The Use of Primary Care Big Data in Understanding the Pharmacoepidemiology of COVID-19: A Consensus Statement From the COVID-19 Primary Care Database Consortium

Hajira Dambha-Miller, MRCGP, PhD
Simon J. Griffin, DM
Duncan Young, PhD
Peter Watkinson, PhD
Pui San Tan, PhD
Ashley K. Clift, MBBS
Rupert A. Payne, MRCGP
Carol Coupland, PhD
Jemma C. Hopewell, PhD
Jonathan Mant
Richard M. Martin, MRCGP, PhD
Julia Hippisley-Cox, MRCP, MD

ABSTRACT

The use of big data containing millions of primary care medical records provides an opportunity for rapid research to help inform patient care and policy decisions during the first and subsequent waves of the coronavirus disease 2019 (COVID-19) pandemic. Routinely collected primary care data have previously been used for national pandemic surveillance, quantifying associations between exposures and outcomes, identifying high risk populations, and examining the effects of interventions at scale, but there is no consensus on how to effectively conduct or report these data for COVID-19 research. A COVID-19 primary care database consortium was established in April 2020 and its researchers have ongoing COVID-19 projects in overlapping data sets with over 40 million primary care records in the United Kingdom that are variously linked to public health, secondary care, and vital status records. This consensus agreement is aimed at facilitating transparency and rigor in methodological approaches, and consistency in defining and reporting cases, exposures, confounders, stratification variables, and outcomes in relation to the pharmacoepidemiology of COVID-19. This will facilitate comparison, validation, and meta-analyses of research during and after the pandemic.

Ann Fam Med 2021;19:135-140. https://doi.org/10.1370/afm.2658.

MORE ONLINE
www.annfammed.org

INTRODUCTION

Primary care big data refers to routinely collected anonymized general practitioner (GP) electronic health records that form large and complex longitudinal databases, often with hundreds of variables at an individual level. These can often be linked to secondary care records, registries (eg, cancer), or to the UK's Office for National Statistics which records births and deaths.[1-3] A coronavirus disease 2019 (COVID-19) primary care database consortium was established in April 2020 and its researchers have ongoing COVID-19 projects in overlapping data sets with over 40 million UK primary care records that are variously linked to public health, secondary care, and vital status records. We summarize the UK databases being utilized by the COVID-19 consortium in Table 1. It is likely that additional data sources will be forthcoming.

The potential of primary care and linked data for understanding COVID-19 is vast, and includes descriptive epidemiology; testing associations with prescribed drugs, including drugs that influence risk; clinical prediction tools for COVID-19 risk and outcome; the impact of and effects of health inequalities; or examining indirect immediate and long-term effects of the infection, such as delayed clinical diagnoses, domestic abuse, or mental health sequelae.

The focus of this consensus statement is on the pharmacoepidemiology of COVID-19, ie, the potential influence of old and new drug therapies on COVID-19 outcomes. The vast majority of drugs for common conditions

Conflicts of interests: P.T. has consulted for AstraZeneca and Duke-NUS. J.H. reports the Clinical Trial Service Unit receives research grants from the pharmaceutical industry. In addition to University of Oxford affiliation, J.H-C. is founder and director of QResearch database, co-owner of ClinRisk Ltd, and was a paid director there until June 2019. No other authors have any competing interests to declare. The authors declare that no support from any organization and no financial relationships have influenced the submitted work.

CORRESPONDING AUTHOR

Hajira Dambha-Miller
Primary Care and Population Health
University of Southampton
Southampton, SO16 5ST United Kingdom
H.Dambha-Miller@soton.ac.uk

