The covidization of medical science

October 2023, EBHC conference, Taormina

John P.A. Ioannidis, MD, DSc
Professor of Medicine, of Epidemiology and Population Health, and (by courtesy) of Biomedical Data Science and of Statistics
Co-Director, Meta-Research Innovation Center at Stanford (METRICS)
Stanford University

Major scientific response
to a major crisis
~2 million scientists
published
~1 million scientific
papers on COVID-19

(Ioannidis J. et al, Royal Society Open Science 2021)

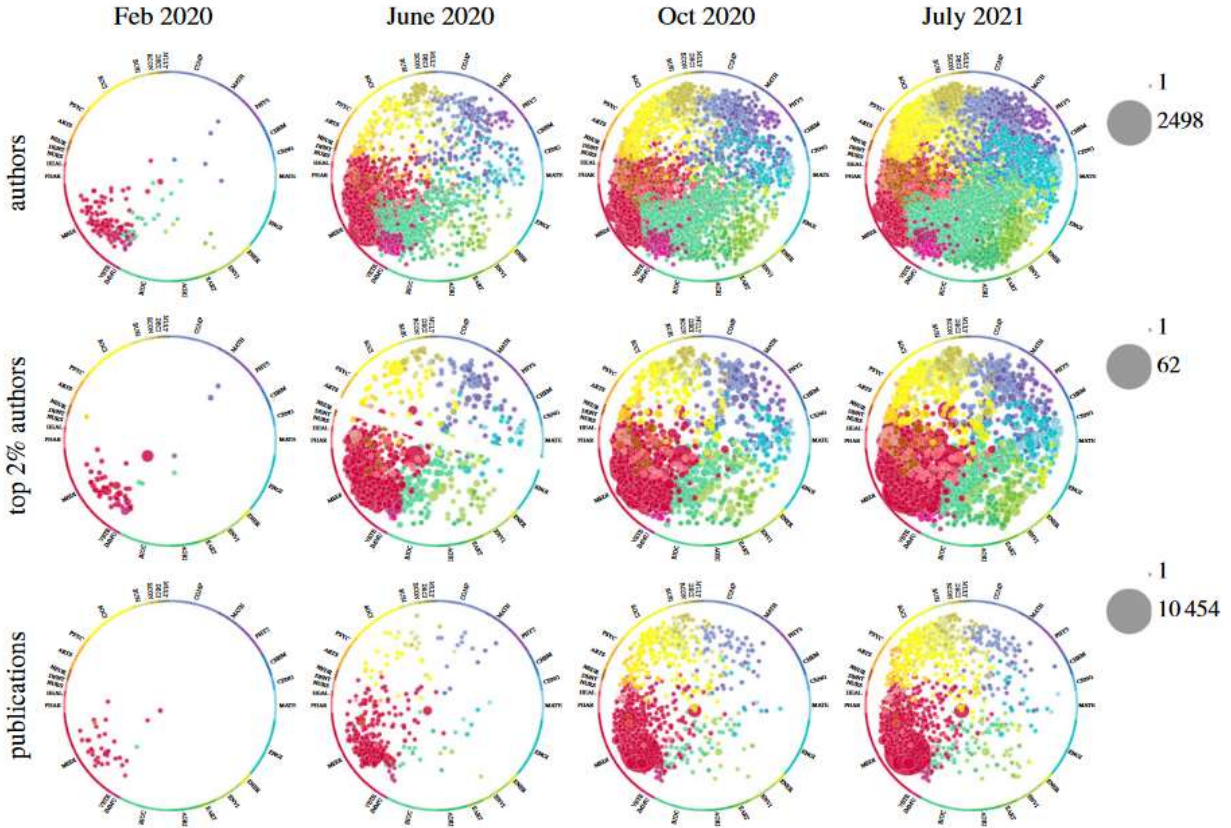


Figure 1. Topics of prominence for COVID-19 authors and publications. The columns represent the progress of the spread at three different measuring points: by end of February 2020, end of June 2020, end of October 2020 and end of July 2021. The first row represents the spread of authors of COVID-19 papers. The authors are assigned to their most dominant topic in their career. The data are filtered to include only topics with greater than or equal to five authors assigned. The second row shows similarly the topics of

