0} m_i(t)}{P_i} \quad (7)$$

Compared with the total population in Afghanistan prior to 2020 (approximately 38 million people), the cumulative mortality associated with COVID-19 by the date of serology survey, $\sum_{t=0}^{t_0} m(t)$, is small. Therefore, it is reasonable to neglect it from Equation (7), which then gives:

$$\tilde{q}_i \approx \frac{(1 - \beta_i) e^{-\alpha t_0} \sum_{t=0}^{t_0} e^{\alpha t} m_i(t)}{x_i(t_0) \beta_i P_i} \quad (8)$$

Combining Equations (4), (5) and (8) we can obtain samples of seroprevalence over time, $\tilde{x}_i(t)$:

$$\tilde{x}_i(t) \approx \frac{x_i(t_0) \sum_{r=0}^t e^{\alpha r} m_i(r)}{e^{\alpha(t-t_0)} \sum_{t=0}^{t_0} e^{\alpha t} m_i(t)} \quad (9)$$

From Equations (6) and (8) we can obtain samples of the total proportion of the population that has been exposed over time, $\tilde{\varepsilon}_i(t - \delta_\epsilon)$:

$$\tilde{\varepsilon}_i(t - \delta_\epsilon) \approx x_i(t_0) e^{\alpha t_0} \frac{\sum_{r=0}^t m_i(r)}{\sum_{t=0}^{t_0} e^{\alpha t} m_i(t)} \quad (10)$$

Note that the seroprevalence (9) and exposure (10) over time are not dependent on β . We use the median estimation of α from the constant infection fatality ratio (IFR) model from Chen et al⁴ as an input to Equations (9) and (10).

81

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

1. Healgen Scientific LLC. COVID-19 IgG/IgM Rapid Test Cassette (Whole Blood/Serum/Plasma) Instruction for Use 2020 [Available from: <https://www.fda.gov/media/138438/download>].
2. Seow J, Graham C, Merrick B, et al. Longitudinal observation and decline of neutralizing antibody responses in the three months following SARS-CoV-2 infection in humans. *Nature Microbiology* 2020;5(12):1598-607. doi: 10.1038/s41564-020-00813-8
3. Stan Development Team. RStan: the R interface to Stan; R package version 2.21.2. 2020 [Available from: <http://mc-stan.org/>].
4. Chen S, Flegg JA, White LJ, et al. Levels of SARS-CoV-2 population exposure are considerably higher than suggested by seroprevalence surveys. *medRxiv* 2021:2021.01.08.21249432. doi: 10.1101/2021.01.08.21249432

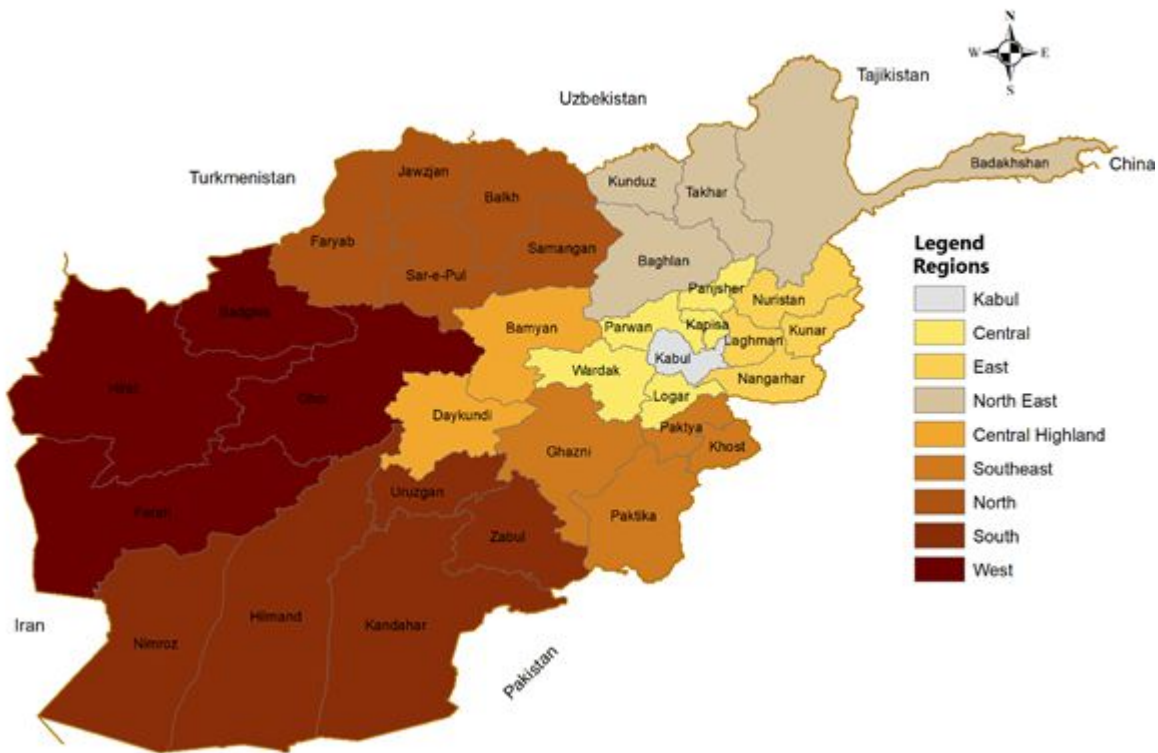
94 **Supplementary Table S2. Sample size for the regional serology survey**

Region	Sample size
Kabul	1104
Central	1056
Central highlands	902
East	1233
North	1071
South	738
North-east	1265
South-east	969
West	1176

95
96 **Supplementary Table S2**

Respondents' characteristics	Number	Percentage
Total respondents	9514	100%
Sex		
Male	5128	53.9%
Female	4386	46.1%
Age		
5–17 years	4346	45.7%
18 years or more	5168	54.3%
Geographical area		
Urban	2574	27%
Rural	6940	73%
Region		
Kabul	1104	11.6%
Central	1056	11.1%
Central highlands	902	9.4%
East	1233	13.0%
North	1071	11.2%
South	738	7.8%
North-east	1265	13.3%
South-east	969	10.2%
West	1176	12.4%

97
98



99
100
101 **Supplementary Figure S1.** A map showing the nine regions in Afghanistan where the study was
102 conducted (the eight regions of Afghanistan plus Kabul province).
103

104 **Supplementary Table S3. Seroprevalence of SARS-CoV-2 antibodies in Afghanistan by region,**
 105 **indicating whether herd immunity has been reached**

Region	Seroprevalence	Herd immunity reached (based on minimum to maximum of 43% and 85%, respectively)	Number of individuals at risk of infection based on the average of all reported herd immunity thresholds
National	31.5%	No	8 462 611
Kabul	53%	Yes, if based on a herd immunity threshold of 43%; otherwise, no	352 090
East	42.9%	No	479 674
Central	36.3%	No	579 968
West	34.1%	No	1 020 314
North-east	32.4%	No	1 164 297
South-east	32.2%	No	925 019
North	30.7%	No	1 227 256
South	25.8%	No	925 019
Central highlands	21.1%	No	386 875

106

107

STROBE Statement—Checklist of items that should be included in reports of *cross-sectional studies*

	Item No	Recommendation	Page No
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	2
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3-4
Objectives	3	State specific objectives, including any prespecified hypotheses	4
Methods			
Study design	4	Present key elements of study design early in the paper	5-8
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5-6
Participants	6	(a) Give the eligibility criteria, and the sources and methods of selection of participants	6
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6-7
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6
Bias	9	Describe any efforts to address potential sources of bias	8
Study size	10	Explain how the study size was arrived at	6
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	6-7
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	7-8
		(b) Describe any methods used to examine subgroups and interactions	
		(c) Explain how missing data were addressed	7
		(d) If applicable, describe analytical methods taking account of sampling strategy	7
		(e) Describe any sensitivity analyses	8
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	8
		(b) Give reasons for non-participation at each stage	8
		(c) Consider use of a flow diagram	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	8-9
		(b) Indicate number of participants with missing data for each variable of interest	8
Outcome data	15*	Report numbers of outcome events or summary measures	9
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	9-13

		(b) Report category boundaries when continuous variables were categorized	
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	13
Discussion			
Key results	18	Summarise key results with reference to study objectives	13-17
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	17
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	17
Generalisability	21	Discuss the generalisability (external validity) of the study results	17
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	18

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

COVID-19 morbidity in Afghanistan: a nationwide, population-based seroepidemiological study

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2021-060739.R1
Article Type:	Original research
Date Submitted by the Author:	30-May-2022
Complete List of Authors:	Saeedzai, Sayed Ataullah; Ministry of Public Health, Monitoring, Evaluation and Health Information System Sahak, Mohammad; World Health Organization, WHE Arifi, Fatima; WHO Aly, Eman; World Health Organisation Regional Office for the Eastern Mediterranean Van Gurp, Margo; KIT Royal Tropical Institute White, Lisa; University of Oxford, Nuffield Department of Medicine; University of Oxford Chen, Siyu; University of Oxford, Discovery, Nuffield Department of Medicine Barakat, Amal; World Health Organisation Regional Office for the Eastern Mediterranean Azim, Giti; Ministry of Public Health Rasoly, Bahara; Ministry of Public Health Safi, Soraya; Ministry of Public Health Flegg, Jennifer; University of Melbourne Ahmed, Nasar; Florida International University, Department of Epidemiology Ahadi, Mohmmad Jamaluddin; Ministry of Public Health Achakzai, Niaz; Ministry of Public Health AbouZeid, Alaa; Cairo University Kasr Alainy Faculty of Medicine
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Infectious diseases, Public health, Epidemiology, Evidence based practice
Keywords:	Epidemiology < INFECTIOUS DISEASES, COVID-19, Public health < INFECTIOUS DISEASES