Table 1. Summary of Database Characteristics

Characteristics	QResearch	RCGP Research & Surveillance Network Center	Clinical Practice Research Datalink	UK Biobank
Established, y	2003	1957	1989	2006
GP practices, No.	1,500 (increasing to 2,519 from Sept 2020)	700	1,841	Partial cohort coverage
Current patient records as of Jan 1, 2020, No.	10.6 million (21 million from Sept 2020)	5 million	14 million	0.5 million
Coverage, countries	England, Scotland	England	All of UK	England, Scotland, Wales
Age groups	All	All	All	40-69 years at recruitment
Clinical system	EMIS Web	EMIS Web, INPS Vision, TPP System One	EMIS Web, INPS Vision	Bespoke system
Birth registration	Yes	Yes	Yes	No
Death registration	Yes	Yes	Yes	Yes
Sociodemographic data	Yes	Yes	Yes	Yes
Ethnicity	Yes	Yes	Yes	Yes
Genome-wide genotyping data	No	No	No	Yes
Geographical location	Yes	Yes	Yes	Yes
Lab tests incl COVID-19 results	Yes	Yes	Yes	Yes
Anthropometric data	Yes	Yes	Yes	Yes
Clinical signs and symptoms	Yes	Yes	Yes	Yes
Drugs prescribed	Yes	Yes	Yes	Yes
Radiology reports	Yes	Yes	Yes	No
Hospital referral	Yes	Yes	Yes	Yes
Hospital diagnosis	Yes	Yes	Yes	Yes
GP attendances	Yes	Yes	Yes	Partial
Hospital attendance	Yes	Yes	Yes	Yes
Additional key linkages to other data sets[a]				
Hospital episode statistics	Yes	Yes	Yes	No
HES outpatient data	Yes	Yes	Yes	No
HES accident and emergency data	Yes	Yes	Yes	No
HES diagnostic imaging data set	Yes	Yes	Yes	No
Death registration data from the Office for National Statistics	Yes	Yes	Yes	Yes
Intensive care data set: ICNARC Case Mix Program	Yes	No	Pending	No
URL for data access	https://www.qresearch.org	https://www.qresearch.org	https://www.cprd.com	https://www.ukbiobank.ac.uk/aboutbiobank-uk

COVID-19 = coronavirus disease 2019; GP = general practitioner; HES = hospital episode statistics; ICNARC = Intensive Care National Audit and Research Centre; INPS = In Practice Systems Limited; RCGP = Royal College of General Practitioners; TPP = The Phoenix Partnership.

Note: For more information see relevant websites. The data in these databases are likely to overlap (about 20% of patients will fall into at least 2 of the data sets). Additional governance and approvals will be needed to remove duplicate entries so that a single patient record and characteristics are included.

[a] Full lists available from each database on request.

such as hypertension, diabetes, or heart failure are prescribed in primary care; eg, 15 million prescriptions of angiotensin-converting enzyme inhibitors were prescribed by UK primary care practitioners last year alone—this drug is now hypothesized to be significant to COVID-19 outcomes.[1]

With an increasing number of studies using primary care big data to examine the influence of drugs on COVID-19 outcomes, it is timely to consider how to best conduct studies. This will facilitate study consistency and rigor, improve transparency, and reduce ambiguity in both methods and reporting. As the pandemic progresses, with the urgency to find solutions, rapid and rigorous research must be conducted with emergent findings externally validated. Consensus on definitions of COVID-19, exposures, outcomes, and consistency in considering potential confounders and stratification variables, will enable meaningful comparisons between findings and facilitate the potential for pooling results in meta-analyses.

Our collective efforts and agreement on transparency in reporting and methodological approaches may contribute toward improved clinical decisions and in turn, improved population health.

METHODS: CONSENSUS DEVELOPMENT

This statement was developed by our primary care database consortium: UK big data experts, epidemiologists, intensive care and primary care researchers, statisticians, patient and public representatives, journal editors, and frontline clinical staff. Universities represented are Bristol, Cambridge, Manchester, Nottingham, Oxford, and Southampton. After initial discussions on the need for such a statement in our projects, we met weekly to refine ideas and reach agreement on item inclusion. The recommendations are entirely those of the consortium with sponsors and funders having no role in their development or reporting.

RESULTS: ITEMS FOR INCLUSION

Protocols

Wherever possible, protocols and analysis plans for COVID-19 research using primary care big data should be made widely available online before commencing the data analysis. Protocols may be published through open-access journals or publicly available institutional repositories. When a study has to adapt or respond to changes, for example, in data availability, published protocols should be updated accordingly. This will facilitate transparency and scrutiny of findings, and a priori analysis plans should reduce false-positive findings. Given the urgency of the research during the current pandemic, publicly available protocols will encourage greater efficiency and less duplication of efforts, except when duplication is conducted for validation purposes. Further, when the same data releases are being reused to answer additional or new research questions, this should be set out in the protocol with adequate consideration of the Findable, Accessible, Interoperable, and Resuable (FAIR) Guiding Principles for scientific data management and stewardship.[2] These principals, however, are generic. In most cases individual patient level data are unlikely to be published in repositories; reuse of the data would thus require applying to the data owner and working with them to ensure an identical data cut is obtained.