98/100 most-cited papers in 2020-2021 across science were on COVID-19

PNAS RESEARCH ARTICLE | MEDICAL SCIENCES OPEN ACCESS

Massive covidization of research citations and the citation elite

John P. A. Ioannidis[a,b,c,d,e,1], Eran Bendavid[a], Maia Salholz-Hillel[f], Kevin W. Boyack[g], and Jeroen Baas[h]

Edited by Kenneth Wachter, University of California, Berkeley, CA; received March 7, 2022; accepted May 31, 2022

Massive scientific productivity accompanied the COVID-19 pandemic. We evaluated the citation impact of COVID-19 publications relative to all scientific work published in 2020 to 2021 and assessed the impact on scientist citation profiles. Using Scopus data until August 1, 2021, COVID-19 items accounted for 4% of papers published, 20% of citations received to papers published in 2020 to 2021, and >30% of citations received in 36 of the 174 disciplines of science (up to 79.3% in general and internal medicine). Across science, 98 of the 100 most-cited papers published in 2020 to 2021 were related to COVID-19; 110 scientists received ≥10,000 citations for COVID-19 work, but none received ≥10,000 citations for non–COVID-19 work published in 2020 to 2021. For many scientists, citations to their COVID-19 work already accounted for more than half of their total career citation count. Overall, these data show a strong covidization of research citations across science, with major impact on shaping the citation elite.

Significance

The COVID-19 pandemic saw a massive mobilization of the scientific workforce. We evaluated the citation impact of COVID-19 publications relative to all scientific work published in 2020 to 2021, finding that 20% of citations received to papers published in 2020 to 2021 were to COVID-19–related papers. Across

Yet, quality of science suffered

Methodological quality of COVID-19 clinical research

Richard G. Jung[1,2,3,13], Pietro Di Santo[1,2,4,5,13], Cole Clifford[6], Graeme Prosperi-Porta[7], Stephanie Skanes[6], Annie Hung[8], Simon Parlow[4], Sarah Visintini[9], F. Daniel Ramirez[1,4,10,11], Trevor Simard[1,2,3,4,12] & Benjamin Hibbert[2,3,4]✉

The COVID-19 pandemic began in early 2020 with major health consequences. While a need to disseminate information to the medical community and general public was paramount, concerns have been raised regarding the scientific rigor in published reports. We performed a systematic review to evaluate the methodological quality of currently available COVID-19 studies compared to historical controls. A total of 9895 titles and abstracts were screened and 686 COVID-19 articles were included in the final analysis. Comparative analysis of COVID-19 to historical articles reveals a shorter time to acceptance (13.0[IQR, 5.0–25.0] days vs. 110.0[IQR, 71.0–156.0] days in COVID-19 and control articles, respectively; $p < 0.0001$). Furthermore, methodological quality scores are lower in COVID-19 articles across all study designs. COVID-19 clinical studies have a shorter time to publication and have lower methodological quality scores than control studies in the same journal. These studies should be revisited with the emergence of stronger evidence.

Scientific quality of COVID-19 and SARS CoV-2 publications in the highest impact medical journals during the early phase of the pandemic: A case control study

Marko Zdravkovic[1,4], Joana Berger-Estilita[2,6], Bogdan Zdravkovic[1], David Berger[3,5]

COVID-19-related medical research: a meta-research and critical appraisal

Marc Raynaud[1†], Huanxi Zhang[2†], Kevin Louis[1†], Valentin Goutaudier[1,3†], Jiali Wang[2], Quentin Dubourg[4], Yongcheng Wei[2], Zeynep Demir[1,3], Charlotte Debiais[1], Olivier Aubert[1], Yassine Bouatou[1], Carmen Lefaucheur[6], Patricia Jabre[7], Longshan Liu[2], Changxi Wang[2], Xavier Jouven[1], Peter Reese[1,8], Jean-Philippe Empana[1] and Alexandre Loupy[1*]

Abstract

Background: Since the start of the COVID-19 outbreak, a large number of COVID-19-related papers have been published. However, concerns about the risk of expedited science have been raised. We aimed at reviewing and categorizing COVID-19-related medical research and to critically appraise peer-reviewed original articles.

Methods: The data sources were Pubmed, Cochrane COVID-19 register study, arXiv, medRxiv and bioRxiv, from 01/11/2019 to 01/05/2020. Peer-reviewed and preprints publications related to COVID-19 were included, written in English or Chinese. No limitations were placed on study design. Reviewers screened and categorized studies according to *i*) publication type, *ii*) country of publication, and *iii*) topics covered. Original articles were critically appraised using validated quality assessment tools.