SCHOLARONE™
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

1 **COVID-19 morbidity in Afghanistan: a nationwide, population-based**
2 **seroepidemiological study**

3 **Authors & affiliations**

4 Sayed Ataullah Saeedzai¹, Mohammad Nadir Sahak², Fatima Arifi^{2*}, Eman Abdelkareem Aly³,
5 Margo Van Gurp⁴, Lisa White⁵, Siyu Chen⁵, Amal Barkat⁶, Giti Azim¹, Bahara Rasoly¹, Soraya
6 Safi¹, Jennifer Flegg⁷, Nasar U. Ahmed⁸, Mohammad Jamaluddin Ahadi¹, Niaz M. Achakzai⁹,
7 Alaa Abouzeid¹⁰

8 ¹General Directorate of Monitoring, Evaluation and Health Information System (M&E-HIS),
9 Ministry of Public Health, Kabul, Afghanistan

10 ²WHO Health Emergencies Programme, WHO, Kabul, Afghanistan

11 ³ Information Systems for Health Unit, WHO Eastern Mediterranean Regional Office, Cairo, Egypt

12 ⁴ KIT Royal Tropical Institute, Amsterdam, the Netherlands

13 ⁵ Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of
14 Medicine, University of Oxford, United Kingdom

15 ⁶ Infectious Hazard Preparedness Unit, WHO Eastern Mediterranean Regional Office, Cairo, Egypt

16 ⁷ School of Mathematics and Statistics, University of Melbourne, Melbourne, Australia

17 ⁸ Department of Epidemiology, Robert Stempel College of Public Health, Florida International
18 University, Miami, FL

19 ⁹ Department of Molecular Biology, Forensic Medicine Directorate. Ministry of Public Health, Kabul,
20 Afghanistan

21 ¹⁰ Department of Public Health, Faculty of Medicine, Cairo University

22 *Corresponding author:

23 Fatima Arifi (Fatima.arefi@gmail.com, arifif@who.int)

24 **Word count: 4135 words**

25 List of Abbreviations

26 CoMo Consortium – COVID-19 International Modelling Consortium

27 EA – enumeration area

28 ELISA – enzyme-linked immunosorbent assay

29 IFR – infection fatality ratio

30 IgG – immunoglobulin G

31 IgM – immunoglobulin M

32 MCMC – Markov chain Monte Carlo

33 MoPH – Ministry of Public Health

34 NPI – nonpharmaceutical intervention

35 NSIA – National Statistics and Information Authority

36 R0 – basic reproduction number

37 RDT – rapid diagnostic test

38 WHO – World Health Organization

39

40 Abstract (300 words)

41 Objective

42 The primary objectives were to determine the magnitude of COVID-19 infections in the general

43 population and age-specific cumulative incidence, as determined by seropositivity and clinical

44 symptoms of COVID-19, and to determine the magnitude of asymptomatic or subclinical infections.

45 Design, setting and participants

46 We describe a population-based, cross-sectional, age-stratified sero-epidemiological study

47 conducted throughout Afghanistan during June/July 2020. Participants were interviewed to

1
2
3 48 complete a questionnaire, and rapid diagnostic tests were used to test for SARS-CoV-2 antibodies.
4
5 49 This national study was conducted in eight regions of Afghanistan plus Kabul province, considered a
6
7 50 separate region. The total sample size was 9514, and the number of participants required in each
8
9 51 region was estimated proportionally to the population size of each region. For each region, 31 to 44
10
11 52 enumeration areas (EAs) were randomly selected, and a total of 360 clusters and 16 households per
12
13 53 EA were selected using random sampling. To adjust the seroprevalence for test sensitivity and
14
15 54 specificity, and seroreversion, Bernoulli's model methodology was used to infer the population
16
17 55 exposure in Afghanistan.
18
19
20
21

22 56 Outcome measures

23
24 57 The main outcome was to determine the prevalence of current or past COVID-19 infection
25

26 58 Results

27
28 59 The survey revealed that, to July 2020, around 10 million people in Afghanistan (31.5% of the
29
30 60 population) had either current or previous COVID-19 infection. By age group, COVID-19
31
32 61 seroprevalence was reported to be 35.1% and 25.3% among participants aged ≥ 18 and 5–17 years,
33
34 62 respectively. This implies that most of the population remained at risk of infection. However, a large
35
36 63 proportion of the population had been infected in some localities, e.g. Kabul province, where more
37
38 64 than half of the population had been infected with COVID-19.
39
40
41

42 65 Conclusion

43
44 66 As most of the population remained at risk of infection at the time of the study, any lifting of public
45
46 67 health and social measures needed to be considered gradually.
47
48
49

50 68 Strengths and Limitations

- 51
52
53
54 69
- 55 70 • This is a large-scale, large sample-size, nationwide, population-based, sero-epidemiological study conducted in Afghanistan
 - 56 71 • Further analysis is conducted to adjust the seroprevalence for test sensitivity and specificity
 - 57 72 • Due to security concerns, not all areas could be surveyed where the government lacked
 - 58 73 control, and this may have affected the findings
- 59
60

- 1
2
3 74 • The findings may not reflect the current situation with regards to the new SARS-CoV-2 delta
4 75 and omicron variants of concern
5 76 • The data were entered in the DHIS2 database, which created many challenges for data
6 77 verification, household matching and the subsequent analysis.
7
8 78
9
10
11

12 79 INTRODUCTION

13
14
15
16 80 The COVID-19 pandemic has resulted in more than 248 million confirmed cases and in excess of 5
17
18 81 million deaths globally to November 2021.¹ Many countries are continuing to experience epidemic
19
20 82 waves of COVID-19, including Brazil, India and Nepal.²⁻⁴ The first reported case of COVID-19 in
21
22 83 Afghanistan was in Herat province on 24 February 2020; as of 20 July 2021, Afghanistan has reported
23
24 84 156 363 confirmed cases of COVID-19 and 7284 deaths from the disease.⁵
25
26
27

28 85 When the COVID-19 pandemic began, there were no vaccines or specific treatments available for
29
30 86 COVID-19, so nonpharmaceutical interventions (NPIs) were recommended, including social
31
32 87 distancing, home quarantine, closure of schools and universities, and bans on public gatherings.
33
34 88 Afghanistan introduced NPIs as soon as the first case of COVID-19 was detected in the country. Case
35
36 89 detection and isolation were seen as key features in helping to reduce the spread of COVID-19. With
37
38 90 the recent political transition in the country and disruption of the health system, public health and
39
40 91 social measures to tackle COVID-19 have been completely neglected, which may pose a major risk
41
42 92 for increasing the spread of COVID-19 in Afghanistan.
43
44
45

46 93 The initial focus of the Afghanistan Ministry of Public Health (MoPH) was on patients with severe
47
48 94 COVID-19 disease and ways to decrease mortality associated with the disease. Serological testing of
49
50 95 patients can be used to provide useful information about an individual's status in terms of a current
51
52 96 or previous COVID-19 infection. Immunoglobulin M (IgM) and G (IgG) antibodies arise at around the
53
54 97 same time, between 1 to 3 weeks after infection; however, IgM antibodies decay more rapidly than
55
56 98 IgG antibodies.⁶ Therefore, for public health studies, IgM is used as a marker of current infection
57
58 99 while IgG is used as a marker of previous infection, i.e. within the previous few months. There are
59
60

1
2
3 100 various rapid diagnostic tests (RDTs) available that can be used to simultaneously test blood samples
4
5 101 for IgM/IgG antibodies against COVID-19.
6
7

8 102 Due to the limited testing and surveillance capacity in Afghanistan, it seemed likely there was
9
10 103 considerable under-reporting of cases and deaths; therefore, robust scientific studies are required to
11
12 104 determine the actual burden of COVID-19 in the country. Serological studies can be used to estimate
13
14 105 levels of past exposure and thus position a population in their epidemic timeline. However, serology
15
16 106 results might underestimate the total exposure in a population⁷ because of decaying antibody titres
17
18 107 over time.⁸⁻¹⁰ Here, we describe a national seroepidemiological survey initiated by the MoPH and
19
20 108 conducted throughout Afghanistan between June and July 2020, involving a questionnaire survey
21
22 109 and antibody testing of participants for COVID-19 infection using RDTs. The primary objectives of the
23
24 110 study were: 1) to determine the magnitude of COVID-19 infection in the general population and age-
25
26 111 specific cumulative incidence, as determined by seropositivity and clinical symptoms of COVID-19;
27
28 112 and 2) to determine the magnitude of asymptomatic or subclinical infections. The World Health
29
30 113 Organization (WHO) protocol for population-based age-stratified seroepidemiological investigations
31
32 114 for COVID-19 infection was adapted for the Afghanistan context to obtain seroprevalence
33
34 115 estimates.¹¹ To adjust the seroprevalence for test sensitivity and specificity, as well as seroreversion,
35
36 116 we further adapted a methodology¹² that was originally developed for the England setting and used
37
38 117 this to infer the population exposure and undocumented mortality associated with COVID-19 in
39
40 118 Afghanistan.
41
42
43
44
45
46
47

48 119 **METHODS**

52 120 **Patient and Public Involvement statement**

53
54
55
56 121 As this study was not a clinical trial and it did not involve patients, no members of the public or
57
58 122 patients were directly involved . The study results were disseminated through public workshops in
59
60