COVID-19 Infection Definition

We considered the World Health Organization (WHO) and Public Health England definitions of COVID-19 infection as of March 13, 2020.[4,5] These definitions have several limitations in terms of operationalizing them within databases of electronic medical records. For example, clinical symptoms such as nasal discharge or sneezing may not be coded within GP records. Virology and serology tests that are being used to define cases are still being validated with varying reports on sensitivity and specificity.[6] UK national testing has been sparse in the initial phases of the pandemic, and has varied between hospital and community settings at different stages of the pandemic. Definitions are evolving each day and across different regions. They are also likely to evolve over time as more standardization of GP coding is introduced, testing rates are widened from largely secondary to community settings, and more accurate tests of infection and immunity are developed. Acknowledging these limitations, at this time, we have agreed on the definitions which follow.

Confirmed COVID-19 Case

Positive result from real-time reverse transcription polymerase chain reaction (RT-PCR) assay of nasal or pharyngeal swab specimen ever *OR* positive serology ever is considered a confirmed case.

Suspected COVID-19 Case

Cases may be only be classified as suspected if they occurred after January 20, 2020, the date first confirmed case outside mainland China was reported to the World Health Organization, *AND* patients meet the following additional criteria:

- Required admission to hospital *AND* 1 or more of the following:
  - Clinical evidence of pneumonia *OR*
  - Radiological evidence of pneumonia *OR*
  - Acute respiratory distress syndrome *OR*
  - Had an influenza-like illness (fever ≥37.8°C) and at least 1 of the following acute onset respiratory symptoms: persistent cough (with or without sputum), hoarseness, nasal discharge or congestion, shortness of breath, sore throat, wheezing, sneezing

  OR
- Did not require admission to hospital *AND* had 1 or both of the following:
  - New continuous cough *AND/OR* high temperature

Suspected COVID-19 With a Negative Test

Negative result (never positive) on RT-PCR assay of nasal or pharyngeal swab specimen *OR* negative (never positive) serology.

Drug Exposures

Where prescribed drugs are being examined in COVID-19 infection, the following reporting principles should be applied:

- Report generic drug name
- List distinct classes of drugs (eg, angiotensin receptor blockers and angiotensin-converting enzyme-inhibitors are listed as separate drug classes)

AND

- Provide individual drug chemical names
- When combination preparations have been prescribed, consider the component ingredients as separate for the purposes of the analysis
- Provide clear definitions of drug exposure including:
  - Exposure time—describe relevant dates of prescription for drug being investigated in relation to COVID-19 case-definition date; ie, time duration before/during/after infection. For research questions with specific aim of altering outcomes, drug exposure during/after infection will be most informative
  - Repeat prescriptions for long-term medications—list number of prescriptions within a defined time period
  - Dosage—describe how different drug dosage regimens are being treated in analysis

A list of UK-prescribed drugs that require urgent characterization for their potential in treating or altering outcomes in COVID-19 is given in Supplemental Appendix 1., available at https://www.AnnFamMed.org/content/19/2/135/suppl/DC1/. This list was compiled from the limited existing literature on the subject, our ongoing systematic review, and anecdotal evidence from front line clinical staff treating patients with COVID-19 infections as of April 18, 2020.[7-11] The list is not exhaustive and will be updated on the QResearch website (www.qresearch.org/) as more data become available. Drug names within classes have been extracted from the British National Formulary.

Confounding Variables

A list of variables that we recommend reporting and considering for inclusion as confounders within statistical models can be found in the Supplemental Appendix 2, available at https://www.AnnFamMed.org/content/19/2/135/suppl/DC1/. We have not provided restrictive recommendations on how these variables should be categorized as this is dependent on the data available and researcher discretion. We suggest considering these variables when determining confounders with a clear description of how categorization of the variable was determined and rationale for its inclusion.