Results: Among the 11,452 publications identified, 10,516 met the inclusion criteria, among which 7468 (71.0%) were peer-reviewed articles. Among these, 4190 publications (56.1%) did not include any data or analytics (comprising expert opinion pieces). Overall, the most represented topics were infectious disease ($n = 2326$, 22.1%), epidemiology ($n = 1802$, 17.1%), and global health ($n = 1602$, 15.2%). The top five publishing countries were China (25.8%), United States (22.3%), United Kingdom (8.8%), Italy (8.1%) and India (3.4%). The dynamic of publication showed that the exponential growth of COVID-19 peer-reviewed articles was mainly driven by publications without original data (mean 261.5 articles ± 51.1 per week) as compared with original articles (mean of 69.3 ± 22.3 articles per week). Original articles including patient data accounted for 713 (9.5%) of peer-reviewed studies. A total of 576 original articles (80.8%) showed intermediate to high risk of bias. Last, except for simulation studies that mainly used large-scale open data, the median number of patients enrolled was of 102 (IQR = 37–337).

Conclusions: Since the beginning of the COVID-19 pandemic, the majority of research is composed by publications without original data. Peer-reviewed original articles with data showed a high risk of bias and included a limited number of patients. Together, these findings underscore the urgent need to strike a balance between the velocity and quality of research, and to cautiously consider medical information and clinical applicability in a pressing pandemic context.

(Continued on next page)

Overview of large-scale quality assessments

Aim

- To systematically collect and summarize
  - all the meta-epidemiological assessments of COVID-19 literature
  - the large-scale SRs and MAs (>150 articles) with a risk-of-bias appraisal

Methods

- PubMed search
- From January 2020 to August 2022
  - 81 articles on research quality of COVID-19 research
  - 66 large-scale SRMAs (>150 articles) including quality assessment

(Ongoing project with Lazaros Belbasis)

Risk of bias assessment

- Among **606 prediction models**, only 5% were at low risk of bias based on PROBAST.

Wynants L et al. BMJ 2020;369:m1328

- Among 463 **RCTs**, 26% were at low risk of bias based on Cochrane risk of bias tool.

Siemieniuk RAC et al. BMJ 2020;370:m2980

- Among 968 **seroprevalence studies**, only 10% were at low risk of bias based on JBI tool for prevalence studies.

Bobrovitz N et al. PLoS One 2021;16(6):e0252617

- Among 243 **systematic reviews**, none of them was at low risk of bias based on AMSTAR-2 tool.

Li Y et al. J Clin Epidemiol 2021;135:17-28

Comparative assessment of research quality

- Compared to a historical control, COVID-19 research articles have a shorter time of acceptance and lower methodological quality.

Jung RG et al. Nat Commun 2021;12(1):943

- In the highest impact medical journals, COVID-19 research articles are of lower level of evidence and of lower methodological quality compared to nonCOVID-19 articles published in the same time period.

Zdravkovic M et al. PLoS One 2020;15(11):e0241826

- In the highest impact medical journals, COVID-19 research articles present lower adherence to reporting guidelines compared to nonCOVID-19 articles published in the same time period.

Quinn TJ et al. BMC Med 2021;19(1):46

Fabrication: even major studies in major journals
Lancet and its peer-reviewers could not realize a study claimed to have happened in 671 hospitals was entirely fake

Articles

Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis

Mandeep R Mehra, Sapan S Desai, Frank Ruschitzka, Amit N Patel

Summary

Background Hydroxychloroquine or chloroquine, often in combination with a second-generation macrolide, are [illegible] widely used for treatment of COVID-19, despite no conclusive evidence of their benefit. Although generally well [illegible] used for approved indications such as autoimmune disease or malaria, the safety and benefit of these treatment regimens are poorly evaluated in COVID-19.

Methods We did a multinational registry analysis of the use of hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19. The registry comprised data from 671 hospitals in six continents. We included patients hospitalised between Dec 20, 2019, and April 14, 2020, with a positive laboratory test for SARS-CoV-2. Patients who received one of the treatments of interest within 48 h of diagnosis were included in one of four treatment groups (chloroquine alone, chloroquine with a macrolide, hydroxychloroquine alone, or hydroxychloroquine with a macrolide), and patients who received none of these treatments formed the control group. Patients for whom one of the treatments of interest was initiated more than 48 h after diagnosis or who were on mechanical ventilation, as well as patients who received remdesivir, were excluded. The main outcome and [illegible] were in-hospital mortality and the occurrence of de-novo ventricular arrhythmias (sustained or non-sustained ventricular tachycardia or ventricular fibrillation).