1
2
3 123 universities, seminars and workshops, and through the media for the general public. Consent was
4
5 124 obtained to be included in the study, and any personal identifier information was excluded during
6
7
8 125 data processing and analysis.
9

10 11 126 **Ethical considerations**

12
13
14 127 Ethical and technical clearance to conduct the survey was obtained from the Institutional Review
15
16 128 Board of the Afghanistan MoPH, **Reference number: A.0321.0278**. Informed consent was obtained
17
18 129 from participants aged ≥ 18 years, and assent from family members was obtained for those aged 5–
19
20 130 17 years. Individuals who did not provide consent were excluded. Survey team members provided
21
22
23 131 advice about home isolation to participants who tested IgM-positive for COVID-19 during the survey.
24
25

26 27 132 **Study design**

28
29
30 133 This was a population-based, cross-sectional, age-stratified seroepidemiological study. Participants
31
32 134 were interviewed to complete a questionnaire, and RDTs were used to test for SARS-CoV-2
33
34 135 antibodies. The survey was conducted during June and July 2020.
35
36

37 38 136 **Population and sampling**

39
40
41 137 This was a national study conducted in the eight regions of Afghanistan plus Kabul province, which
42
43 138 was considered as a separate region, making nine regions in total (online supplemental figure S1).
44
45 139 The total sample size was 9514 and the number of participants required in each region was
46
47 140 estimated proportionate to the population size of each region (online supplemental Table S1). Two-
48
49 141 stage cluster sampling was used. In the first stage, an updated list of enumeration areas (EAs) was
50
51 142 used as the study sampling frame, with 31 to 44 EAs (clusters) randomly selected per region,
52
53 143 resulting in a total of 360 clusters. Due to time constraints and to ensure data validity, insecure or
54
55 144 inaccessible EAs were excluded from the study.
56
57
58
59
60

1
2
3 145 In the second stage, all households in an EA were listed and 16 households per EA were selected
4
5 146 using a random sampling table). For the age-stratification, two individuals from each household
6
7 147 were randomly selected for testing: one aged 5–17 years and one aged ≥18 years.
8
9

10 11 148 **Serological testing**

12
13
14 149 Finger-prick blood samples were collected from the randomly selected household members in each
15
16 150 age category. The antibody RDTs for COVID-19 were performed in the presence of the participant,
17
18 151 and the results were shared with them. The COVID-19 RDT used was the COVID-19 IgG/IgM Rapid
19
20 152 Test Cassette developed by Healgen Scientific LLC, USA. The RDT is US Food and Drug Administration
21
22 153 (FDA)-authorised, with IgM relative sensitivity and specificity of 95.7% and 97.3%, respectively; IgG
23
24 154 relative sensitivity and specificity of 91.8% and 96.4%, respectively; and both IgG-positive and/or
25
26 155 IgM-positive specificity of 97.5%.
27
28
29

30 31 156 **Data collection and analysis**

32
33
34 157 The survey used a validated questionnaire that was initially piloted in Kabul province. All participants
35
36 158 were interviewed by the survey team members, who completed a questionnaire that included
37
38 159 questions about the demographics of each participant and their household members, their history of
39
40 160 exposure to COVID-19, and deaths in the family during the 15-month period beginning in March
41
42 161 2019.
43
44
45

46 162 Data collection teams comprised two members, one male and one female; there were 191 teams in
47
48 163 total. Due to the need for blood-drawing for samples, the team members were either nurses,
49
50 164 midwives or laboratory technicians.
51
52

53
54 165 Regional COVID-19 data were entered into DHIS2 (District Health Information Software-2) by disease
55
56 166 surveillance officers in the provinces. DHIS2 is the national data warehouse for Afghanistan's health
57
58 167 information and includes data that inform the country's COVID-19 dashboard.⁵ Various steps were
59
60

1
2
3 168 taken for data quality assurance at both regional and central levels within the MoPH; data collection
4
5 169 teams were monitored by master trainers in the regional capitals and by disease surveillance staff in
6
7 170 the provinces. Prior to being entered into the system, questionnaires were quality checked and
8
9
10 171 some participants whose phone numbers were available in the questionnaire were contacted at
11
12 172 random by phone call to confirm that their details were correct.

13
14
15 173 Data were imported into STATA version 15¹³ for the statistical analyses. To ensure a representative
16
17 174 sample and results, weighted analysis was applied to adjust for the complex survey design. Sample
18
19 175 weighting, non-response weighting and post-stratification weighting were performed. The
20
21 176 proportions of infections and 95% confidence intervals were calculated and adjusted to take the
22
23 177 survey design into account. The H0 was tested against alternative/research hypothesis at there are
24
25 178 differences in prevalence COVID among social demographic and regional characteristics. To
26
27 179 determine the overall levels of current and past infection of COVID-19, individuals who tested
28
29 180 positive for IgG, IgM or both were summed. To determine the incidence of COVID-19 during the
30
31 181 survey period, IgM positivity alone was used.

32 33 34 35 36 182 **Adjustment of seroprevalence and exposure inference**

37
38
39 183 We first used a simple Bernoulli model to estimate the regional SARS-CoV-2 (the virus that causes
40
41 184 COVID-19) seroprevalence, after adjusting the proportion of individuals in each region with current
42
43 185 or past COVID-19 infection according to the sensitivity and specificity of the serology test.¹⁴ (The
44
45 186 term 'seroprevalence' below denotes the serology-positive ratio already adjusted by the test.)
46
47 187 Further details of the method used can be found in the online supplemental method, appendix 1.
48
49 188 We revised the mathematical model¹² to estimate the total exposure in the population by region
50
51 189 after taking into account waning antibody levels. Further details of the method used can be found in
52
53 190 the online supplemental method, appendix 1.
54
55
56
57
58
59
60

191 RESULTS

192 Demographic details

193 This seroepidemiological study has provided estimates of the prevalence of SARS-CoV-2 antibodies
194 across Afghanistan, in urban and rural areas, and in the nine regions of the country. Of the 360
195 clusters identified for participation in the study, 338 (94%) were included; the remainder were
196 excluded due to insecure or inaccessible EAs and time constraints. Similarly, of the total planned
197 5408 households in 338 clusters, 5177 (96%) households completed the survey. A total of 9514
198 household members from these 338 clusters were interviewed and tested for COVID-19. The mean
199 age of respondents was 27 years, 53.9% were male and 46.1% were female, 73% were from rural
200 areas (online supplemental table S2), and most participants (79.2%) were married.

201 COVID-19 infections in Afghanistan

202 The total proportion of COVID-19 infections (including all positive results, the average of both
203 current and past infection) for the whole of Afghanistan was 31.5%. By region, Kabul had the highest
204 proportion of COVID-19 infections (53%), while the Central highlands region had the lowest
205 proportion, at 21.1% (figure 1).

206 Based on further analysis, the adjusted seroprevalence by region was consistent with the serosurvey
207 results. Kabul still had the highest adjusted seroprevalence (51.8%) (table 1 and figure 2).

208 RDT results for participants aged 18 years or more

209 In total, 5618 participants aged ≥ 18 years were interviewed and tested for this survey. Among this
210 age group, 2056 (35.1%) of individuals tested positive for antibodies against SARS-CoV-2 (table 2).

211 There were 885 (37.2%) females and 1170 (33.9%) males who tested positive, and there was a
212 higher proportion of positive tests in individuals who lived in urban areas compared with the

213 proportion in people who lived in rural areas (773, 42.3% versus 1323, 31.7%, respectively) (table 2).
 214 Kabul region had the highest proportion of participants aged ≥ 18 years who tested positive for
 215 antibodies against SARS-CoV-2 (357, 56.8%) (table 2). The survey results revealed that 164 (2.6%) of
 216 participants aged ≥ 18 years were IgM-positive for COVID-19, i.e. they had a current infection, with
 217 the highest proportion of current infections in the South-east region (37, 7.0%) (table 1).

218 **Table 1. Seroprevalence of SARS-CoV-2 antibodies and proportion of IgM-seropositive in**
 219 **participants aged ≥ 18 years by region, area of residence and sex**

	Number of positive COVID-19 tests [#]	Seroprevalence % [95% CI]	Adjusted seroprevalence [95% CI]	Number of IgM-positive COVID-19 tests	IgM-seropositive % COVID-19 tests [95% CI]
National	2056	35.1 [31–39.5]	29.8 [28.8, 30.7]	164	2.6 [2.0–3.5]
Region***				Region* **	
Central	254	45.5 [37.8–53.4]	34.6 [31.6, 37.6]	28	4.3 [2.4–7.6]
Central highlands	105	24.9 [17.9–33.7]	19.0 [16.4, 21.8]	5	1.0 [0.4–2.3]
East	294	49.1 [41.5–56.8]	41.5 [38.6, 44.4]	16	2.5 [1.4–4.5]
Kabul	357	56.8 [52.0–62.0]	51.8 [48.8, 54.8]	17	2.7 [1.4–5.0]
North	212	35.3 [28.1–43.4]	28.9 [26.3, 31.8]	7	1.4 [0.6–3.4]
North-east	263	39.3 [31.9–47.4]	30.7 [28.1, 33.3]	26	4.0 [2.1–7.8]
South	115	26.6 [19.0–36.0]	23.9 [20.7, 27.1]	8	1.6 [0.7–3.4]
South-east	221	40.9 [34.4–47.9]	30.5 [27.4, 33.6]	37	7.0 [3.7–12.9]
West	235	39.8 [34.8–45.1]	32.4 [29.7, 35.2]	20	3.4 [1.8–6.3]
Area of residence					
Rural		31.7 [26.5–37.4]		121	3.7 [1.7–7.9]
Urban		42.3 [35.7–49.2]		43	2.3 [1.2–4.2]
Sex					
Male	1170	33.9 [29–39.2]		104	2.4 [1.4–4.0]
Female	885	37.2 [32–42.6]		60	4.1 [1.8–9.2]
Age (years)					
18–39	1109	33.7 [28.5–39.2]		96	2.7 [1.9–3.9]
40–59	657	36.5 [31.9–41.3]		50	2.4 [1.6–3.7]
60+	290	40.0 [31.8–48.2]		18	2.1 [1.1–4.2]