Stratification Variables

In response to emergent findings on particular subgroups within the populations being disproportionately affected by COVID-19 infection,[12] we propose examining outcomes stratified by age, sex, ethnic group, and domicile (own home vs care/nursing home). As more data are published, further stratification variables should be considered.

Outcome Reporting

We endorse the use of appropriate reporting guidelines such as the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE), REporting of studies Conducted using Observational Routinely-collected Data (RECORD), or Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) checklist.[13-15] We have considered outcomes from both intensive care unit (ICU) and primary care data as some linked data sets are used within our consortium. Wherever possible, there are minimal data (Table 2) to be reported. We acknowledge, as the short and long-term sequela of COVID-19 are better understood, additional outcomes in primary care will need to be added to this list.

Analytical Methods

Individual study analyses are likely to vary from project to project. We agree on ensuring analytical

Table 2. Recommended Outcome Reporting for COVID-19 Studies

Treatment Setting	Time Period	Type of Outcomes to Report
Primary care outcomes	Short-term	Lower respiratory tract infection (pneumonia)
		Emergency admission
		ICU admission
	Long-term	All-cause mortality and cause-specific mortality including COVID-19 specific mortality
ICU outcomes	Short-term	Vital status at ICU discharge (alive/dead)
		Vital status at acute hospital discharge (alive/dead)
		Days of advanced respiratory support (artificial ventilation)
		Days of advanced cardiovascular support (inotropes, pressors, or mechanical cardiovascular support)
		Days of renal support (use of renal replacement therapy)
		Days of ICU care (reported from ICU admission to discharge)
		Days of acute hospital care after ICU discharge (for repeat ICU admissions in the same acute hospital admission the total days not on ICU should be used)
	Long-term	Vital status (alive/dead) at 30 and 90 days after ICU admission
		All-cause and COVID-19 specific mortality at 6 and 12 months after ICU admission

COVID-19 = coronavirus disease 2019; ICU = intensive care unit.

methods are transparent and reported in full, with the following guiding principles:

- State an a priori hypothesis wherever possible
- Report descriptive characteristics including age, sex, ethnic group, measures of deprivation (such as the Index of Multiple Deprivation or Townsend deprivation score), comorbidities, and medication use
- State sample size considerations, power calculations, and multiple testing considerations
- Consider clustering by ICU or general practices or physician and employ appropriate methods (eg, robust standard errors)
- Check assumptions for any models (eg, proportional hazards assumption)
- Report how missing data were managed (eg, multiple imputation method to replace missing data)
- Report both unadjusted and adjusted models
- Report methods used to examine subgroups and interactions
- Report causal analysis methods (eg, instrumental variables analysis[16])
- Report sensitivity analyses
- Report steps taken to mitigate time-window bias (in case-control study designs) or immortal time bias (in cohort study designs)[17]
- Consider propensity score weighting methods to account for multiple differences between groups

DISCUSSION: KEY CHALLENGES AND SHORTCOMINGS

Across the primary care data sets, we acknowledge the potential limitations of using big data for COVID-19 research. All data are collected from routine clinical care records. They are dependent on accurate coding by individual clinicians which does not guarantee consistency or accuracy of codes. The precision and quality of each variable may be different and it is necessary to ensure adequate preliminary work be done on each variable's quality, completeness, and accuracy.[3,4] Further, some data on exposures and confounders will have been entered before the pandemic, and there might be a delay in outcome data reaching GP records. Uptake of newly introduced clinical codes that are specific to COVID-19 may not be universal. Historically, however, UK primary care records have been of high quality in terms of accuracy, completeness of clinical diagnosis, and medication prescribing.[18,19] The use of non-randomized observational data to make causal inferences still requires careful interpretation and appropriate analyses.[20] Other considerations relate to the case definition of COVID-19.

Our definitions have been informed by those proposed by Public Health England and the WHO. We will use positive RT-PCR or serology results as a definitive for confirmed cases. UK testing for COVID-19 has been limited and to date, we are not aware of any established serology or virology test with high sensitivity or specificity. Moreover, recent modeling suggests that there might be a substantial proportion of asymptomatic COVID-19 cases.[21] These individuals will not have presented to the health services and or identified within our data sets. It is possible that patients who are in trials on experimental treatments may be included in the large database analyses. Although they may be few and may have a flag in their record to identify their inclusion within a trial, this may not always be the case.