Findings 96032 patients (mean age 53·8 years, 46·3% women) with COVID-19 were hospitalised during the study [illegible]

RETRACTED

This online publication has been corrected. The corrected version first appeared at thelancet.com on May 29, 2020

See Online/Comment

Brigham and Women's Hospital Heart and Vascular Center and Harvard Medical School, Boston, MA, USA (Prof M R Mehra MD); Surgisphere Corporation, Chicago, IL, USA (S S Desai MD); University Heart Centre, University Hospital Zurich, Zurich, Switzerland

doi:10.1093/eurheartj/ehab398

Global spotlights

Retractions in medicine: the tip of the iceberg

Ivan Oransky[1], Stephen E. Fremes[2], Paul Kurlansky[3], and Mario Gaudino [4]*

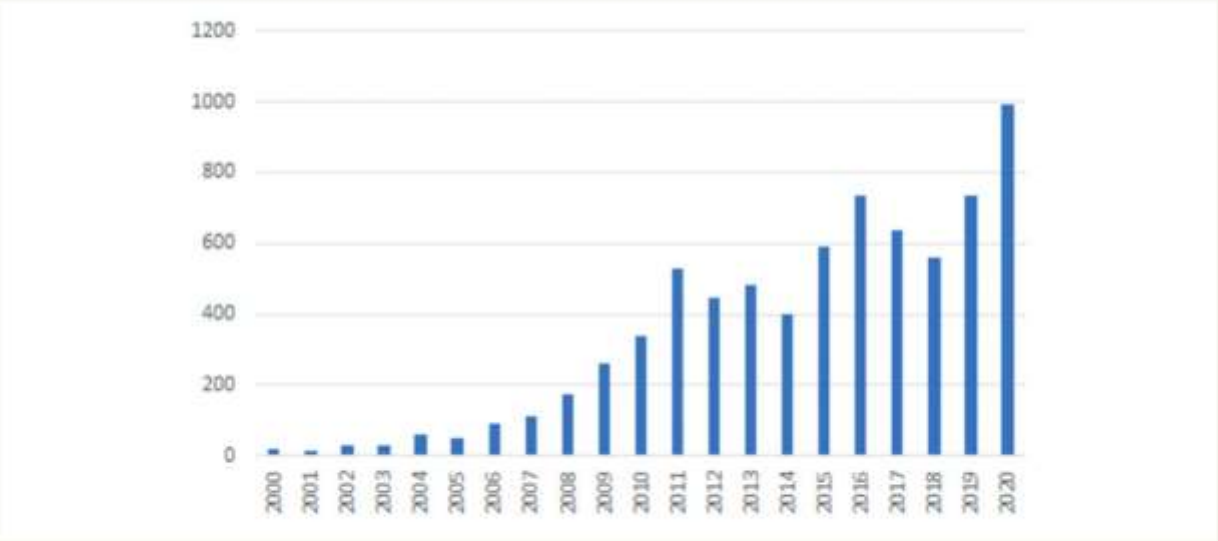


Figure 1 Retractions in the health sciences by year since 2000. Data are from the Retraction Watch Database at retractiondatabase.org.

Retracted papers originating from paper mills: cross sectional study

Cristina Candal-Pedreira,[1,2] Joseph S Ross,[3,4,5] Alberto Ruano-Ravina,[1,2,6] David S Egilman,[7] Esteve Fernández,[8,9] Mónica Pérez-Ríos[1,2,6]