220 [#]The total number of positive COVID-19 tests includes all positive results: both current and past infections i.e. IgG-positive,
 221 IgM-positive or both. * $p < .05$, ** $p < .01$, *** $p < .001$

222 CI, confidence interval

224 RDT results for participants aged 5 to 17 years

225 There were 4346 participants aged 5–17 years interviewed and tested for this survey. Among this
 226 age group, a total of 850 (25.3%) individuals tested positive for antibodies against SARS-CoV-2 (table
 227 2), 401 (27.8%) females and 446 (24.2%) males. Again, there was a higher proportion of positive
 228 tests in individuals who lived in urban areas compared with the proportion among people who lived

229 in rural areas (322, 30.8% versus 528, 23.4%, respectively) (table 2). Kabul region had the highest
 230 proportion of participants aged 5–17 years who tested positive for antibodies against SARS-CoV-2
 231 (177, 46.4%) (table 2). There were 89 (3.3%) participants aged 5–17 years who were IgM-positive for
 232 COVID-19, with the highest proportion of current infections in the South region (7, 4.7%) (table 2).

233

234 **Table 2. Seroprevalence of SARS-CoV-2 antibodies and proportion of IgM-seropositive results in**
 235 **participants aged 5–17 years by region, area of residence and sex.**

	Number of positive COVID-19 tests [#]	Seroprevalence % [95% CI]	Number of IgM-positive COVID-19 tests	IgM-seropositive % COVID-19 tests [95% CI]
National	850	25.3 [20.5–30.8]	89	3.3 [1.8–6.3]
Region*			Region**	
Central	79	21.0 [14.5–29.3]*	10	2.8 [1.2–6.3]
Central highlands	42	14.6 [8.6–23.8]	3	1.6 [0.4–6.6]
East	172	32.4 [26.8–38.6]	10	1.4 [0.7–3.1]
Kabul	177	46.4 [40.8–52.1]	14	3.5 [1.6–7.3]
North	96	23.0 [16.8–30.8]	6	1.2 [0.4–3.7]
North-east	108	20.9 [15.1–28.2]	18	2.8 [1.0–7.6]
South	55	24.4 [14.5–38.0]	7	4.7 [1.6–13.1]
South-east	42	17.6 [10.6–27.6]	9	2.4 [0.8–6.8]
West	79	24.5 [18.4–31.8]	12	3.2 [1.7–6.0]
Area of residence				
Rural	528	23.4 [17.5–30.6]	60	3.7 [1.7–7.9]
Urban	322	30.8 [24.8–37.5]	29	2.3 [1.2–4.2]
Sex				
Male	446	24.2 [18.5–31]	47	2.4 [1.4–4.0]
Female	401	27.8 [21.3–33]	42	4.1 [1.8–9.2]
Age (years) **			Age **	
5–9	175	[13.4–26.2]	20	3.3 [1.1–9.5]
10–14	365	[20.8–33.8]	40	3.7 [1.7–7.9]
15–17	310	[23.5–35.6]	29	2.8 [1.5–5.2]

236 [#]The total number of positive COVID-19 tests includes all positive results: both current and past infections i.e. IgG-positive,
 237 IgM-positive or both. *p<.05, **p<.01, *** p<.001

238 CI, confidence interval

239

240 Predictions for cumulative exposure in the population up to 21 July 2020 in the nine regions of

241 Afghanistan are shown in figure 3. The method used for the modelling analysis, which was

242 developed by the COVID-19 International Modelling Consortium (CoMo Consortium), is detailed in

243 the online supplemental method, appendix 1.

1
2
3 244 The solid orange circles and black error bars in the panel for each region represent the observed
4
5 245 seroprevalence data and the associated credible interval (CrI) after adjusting for the sensitivity and
6
7 246 specificity of the antibody test. The green and orange lines show the median predictions for
8
9 247 exposure and seroprevalence, respectively, while the shaded areas correspond to 95% CrI. The
10
11 248 median predicted exposure levels by region (expressed as the proportion of the population that has
12
13
14 249 been infected) as of 21 July 2020 are shown on the map of Afghanistan.

16 251 **DISCUSSION**

17
18
19
20
21 252 This national survey of COVID-19 morbidity in Afghanistan, which was conducted during June and
22
23 253 July 2021, revealed that around 10 million people (31.5% of the population) were seropositive for
24
25 254 antibodies against SARS-CoV-2, reflecting either current or previous COVID-19 infection. The
26
27 255 population of Afghanistan is estimated to comprise approximately 33.6 million people.¹⁵ Our finding
28
29 256 is reasonably consistent with the results of a telephone survey conducted before July 2020 with a
30
31 257 randomly selected sample of 713 healthcare workers to estimate COVID-19 morbidity in the country.
32
33
34 258 The estimated proportion of individuals who had experienced COVID-19 signs and symptoms was
35
36 259 49.6%, which is close to the value for total infections for most regions reported in the present study,
37
38 260 however, no laboratory testing was conducted for the phone survey, which only collected clinical
39
40 261 information about symptoms. There is a discrepancy between our serosurvey results and the
41
42 262 detected number of COVID-19 infections reported to the surveillance system in the country (36 710
43
44 263 cases reported by the surveillance system as of 30 July 2020 and 156 363 cases as of 5 November
45
46 264 2021). The under-reporting of COVID-19 cases is a problem globally due to limited testing
47
48 265 availability, flawed test sensitivity, poor surveillance and the indeterminate proportion of
49
50 266 asymptomatic infections.¹⁶ However, some studies have suggested a lower prevalence of COVID-19
51
52 267 in countries during a similar period.¹⁷ For example, the upper bound of COVID-19 prevalence was
53
54 268 estimated to be 8.2% in Spain, 6.8% in Italy and 6.1% in the UK. However, the contexts, social mixing
55
56 269 and other factors for the demographic scaling model vary across countries, particularly in resource-

1
2
3 270 limited countries. In such contexts, there are close contacts at home due to large family sizes, while
4
5 271 social mixing in schools, communities and society might be more frequent as people rely on daily
6
7 272 wages, and the adopted COVID-19 control measures might be less enforced and effective in such
8
9 273 settings. Population-based seroprevalence studies are helpful to identify the true burden of disease,
10
11 274 which might be higher compared with the burden estimated by modelling studies.
12
13
14

15 275 A modelling exercise was performed using the CoMo model to estimate the peak incidence of
16
17 276 COVID-19 in Afghanistan. The CoMo model was developed by the CoMo Consortium.¹⁸ The CoMo
18
19 277 Consortium adopts a participatory modelling approach,¹⁹ which places in-country subject matter
20
21 278 experts at the forefront of model development to ensure that contextual considerations, such as
22
23 279 local infrastructure, human resources and sociocultural considerations, are fully taken into account.
24
25 280 The CoMo model was used to estimate the peak incidence of COVID-19 in Afghanistan under four
26
27 281 scenarios: good, bad, very bad and appropriate, depending on the coverage of and adherence to the
28
29 282 NPIs. If the use of NPIs (in a very bad scenario) is not considered, then the COVID-19 peak was
30
31 283 predicted to occur in June 2020, with an estimated 69.6% of the population infected and 20 509
32
33 284 deaths by the end of 2020.
34
35
36
37

38 285 In communicable disease epidemiology, one of the key parameters used in decision-making is the
39
40 286 estimate of herd immunity in a population. Herd immunity occurs when a certain proportion of the
41
42 287 population is immune to a given infectious disease, reducing the probability that the disease will be
43
44 288 transmitted from one individual to another, thus helping to protect the entire population from that
45
46 289 disease.²⁰ Herd immunity can be achieved either through individuals being exposed or vaccinated.
47
48 290 Determining a country's herd immunity threshold to a given disease is directly related to estimates
49
50 291 of the basic reproductive number, R_0 , of that disease. R_0 indicates the average number of
51
52 292 individuals one infected individual can go on to infect in a fully susceptible population. Different herd
53
54 293 immunity thresholds in different contexts have been estimated for COVID-19, ranging from 43% to
55
56 294 85%.²⁰⁻²⁵ For example, one study indicated that if $R_0=3$, i.e. one infected individual can infect up to
57
58
59
60

1
2
3 295 three others, 67% of the population must be immune to achieve herd immunity.²¹ Estimates by
4
5 296 Johns Hopkins University suggest that 70% of the population must be immune to achieve herd
6
7 297 immunity and end restrictions on people's day to day lives²⁰, while another study suggested that R0
8
9
10 298 values of 1–2, 2–4 and >4 would require herd immunity thresholds of 50%, 56.1–74.8% and 77.9–
11
12 299 85%, respectively.²² In addition to R0 and the herd immunity threshold, the rate of antibody decline
13
14 300 post-infection must also be considered, with one study suggesting that antibodies to COVID-19
15
16 301 decline within 94 days of infection.¹⁰

17
18
19 302 A study conducted by Eckerle and Meyer revealed that by mid-2020, an insufficient proportion
20
21 303 of the population had been infected globally to achieve herd immunity, and these findings were
22
23 304 confirmed by reports of low COVID-19 morbidity levels from countries such as Sweden, where
24
25 305 an infection rate of 7% was reported by the end of April despite no lockdown; the mentioned
26
27 306 study also states that obtaining herd immunity by exposing the population to the disease results
28
29 307 in the simultaneous infection of the majority of the population and paves the way for a second
30
31 308 wave of the disease.