It is also plausible that big data will over-represent disease severity and the contributing factors because the less severe and asymptomatic cases are not recorded. This issue will be less relevant in subgroup designs or analyses assessing the risk of adverse outcomes in those presenting to hospital or ICU. All observational studies nested within these databases will be subject to the usual risk of statistical error (type 1 or 2), bias, and confounding. These must be considered in terms of magnitude and direction. It is likely that many of the biases will be non-differential and minimized to some extent by the large sample sizes afforded by the data. Moreover, these primary care data lack selection and recall biases and often include multiple linkages to enable best-attainable ascertainment of outcome and exposure data. Large sample sizes will increase precision but could also lead to false positives. In the early stages of the pandemic, the number of people with outcomes recorded in GP records will be small but this is rising rapidly and the timing of analyses will, therefore, be important. If conducted too early the sample size will be inadequate but if too late, opportunities for findings to influence policy will be missed.

CONCLUSIONS

Our consensus statement focused on inferential analytical methods with a recommendation for a priori hypothesis and sample size calculations. However, big data should not ignore the increasing value of exploratory or data mining analyses. These methods have raised concerns around reporting and data interpretation but will inevitably become more widely used especially as new diseases such as COVID-19 emerge. Indeed, many of the principles that we have set out here will apply to the pharmacoepidemiology research questions of the future. Finally, our consortium has thus far included researchers and databases across the UK but our work could inform similar approaches worldwide. Our databases are available to

those outside the UK. Establishing linked primary care records within individual countries has immense potential to answer pressing national research questions. Further, this could subsequently allow between-country comparisons and external validation of findings.

To read or post commentaries in response to this article, go to https://www.AnnFamMed.org/content/19/2/135/tab-e-letters.

Key words: big data; coronavirus; epidemiology; primary health care

Submitted May 1, 2020; submitted, revised, August 7, 2020; accepted August 31, 2020.

Author affiliations: Division of Primary Care and Population Health, University of Southampton, England (H.D-M.); Department of Public Health and Primary Care, University of Cambridge, England (S.J.G.; J.M.); Nuffield Department of Clinical Neurosciences, University of Oxford, England (D.Y., P.W.); Nuffield Department of Primary Care Health Sciences, University of Oxford, England (A.K.C., P.S.T., J.H-C.); Population Health Sciences, University of Bristol, England (R.A.P., R.M.M.); Division of Primary Care, School of Medicine, University of Nottingham, England (C.C.); Clinical Trial Service Unit, Nuffield Department of Population Health, University of Oxford, England (J.C.H.); National Institute of Health Research Bristol Biomedical Research Centre, University of Bristol, England (R.M.M.).

Author contributions: H.D-M. drafted the manuscript and all authors contributed toward revising it.

Funding support: S.J.G. is supported by an MRC Epidemiology Unit program: MC_UU_12015/4. The University of Cambridge has received salary support for S.J.G. from the NHS in the East of England through the Clinical Academic Reserve. H.D-M. is a NIHR funded Academic Clinical Lecturer. R.M.M. is supported in part by the NIHR Bristol Biomedical Research Centre and by a Cancer Research UK (C18281/A19169) program grant (the Integrative Cancer Epidemiology Programme). P.W. is supported in part by the NIHR Oxford Biomedical Research Centre. J.C.H. acknowledges personal support from the British Heart Foundation (FS/14/55/30806) and Cancer Research UK (C5255/A18085) through the Cancer Research UK Oxford Centre. J.H-C. also receives support from the NHS and NIHR. J.M. is an NIHR Senior Investigator.

Disclaimer: The views and opinions expressed by authors in this article are those of the authors and do not necessarily reflect those of the UK NIHR or the Department of Health and Social Care.

Transparency declaration: This manuscript is an honest, accurate, and transparent account; no important aspects have been omitted.

Supplemental materials: Available at https://www.AnnFamMed.org/content/19/2/135/suppl/DC1/.