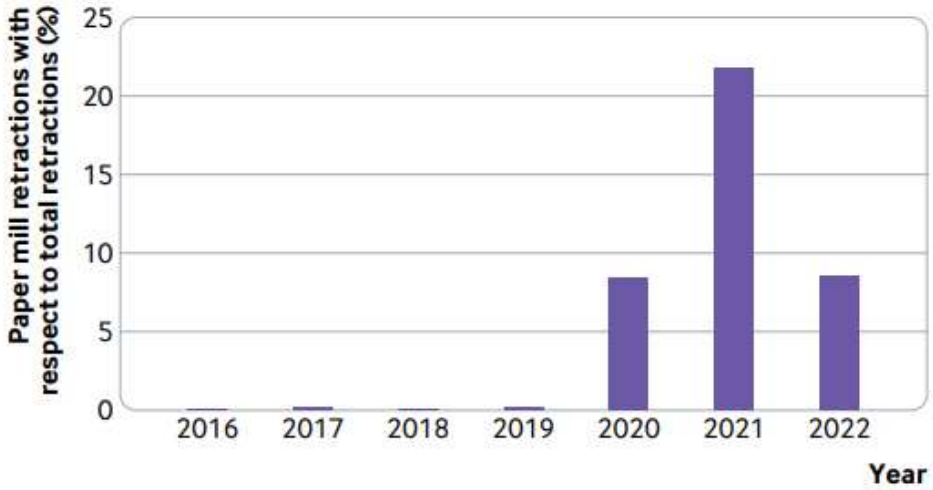


Fig 2 | Percentage of paper mill retractions with respect to total retractions

Science reached unprecedented attention outside of science

TABLE 1 Increase in total published items in the scientific literature and of published items with extreme attention in media and social media (Altmetric scores >4000)[a]

Year	Published items	Items with Altmetric >4000
2022 (first half)	3,099,247	104
2021	6,624,362	327
2020	6,575,801	351
2019	5,829,102	39
2018	5,394,434	39
2017	5,066,175	38
2016	4,617,354	19

[a]Data are derived from the dimensions.ai database with search in July 7, 2022; the year 2022 may be partly incomplete even for the first half due to registration delays.

Dramatis personae

- Journalists
- Social media influencers
- Science journalists
- Politicians
- Big tech stakeholders
- Scientists working in the field
- Scientists outside the field
- People in fear
- People in panic
- People enraged

- Combinations of the above in the same person or in different people

Constructive and obsessive criticism in science

Vinay Prasad[1] | **John P. A. Ioannidis**[2]

[1]Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, California, USA

[2]Departments of Medicine, of Epidemiology and Population Health, of Biomedical Data Science, and of Statistics, and Meta-Research Innovation Center at Stanford (METRICS), Stanford University, San Francisco, California, USA

Correspondence
John P.A. Ioannidis, Departments of Medicine, of Epidemiology and Population Health, of Biomedical Data Science, and of Statistics, and Meta-Research Innovation Center at Stanford (METRICS), Stanford University, SPRC, 1265 Welch Rd, MSOB Room X306, Stanford, CA 94305, USA.
Email: jioannid@stanford.edu

Abstract

Social media and new tools for engagement offer democratic platforms for enhancing constructive scientific criticism which had previously been limited. Constructive criticism can now be massive, timely and open. However, new options have also enhanced obsessive criticism. Obsessive criticism tends to focus on one or a handful of individuals and their work, often includes ad hominem aspects, and the critics often lack field-specific skills and technical expertise. Typical behaviours include: repetitive and persistent comments (including sealioning), lengthy commentaries/tweetorials/responses often longer than the original work, strong degree of moralizing, distortion of the underlying work, argumentum ad populum, calls to suspend/censor/retract the work or the author, guilt-by-association, reputational tarnishing, large gains in followers specifically through attacks, finding and positing sensitive personal information, anonymity or pseudonymity, social media campaigning, and unusual ratio of criticism to pursuit of one's research agenda. These behaviours may last months or years. Prevention and treatment options may include awareness, identifying and working around aggravating factors, placing limits on the volume by editors, constructive pairing of commissioned editorials, incorporation of some hot debates from unregulated locations such as social media or PubPeer to the pages of scientific journals, preserving decency and focusing on evidence and arguments and avoiding personal statements, or (in some cases) ignoring. We need more research on the role of social media and obsessive criticism on an evolving cancel culture, the social media credibility, the use/misuse of anonymity and pseudonymity, and whether potential interventions from universities may improve or further weaponize scientific criticism.