32
33
34
35
36 309 These estimates of herd immunity thresholds suggest that the present survey findings, of a SARS-
37
38 310 CoV-2 antibody seroprevalence of approximately 32% among the population in Afghanistan, mean
39
40 311 that less than half of population was infected and most of the country's population remained at risk
41
42 312 of infection. However, in some provinces, large numbers of individuals have been infected and
43
44 313 recovered from COVID-19. In Kabul province, for example, more than half of the population has
45
46 314 been infected. However, as the majority of the population remains at risk of infection, preventive
47
48 315 measures and NPIs should be lifted gradually, as per WHO guidelines.²⁶ It should also be noted that
49
50 316 this survey was conducted at a time when the SARS-CoV-2 alpha variant was the most prevalent
51
52 317 variant in Afghanistan; it is unclear what effect the arrival of new variants, such as the delta and
53
54 318 omicron variants, and vaccination will have on population immunity.
55
56
57
58
59
60

1
2
3 319 As in many low- and middle-income countries, COVID-19 vaccination rates in Afghanistan are low,
4
5 320 with just 12% of the population currently fully vaccinated.⁵ With the disruptions to the health system
6
7 321 as a result of the evolving political situation in the country, the COVID-19 response may deteriorate
8
9 322 if control measures are not implemented and vigilantly maintained.

10
11
12
13 323 Based on the evidence outlined above, the NPIs currently in place in Afghanistan should not have
14
15 324 been lifted, as large numbers of the population are yet to become immune through natural infection
16
17 325 or vaccination. If the NPIs are lifted, the rates of hospitalisation will increase, as will the number of
18
19 326 patients requiring ventilation; this will place the already fragile health system under considerable
20
21 327 pressure. However, after July 2021, the restrictions were reduced and since then the country has
22
23 328 focused on school closures alone as a mitigation measure to balance the economy, social life and the
24
25 329 impact of COVID-19 on the health system. It is worth mentioning that with the recent transition of
26
27 330 government in Afghanistan and decreased funding for the country's health system, there are
28
29 331 evolving challenges that will ultimately lead to the increased spread of COVID-19 and other
30
31 332 infectious diseases. Greater levels of poverty, a displaced population and poor sanitation will further
32
33 333 exacerbate this problem. The influx of refugees from Afghanistan to other countries might also
34
35 334 facilitate the cross-border spread of disease. Particularly with the emergence of new variants and
36
37 335 low vaccination coverage, it is crucial to have continued public health and social measures to
38
39 336 mitigate the impact of COVID-19 in a conflict-affected and unstable country. For health services to
40
41 337 continue, functional hospitals, surveillance systems and laboratories, as well as a skilled healthcare
42
43 338 workforce, are needed to mitigate the spread of COVID-19 and other infections within Afghanistan
44
45 339 and prevent the regional and even global spread of disease.

46
47
48
49
50
51 340 This study had some limitations. First, the time available to conduct the survey was limited. Second,
52
53 341 security concerns meant that not all areas could be surveyed; the inability to conduct proper
54
55 342 household listing and create maps for enumeration areas in those areas where the government
56
57 343 lacked control may have affected the findings. Third, the findings may not reflect the current
58
59
60

1
2
3 344 situation with regards to the new SARS-CoV-2 delta and omicron variants of concern, as the R0 for
4
5 345 these variants is not well-established. Once a stable estimate of the R0 for these variants has been
6
7 346 established then our findings can be adjusted accordingly to assist with programme planning.
8
9
10 347 Fourth, the data were entered in the DHIS2 database, which created many challenges for data
11
12 348 verification, household matching and the subsequent analysis. All data were re-entered in DHIS2 at
13
14 349 the central level to ensure data quality and to match the households for reliable and valid analysis.
15
16 350 In future surveys, it would be preferable to collect data by entering them directly via a tablet or
17
18 351 similar appropriate research data entry tool to improve the data quality.
19
20
21 352
22
23
24

25 353 **CONCLUSION**

26
27
28
29 354 Although the immunity threshold may have been reached in some localities within Afghanistan,
30
31 355 specifically Kabul, this threshold has not yet been reached among the country's entire population. In
32
33 356 particular, the proportion of the population that is seropositive for antibodies against SARS-CoV-2 is
34
35 357 much lower in rural areas than urban areas. The seroprevalence represents a lower estimate of herd
36
37 358 immunity and the predicted exposure represents an upper limit. Given the large proportion of the
38
39 359 population that remains susceptible to COVID-19 infection, and the limited COVID-19 vaccination
40
41 360 coverage, NPIs and vigilance should remain in place to protect the health system from an
42
43 361 unmanageable burden of hospitalisations. The link between the presence of antibodies and
44
45 362 immunity has yet to be established, as has the link between prior exposure and immunity. As
46
47 363 antibody levels wane, seroprevalence may provide an underestimate of immunity but, conversely, if
48
49 364 immunity wanes, then prior exposure would provide a higher estimate of immunity.
50
51
52
53
54
55
56
57
58
59
60

365 **Data availability statement**

366 Survey serology data are stored in the Ministry of Public Health national database and are available
367 upon reasonable request. All data, code and materials used in the analyses can be accessed at:
368 <https://github.com/SiyuChenOxf/AfghanistanSerologyStudy/tree/master>. All parameter estimates
369 and figures 3 and 4 can be reproduced using the code provided. This work is licensed under a
370 Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use,
371 distribution and reproduction in any medium, provided the original work is properly cited.

372 **Acknowledgments**

373 This SARS-CoV-2 Seroprevalence in Afghanistan: A Population-based Study has been supported by
374 WHO Unity Studies, a global sero-epidemiological standardization initiative, with funding to WHO by
375 the COVID-19 Solidarity Response Fund and the German Federal Ministry of Health (BMG) COVID-19
376 Research and Development Fund. We would like to acknowledge the Monitoring and Evaluation and
377 Health Information System General Directorate, which led the study in Afghanistan. The study was
378 funded by Ministry of Public Health. We would also like to acknowledge the regional master trainers
379 and the National Statistics and Information Authority (NSIA) team for their support in conducting the
380 survey. We would additionally like to thank the following for their efforts during the survey: the WHO
381 team, USAID, World Bank, CDC Atlanta, Performance Management Office (PMO), Grants and
382 Contracts Management Unit (GCMU), Afghanistan National Public Health Institute (ANPHI), all
383 Provincial Public Health Directorates (PPHDs), Health Promotion Department (HPD), all Basic Package
384 of Health Services (BPHS)- and Essential Package of Hospital Service (EPHS)-implementing NGOs, the
385 survey steering committee team, and the data management team. We acknowledge the national and
386 provincial surveillance teams for their extensive support in conducting the survey. We also appreciate
387 all those individuals who supported the data collection, data entry, and listing of households for the

1
2
3 388 survey frame. Medical writing assistance and editorial support was provided by Adam Bodley,
4
5 389 according to Good Publication Practice guidelines. The views expressed in this manuscript are those
6
7 390 of the authors and do not necessarily reflect the views of the authors' affiliated organisations or the
8
9 391 organisations listed here.
10
11
12
13

14 392 **Funding Statement:**

15
16
17 393 There are no funders to report for this submission.
18
19
20
21

22 394 **Authorship & Contributorship statement**

23
24
25 395 S.A. Seedzai, planned the study, lead the study design and tools in the country, coordinated with
26 396 national authorities, provided Training of Trainers (ToT), monitored process of data collection,
27 397 analysed the data and contributed to writing and finalising the report and findings of the study and
28 398 organised dissemination workshops, seminars in the country.
29
30

31 399 M.N. Sahak, supported in planning the study, supported in scientific study design, coordinated
32 400 internally for sharing scientific WHO tools and protocol, supported in data analysis, contributed to
33 401 writing and finalising the study report.
34

35 402 F. Arifi, supported in the design and tools of the study, contributed in analysing and interpreting the
36 403 results, supported in drafting the report and communicated with all authors for compiling and
37 404 incorporating the feedback in the final version.
38

39 405 E. Aly, supported in the study design and tools, contributed to data cleaning, analysis &
40 406 interpretation, and writing the report and finalisation of the study report.
41
42

43 407 M.V. Gulp, supported in study design, complex and multi-level analysis, reviewed and revised the
44 408 study findings and report.
45

46 409 L.J. White, supported in study method, analysis, interpretation of findings and modelling based on
47 410 the serosurvey data, reviewed and revised the final study report.
48

49 411 S. Chen, supported in study method, analysis, interpretation of findings and modelling based on the
50 412 serosurvey data, reviewed and revised the final study report.
51

52 413 A. Barakat, supported in proper laboratory testing of the study method, selection of testing kits,
53 414 contributed to the interpretation of findings, reviewed the draft report and provided inputs to the
54 415 revision and finalisation of the study report.
55

56 416 G. Azim, supported in planning the study, facilitated internal coordination within MoPH, provided
57 417 ToT training, supervised the data collection in a cluster, supported in drafting the study report and
58 418 contributed to finalising the study findings.
59
60

1
2
3 419 B. Rasoly, supported in planning the study, provided ToT training, supervised the data collection in a
4 420 cluster, supported data entry and cleaning, drafting the study report and contributed to finalising
5 421 the study findings.