References

1. Williams T, Van Staa T, Puri S, *et al.* Recent advances in the utility and use of the General Practice Research Database as an example of a UK Primary Care Data resource. *Thera Adv Drug Saf.* 2012;3(2):89-99. doi:10.1177/2042098611435911
2. Herrett E, Gallagher AM, Bhaskaran K, et al. Data resource profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol.* 2015;44(3):827-836. 10.1093/ije/dyv098
3. de Lusignan S, Correa A, Smith GE, et al. RCGP Research and Surveillance Centre: 50 years' surveillance of influenza, infections, and respiratory conditions. *Br J Gen Pract.* 2017;67(663):440-441.
4. Global Surveillance for human infection with novel coronavirus disease (2019-nCoV). World Health Organization. Published Jan 31, 2020. Accessed Apr 13, 2020. https://apps.who.int/iris/handle/10665/330857
5. COVID-19: investigation and initial clinical management of possible cases. Public Health England. Updated Oct 2, 2020. Accessed Apr 13, 2020. https://www.gov.uk/government/publications/wuhan-novel-coronavirus-initial-investigation-of-possible-cases/investigation-and-initial-clinical-management-of-possible-cases-of-wuhan-novel-coronavirus-wn-cov-infection
6. Hoffman T, Nissen K, Krambrich J, et al. Evaluation of a COVID-19 IgM and IgG rapid test; an efficient tool for assessment of past exposure to SARS-CoV-2. *Infect Ecol Epidemiol.* 2020;10(1):1754538.
7. Patel AB, Verma A. COVID-19 and angiotensin-converting enzyme inhibitors and angiotensin receptor blockers: what is the evidence? *JAMA*. 2020;323(18):1769-1770. 10.1001/jama.2020.4812
8. Gbinigie K, Frie K. Should chloroquine and hydroxychloroquine be used to treat COVID-19? A rapid review. *BJGP Open*. 2020;4(2):bjgpopen20X101069. doi:10.3399/bjgpopen20X101069
9. Little P. Non-steroidal anti-inflammatory drugs and COVID-19. *BMJ*. 2020;368:m1185. 10.1136/bmj.m1185
10. Cortegiani A, Ingoglia G, Ippolito M, Giarratano A, Einav S. A systematic review on the efficacy and safety of chloroquine for the treatment of COVID-19. *J Crit Care.* 2020;57:279-283.
11. Dong L, Hu S, Gao J. Discovering drugs to treat coronavirus disease 2019 (COVID-19). *Drug Discov Ther.* 2020;14(1):58-60.
12. Rimmer A. COVID-19: Disproportionate impact on ethnic minority healthcare workers will be explored by government. *BMJ.* 2020;369:m1562. doi:10.1136/bmj.m1562
13. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med.* 2007;147(8):573-577. 10.7326/0003-4819-147-8-20071016000010
14. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med.* 2015;162(1):55-63. 10.7326/M14-0697
15. Benchimol EI, Smeeth L, Guttmann A, et al; RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med.* 2015;12(10):e1001885. 10.1371/journal.pmed.1001885
16. Rassen JA, Brookhart MA, Glynn RJ, Mittleman MA, Schneeweiss S. Instrumental variables II: instrumental variable application-in 25 variations, the physician prescribing preference generally was strong and reduced covariate imbalance. *J Clin Epidemiol.* 2009;62(12):1233-1241. 10.1016/j.jclinepi.2008.12.006
17. Suissa S. Immortal time bias in observational studies of drug effects. *Pharmacoepidemiol Drug Saf.* 2007;16(3):241–249.
18. Jick H, Jick SS, Derby LE. Validation of information recorded on general practitioner based computerised data resource in the United Kingdom. *BMJ.* 1991;302(6779):766-768.
19. Majeed A. Sources, uses, strengths and limitations of data collected in primary care in England. *Health Stat Q.* 2004;(21):5-14.
20. Hernán M, Robins J. Instruments for causal inference: an epidemiologist's dream? *Epidemiology.* 2006;17(4):360-372.
21. Lourenco J, Paton R, Ghafari M, et al. Fundamental principles of epidemic spread highlight the immediate need for large-scale serological surveys to assess the stage of the SARS-CoV-2 epidemic. *medRxiv.* Preprint published online Mar 26, 2020.