7 422 S. Safi, supported in planning the study, provided ToT training, supervised the data collection in a
8 423 cluster and collected data where needed, supported data entry and cleaning, drafting the study
9 424 report and contributed to finalising the study findings.

11 425 J. Flegg, contributed to the study method, interpretation of serosurvey findings for the context,
12 426 analysing the data, reviewed the draft report and provided inputs to the finalisation of the study
13 427 report.

16 428 N. Ahmed, contributed to the study method, interpretation of serosurvey findings for the context,
17 429 analysing the data, reviewed the draft report and provided inputs to the finalisation of the study
18 430 report.

20 431 M.J. Ahadi, supported the planning of the study, provided ToT training, supervised the data
21 432 collection in clusters and collected data where needed, supported data entry and cleaning, drafting
22 433 the study report and contributed to finalising the study findings.

24 434 N.M. Achakzai, supported correct laboratory testing protocol of the study method, selection of
25 435 testing kits, contributed to interpretation of findings, reviewed the draft report and provided inputs
26 436 to the revision and finalisation of the study report.

29 437 A. AbouZeid, supported the scientific study design, internal WHO coordination, supported data
30 438 analysis, and contributed to writing and finalising the study report

32 439 **Competing interests**

34 440 No authors declare any competing interests.

37 441 **REFERENCES**

- 38 442
- 42 443 1. COVID-19 dashboard. Johns Hopkins COVID-19 dashboard Johns Hopkins University; 2021
43 444 [Available from: <https://coronavirus.jhu.edu/map.html>.
 - 44 445 2. Burki T. No end in sight for the Brazilian COVID-19 crisis. *Lancet Microbe* 2021;2(5):e180. doi:
45 446 10.1016/s2666-5247(21)00095-1 [published Online First: 2021/05/11]
 - 46 447 3. Thiagarajan K. Why is India having a covid-19 surge? *BMJ* 2021;373:n1124. doi:
47 448 10.1136/bmj.n1124
 - 48 449 4. ReliefWeb. Second wave of Covid-19 causes alarm: ReliefWeb; 2021 [Available from:
49 450 <https://reliefweb.int/report/nepal/second-wave-covid-19-causes-alarm>.
 - 50 451 5. Ministry of Public Health. COVID-19 Dashboard 2021 [Available from: <http://covid.moph-dw.org/>
51 452 accessed 11/06/2021 2021.
 - 52 453 6. CDC. Interim Guidelines for COVID-19 Antibody Testing Atlanta: Centers for Disease Control and
53 454 Prevention; 2021 [Available from: [https://www.cdc.gov/coronavirus/2019-
54 455 ncov/lab/resources/antibody-tests-guidelines.html2021](https://www.cdc.gov/coronavirus/2019-ncov/lab/resources/antibody-tests-guidelines.html2021).
 - 55 456 7. Burgess S, Ponsford MJ, Gill D. Are we underestimating seroprevalence of SARS-CoV-2? *BMJ*
56 457 2020;370:m3364. doi: 10.1136/bmj.m3364
- 58
59
60

- 1
2
3 458 8. Ibarondo FJ, Fulcher JA, Goodman-Meza D, et al. Rapid Decay of Anti-SARS-CoV-2 Antibodies in
4 459 Persons with Mild Covid-19. *New England Journal of Medicine* 2020;383(11):1085-87. doi:
5 460 10.1056/NEJMc2025179
6 461 9. Harris RJ, Whitaker HJ, Andrews NJ, et al. Serological surveillance of SARS-CoV-2: Six-month trends
7 462 and antibody response in a cohort of public health workers. *The Journal of infection*
8 463 2021;82(5):162-69. doi: 10.1016/j.jinf.2021.03.015 [published Online First: 2021/03/22]
9 464 10. Seow J, Graham C, Merrick B, et al. Longitudinal observation and decline of neutralizing antibody
10 465 responses in the three months following SARS-CoV-2 infection in humans. *Nature*
11 466 *Microbiology* 2020;5(12):1598-607. doi: 10.1038/s41564-020-00813-8
12 467 11. WHO. Population-based age-stratified seroepidemiological investigation protocol for COVID-19
13 468 virus infection, 17 March 2020. Geneva: WHO, 2020.
14 469 12. Chen S, Flegg JA, White LJ, et al. Levels of SARS-CoV-2 population exposure are considerably
15 470 higher than suggested by seroprevalence surveys. *medRxiv* 2021:2021.01.08.21249432. doi:
16 471 10.1101/2021.01.08.21249432
17 472 13. StataCorp. Stata Statistical Software: Release 15 College Station, TX: StataCorp LLC; 2017
18 473 [Available from: <https://www.stata.com/stata15/>.
19 474 14. Healgen Scientific LLC. COVID-19 IgG/IgM Rapid Test Cassette (Whole Blood/Serum/Plasma)
20 475 Instruction for Use 2020 [Available from: <https://www.fda.gov/media/138438/download>.
21 476 15. NSIA. Estimated Population of Afghanistan 2021-22: NSIA; 2021 [Available from:
22 477 [https://www.nsia.gov.af:8080/wp-content/uploads/2021/06/Estimated-Population-2-of-](https://www.nsia.gov.af:8080/wp-content/uploads/2021/06/Estimated-Population-2-of-AFG-New-Design-14.pdf)
23 478 [AFG-New-Design-14.pdf](https://www.nsia.gov.af:8080/wp-content/uploads/2021/06/Estimated-Population-2-of-AFG-New-Design-14.pdf).
24 479 16. Wu SL, Mertens AN, Crider YS, et al. Substantial underestimation of SARS-CoV-2 infection in the
25 480 United States. *Nature Communications* 2020;11(1):4507. doi: 10.1038/s41467-020-18272-4
26 481 17. Bohk-Ewald C, Dudel C, Myrskylä M. A demographic scaling model for estimating the total
27 482 number of COVID-19 infections. *Int J Epidemiol* 2021;49(6):1963-71. doi:
28 483 10.1093/ije/dyaa198 [published Online First: 2020/12/23]
29 484 18. CoMo Consortium. CoMo Consortium: BMJ; 2021 [Available from: <https://como.bmj.com/2021>.
30 485 19. Aguas R, White L, Hupert N, et al. Modelling the COVID-19 pandemic in context: an international
31 486 participatory approach. *BMJ Glob Health* 2020;5(12) doi: 10.1136/bmjgh-2020-003126
32 487 [published Online First: 2020/12/29]
33 488 20. D'Souza G, Dowdy D. What is Herd Immunity and How Can We Achieve It With COVID-19?
34 489 Maryland: Johns Hopkins Bloomberg School of Public Health; 2021 [Available from:
35 490 <https://www.jhsph.edu/covid-19/articles/achieving-herd-immunity-with-covid19.html2021>.
36 491 21. Randolph HE, Barreiro LB. Herd Immunity: Understanding COVID-19. *Immunity* 2020;52(5):737-
37 492 41. doi: 10.1016/j.immuni.2020.04.012 [published Online First: 2020/05/21]
38 493 22. Kwok KO, Lai F, Wei WI, et al. Herd immunity - estimating the level required to halt the COVID-19
39 494 epidemics in affected countries. *J Infect* 2020;80(6):e32-e33. doi: 10.1016/j.jinf.2020.03.027
40 495 [published Online First: 2020/03/27]
41 496 23. Gomes MGM, Corder RM, King JG, et al. Individual variation in susceptibility or exposure to SARS-
42 497 CoV-2 lowers the herd immunity threshold. *medRxiv* 2020:2020.04.27.20081893. doi:
43 498 10.1101/2020.04.27.20081893
44 499 24. Altmann DM, Douek DC, Boyton RJ. What policy makers need to know about COVID-19
45 500 protective immunity. *Lancet* 2020;395(10236):1527-29. doi: 10.1016/s0140-6736(20)30985-
46 501 5 [published Online First: 2020/05/01]
47 502 25. Britton T, Ball F, Trapman P. The disease-induced herd immunity level for Covid-19 is
48 503 substantially lower than the classical herd immunity level. *arXivorg* 2020:
49 504 arXiv:2005.03085v1.
50 505 26. World Health Organization. Considerations for implementing and adjusting public health and
51 506 social measures in the context of COVID-19. Geneva: World Health Organization, 2021.
52
53
54
55
56
57
58
59 507
60

1
2
3 5084
5 509 **Figure Captions/Legends**6
7
8 510 Figure 1. Seroprevalence of SARS-CoV-2 antibodies (including all positive results: IgG-positive, IgM-
9 positive or both) among all age groups by region in Afghanistan.10
11
12
13 512 Figure 2. Adjusted seroprevalence by region by the sensitivity and specificity of the serology test for
14 IgG-positive and/or IgM-positive.15 513
16
17 514 Figure 3. Time course of the COVID-19 pandemic up to 21 July 2020 for the nine regions in
18 Afghanistan, for all age groups.
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

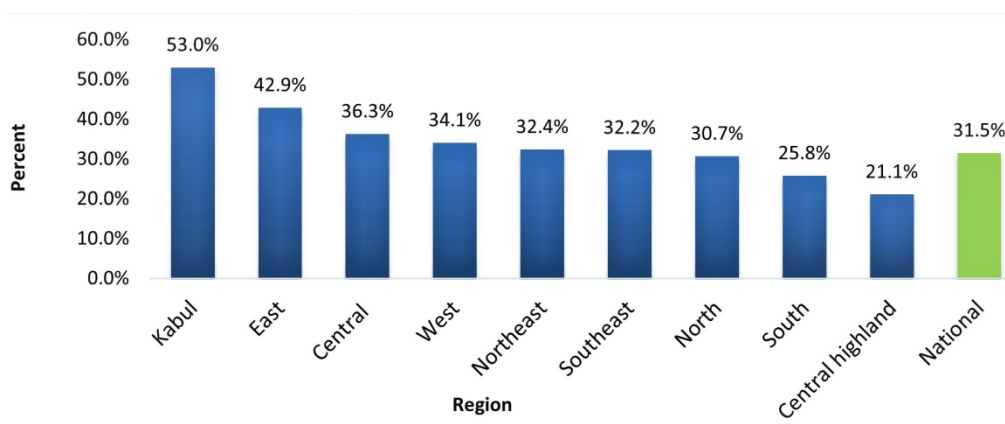


Figure 1. Seroprevalence of SARS-CoV-2 antibodies (including all positive results: IgG-positive, IgM-positive or both) among all age groups by region in Afghanistan.

149x63mm (330 x 330 DPI)

BMJ Open: first published as 10.1136/bmjopen-2021-060739 on 27 July 2022. Downloaded from <http://bmjopen.bmj.com/> on April 22, 2024 by guest. Protected by copyright.

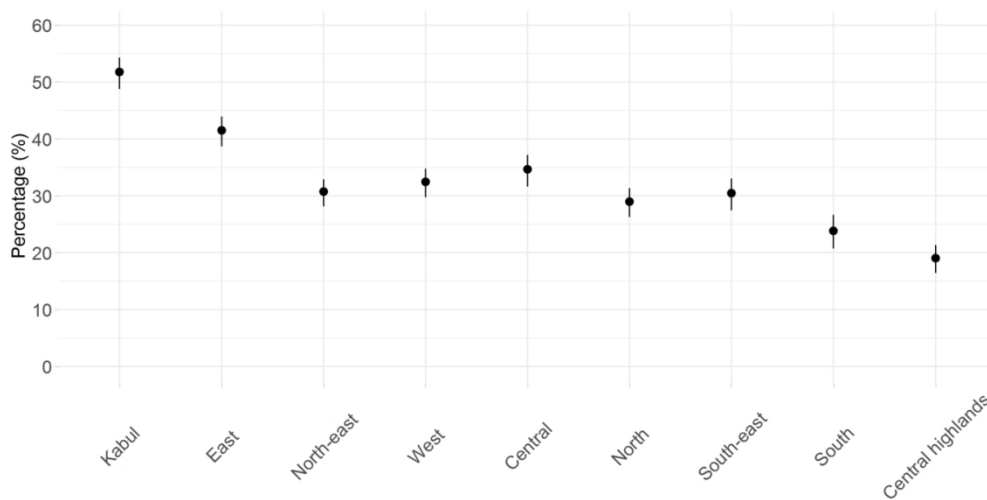
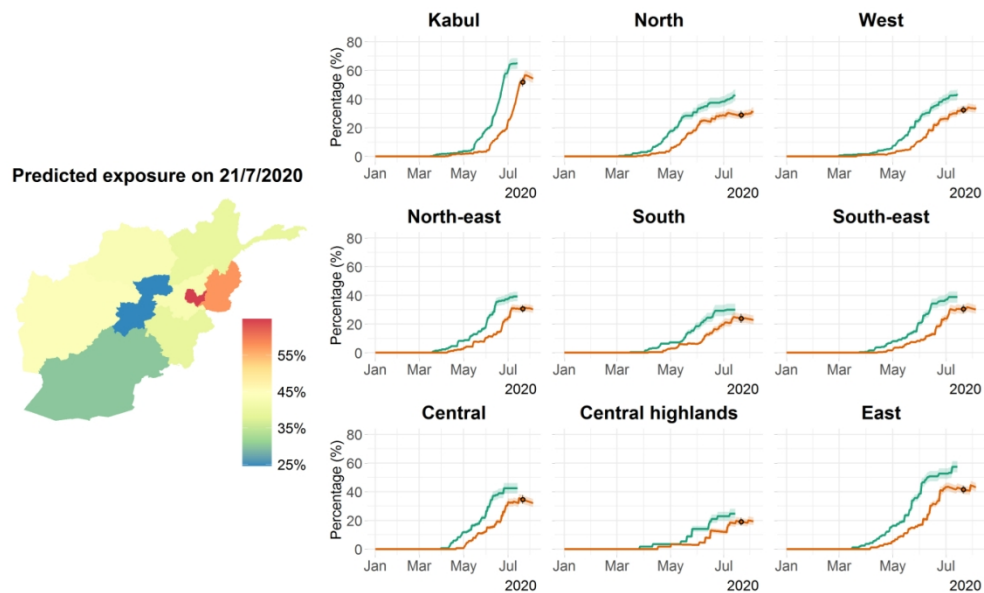


Figure 2. Adjusted seroprevalence by region by the sensitivity and specificity of the serology test for IgG-positive and/or IgM-positive.

158x82mm (220 x 220 DPI)



159x93mm (220 x 220 DPI)

26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Supplementary material

Supplementary appendix 1 – Methods

Data sources

Regional daily deaths

The documented daily mortality data associated with COVID-19, which might be subject to underreporting, for each of the nine Afghanistan regions (Kabul, East, West, North, South, North-east, South-east, Central, and Central highlands) from January 1 to August 4, 2020, were extracted from the Afghanistan Ministry of Health DHIS2 database.

Regional serology data

The proportion of individuals with current or past COVID-19 infection in each region were obtained from the seroepidemiological study data (table 3, main text). The serology survey provided a result for both IgM and IgG antibodies for each participant, using the COVID-19 IgG/IgM Rapid Test Cassette.¹ The dynamics of IgM and IgG antibodies within an infected individual are complicated.² Here, we take the simplified view that an individual who is either IgG positive and/or IgM positive has been exposed to COVID-19 (either past or current infection). Therefore, in the following modelling, the sensitivity and specificity provided by the manufacturer of the imperfect serology test for IgG+ and/or IgM+ was employed.

Adjustment of seroprevalence

We first used a simple Bernoulli model to estimate the regional seroprevalence, after adjusting the proportion of individuals in each region with current or past COVID-19 infection (table 3, main text) according to the sensitivity and specificity of the serology test.¹ (The term ‘seroprevalence’ below denotes the serology positive ratio already adjusted by the test used.) The Bayesian framework was as follows:

$$x_i(t_0) \sim \text{Beta}(1,1)$$

$$w_{ij} \sim \text{bernoulli} \left(k_{se} \times x_i(t_0) + (1 - k_{sp}) \times (1 - x_i(t_0)) \right)$$

$$w_{ij} = \begin{cases} 0, \text{IgG} + \text{ or } \text{IgM} + \\ 1, \text{IgG} - \text{ and } \text{IgM} - \end{cases}, j = \{1, \dots, N_i\} \quad (1)$$

where in the first equation given above we have specified a uniform prior for $x_i(t_0)$, which is the proportion of the population in region i that is serology positive, either for IgM or IgG, on July 21, 2020; w_{ij} is the serology survey result for the j -th participant in the serology study from region i ; N_i is the total number of participants in the serology survey for region i (table S1); and k_{se} (k_{sp}) is the

33 median of the serology test cassette sensitivity (specificity) reported by the manufacturer.¹ The
 34 posterior for seroprevalence on the date the serology survey was conducted, $t = t_0$, was estimated
 35 using a Markov chain Monte Carlo (MCMC) implemented in Rstan³ and denoted as $\tilde{x}_i(t_0)$. The code
 36 associated with this work is publicly available, see the Data availability statement of the paper.

38 Mechanistic model

39 We revised the mathematical model⁴ to account for the underreporting of mortality in the
 40 Afghanistan setting according to the varying serology status of the population, $X_i(t)$, of each
 41 regional population (for each region $i = 1, \dots, 9$, corresponding to Kabul, East, West, North, South,
 42 North-east, South-east, Central, and Central highlands, respectively). The population that has
 43 positive serology status increased with exposure of the population to COVID-19 and decreased due
 44 to the waning of antibodies.

45 Given that the constant age-averaged infection fatality rate by region is β_i , the documented
 46 mortality over time by region is $m_i(t)$, and the reporting rate of mortality associated with COVID-19
 47 by region is q_i , which is assumed to be constant over time, then, at each time step, of the $\frac{1}{q_i\beta_i}m_i(t)$
 48 individuals who were exposed, $\frac{1}{q_i}m_i(t)$ die and the remaining number of individuals, $\frac{1-\beta_i}{q_i\beta_i}m_i(t)$,
 49 seroconvert from negative to positive. Then, assuming that positive individuals convert to negative
 50 at a rate of α , the equation for the rate of change of the number of seropositive individuals is given
 51 by:

$$39 \quad \frac{dX_i(t)}{dt} = \frac{(1-\beta_i)}{q_i\beta_i}m_i(t) - \alpha X_i(t) \quad (2)$$

53 Solving Equation (2), subject to the initial condition $X_i(t = 0) = 0$ where $t = 0$ is time since January
 54 1, 2020, gives:

$$46 \quad X_i(t) = \frac{(1-\beta_i)e^{-\alpha t}}{q_i\beta_i} \int_0^t e^{\alpha r} m_i(r) dr \quad (3)$$

56 Discretising Equation **Error! Reference source not found.** with daily intervals ($\Delta r = 1$) gives:

$$52 \quad X_i(t) = \frac{(1-\beta_i)e^{-\alpha t}}{q_i\beta_i} \sum_{r=0}^t e^{\alpha r} m_i(r) \quad (4)$$

58 Then, the proportion of the population that is serology positive over time, $x_i(t)$, is

$$57 \quad x_i(t) = \frac{X_i(t)}{P_i - \frac{\sum_{r=0}^t m_i(r)}{q_i}} \quad (5)$$

60 Where P_i is the reported population in region i before the COVID-19 outbreak, and the total
61 proportion of the population that has been exposed over time, $\varepsilon(t - \delta_\epsilon)$, is

$$62 \quad \varepsilon_i(t - \delta_\epsilon) = \frac{\frac{1-\beta_i}{q_i\beta_i} \sum_{r=0}^t m_i(r)}{P_i - \frac{\sum_{r=0}^t m_i(r)}{q_i}} \quad (6)$$

63 where δ_ϵ is the time lag between exposure and seroconversion and is fixed at 21 days.⁴

64 Exposure inference

65 We use the posterior samples of seroprevalence at t_0 , $\tilde{x}_i(t_0)$, from the MCMC and combine it with
66 Equations (4) and (5) to calculate the posterior samples of reporting rate for mortality, \tilde{q}_i :

$$67 \quad \tilde{q}_i = \frac{(1-\beta_i)e^{-\alpha t_0} \sum_{t=0}^{t_0} e^{\alpha t} m_i(t)}{x_i(t_0)\beta_i P_i} + \frac{\sum_{t=0}^{t_0} m_i(t)}{P_i} \quad (7)$$

68 Compared with the total population in Afghanistan prior to 2020 (approximately 38 million people),
69 the cumulative mortality associated with COVID-19 by the date of serology survey, $\sum_{t=0}^{t_0} m(t)$, is
70 small. Therefore, it is reasonable to neglect it from Equation (7), which then gives:

$$71 \quad \tilde{q}_i \approx \frac{(1-\beta_i)e^{-\alpha t_0} \sum_{t=0}^{t_0} e^{\alpha t} m_i(t)}{x_i(t_0)\beta_i P_i} \quad (8)$$

72 Combining Equations (4), (5) and (8) we can obtain samples of seroprevalence over time, $\tilde{x}_i(t)$:

$$73 \quad \tilde{x}_i(t) \approx \frac{x_i(t_0) \sum_{r=0}^t e^{\alpha r} m_i(r)}{e^{\alpha(t-t_0)} \sum_{t=0}^{t_0} e^{\alpha t} m_i(t)} \quad (9)$$

74 From Equations (6) and (8) we can obtain samples of the total proportion of the population that has
75 been exposed over time, $\tilde{\varepsilon}_i(t - \delta_\epsilon)$:

$$76 \quad \tilde{\varepsilon}_i(t - \delta_\epsilon) \approx x_i(t_0) e^{\alpha t_0} \frac{\sum_{r=0}^t m_i(r)}{\sum_{t=0}^{t_0} e^{\alpha t} m_i(t)} \quad (10)$$

77 Note that the seroprevalence (9) and exposure (10) over time are not dependent on β . We use the
78 median estimation of α from the constant infection fatality ratio (IFR) model from Chen et al⁴ as an
79 input to Equations (9) and (10).

80

81

82 References

- 83 1. Healgen Scientific LLC. COVID-19 IgG/IgM Rapid Test Cassette (Whole Blood/Serum/Plasma)
84 Instruction for Use 2020 [Available from: <https://www.fda.gov/media/138438/download>].
- 85 2. Seow J, Graham C, Merrick B, et al. Longitudinal observation and decline of neutralizing antibody
86 responses in the three months following SARS-CoV-2 infection in humans. *Nature*
87 *Microbiology* 2020;5(12):1598-607. doi: 10.1038/s41564-020-00813-8
- 88 3. Stan Development Team. RStan: the R interface to Stan; R package version 2.21.2. 2020 [Available
89 from: <http://mc-stan.org/>].
- 90 4. Chen S, Flegg JA, White LJ, et al. Levels of SARS-CoV-2 population exposure are considerably
91 higher than suggested by seroprevalence surveys. *medRxiv* 2021:2021.01.08.21249432. doi:
92 10.1101/2021.01.08.21249432

93

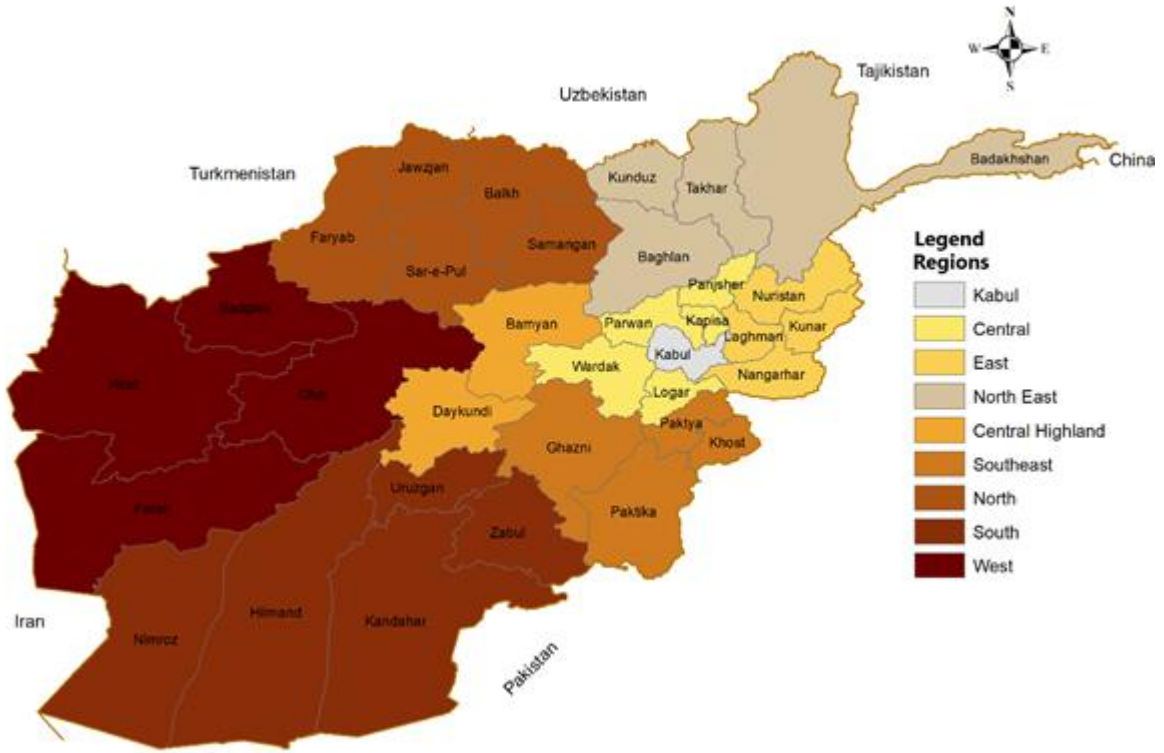
94 **Supplementary Table S1. Sample size for the regional serology survey**

Region	Sample size
Kabul	1104
Central	1056
Central highlands	902
East	1233
North	1071
South	738
North-east	1265
South-east	969
West	1176

95
96 **Supplementary Table S2**

Respondents' characteristics	Number	Percentage
Total respondents	9514	100%
Sex		
Male	5128	53.9%
Female	4386	46.1%
Age		
5–17 years	4346	45.7%
18 years or more	5168	54.3%
Geographical area		
Urban	2574	27%
Rural	6940	73%
Region		
Kabul	1104	11.6%
Central	1056	11.1%
Central highlands	902	9.4%
East	1233	13.0%
North	1071	11.2%
South	738	7.8%
North-east	1265	13.3%
South-east	969	10.2%
West	1176	12.4%

97
98



99

100

101 **Supplementary Figure S1.** A map showing the nine regions in Afghanistan where the study was
102 conducted (the eight regions of Afghanistan plus Kabul province).

102

STROBE Statement—Checklist of items that should be included in reports of *cross-sectional studies*

	Item No	Recommendation	Page No
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	2
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3-4
Objectives	3	State specific objectives, including any prespecified hypotheses	4
Methods			
Study design	4	Present key elements of study design early in the paper	5-8
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5-6
Participants	6	(a) Give the eligibility criteria, and the sources and methods of selection of participants	6
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6-7
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6
Bias	9	Describe any efforts to address potential sources of bias	8
Study size	10	Explain how the study size was arrived at	6
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	6-7
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	7-8
		(b) Describe any methods used to examine subgroups and interactions	
		(c) Explain how missing data were addressed	7
		(d) If applicable, describe analytical methods taking account of sampling strategy	7
		(e) Describe any sensitivity analyses	8
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	8
		(b) Give reasons for non-participation at each stage	8
		(c) Consider use of a flow diagram	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	8-9
		(b) Indicate number of participants with missing data for each variable of interest	8
Outcome data	15*	Report numbers of outcome events or summary measures	9
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	9-13

		(b) Report category boundaries when continuous variables were categorized	
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	13
Discussion			
Key results	18	Summarise key results with reference to study objectives	13-17
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	17
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	17
Generalisability	21	Discuss the generalisability (external validity) of the study results	17
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	18

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at www.strobe-statement.org.