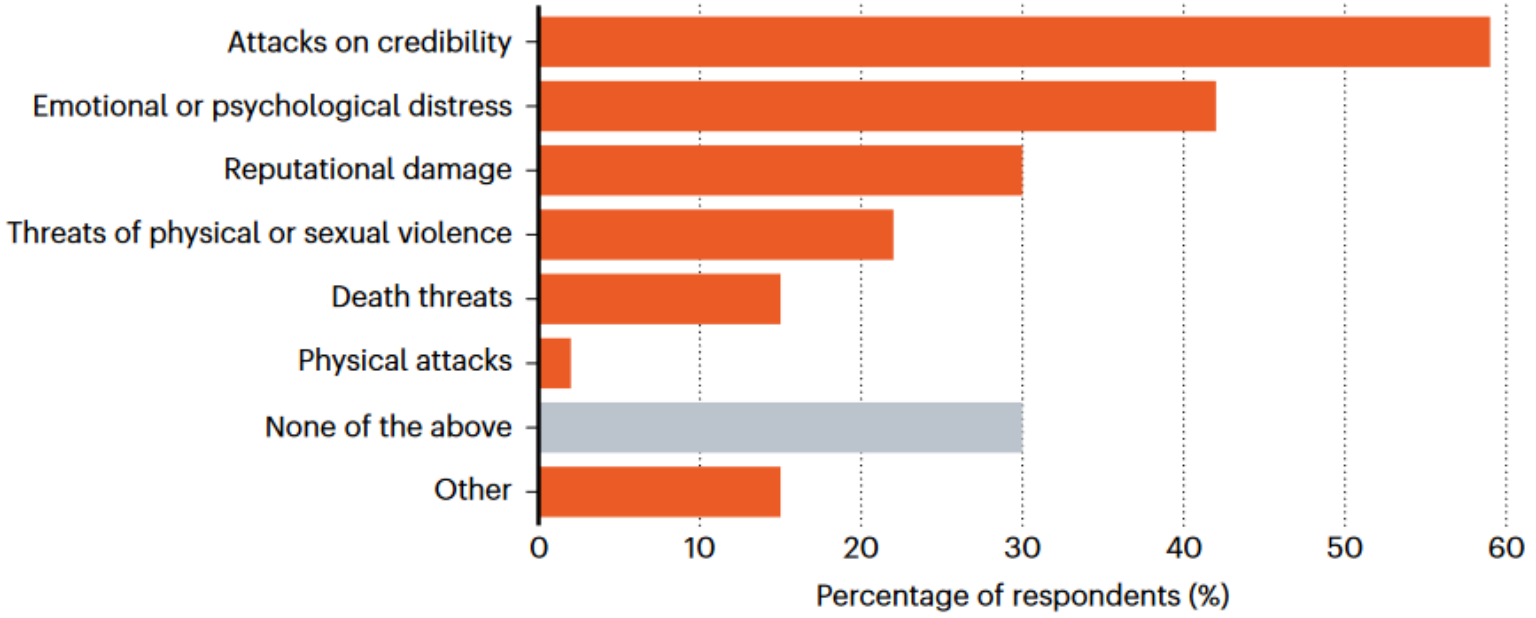
TABLE 2 Proposed diagnostic criteria for obsessive criticism

Focus	Focusing on one or a handful of individuals and their work, as opposed to collections of scientific papers that all point in a single direction
	Tinged with ad hominem: comments about the person who is authoring, including nature of their job, past work, past collaborations
	Lack of track record of field-specific skills and sufficient field-specific technical expertise
Behaviour	Repetitive and persistent comments, including sealioning
	Lengthy commentaries/ twitter threads/ rapid responses often several times longer than the original work
	Strong degree of moralizing: claiming the work will lead to evil or wrong policy choices
	Distortion of the underlying work/ strawman arguments
	Argumentum ad populum: claiming, without evidence, that most scientists disagree or believe the work is harmful
	Calls to suspend/ censor/ retract the work; and suspend/ censor or retract the speaker
	Guilt-by-association: claiming that since some nefarious groups enjoy the work, the work must be incorrect
	Reputational tarnishing of individuals and of their associates: distortion or misrepresentation of conflicts, speculation regarding true motives and funders
	Large gains in followers on social media platforms gained specifically through attacks
	Finding and positing sensitive personal information like home address or annual salary about the target
	Anonymity, pseudonymity, recruitment of fake accounts -- amplifying these accounts through retweets/ quote tweets
	Social media campaigning - Interaction with and retweeting accounts that parody or target the scientist as an individual or reiterate/echo some of the above-listed features
	Unusual ratio of criticism to pursuit of one's research agenda
Duration	The duration of this interaction often lasts months or years

NEGATIVE IMPACTS

In a *Nature* survey of scientists who have commented about COVID-19, 15% of 321 respondents said they had received death threats.

Question: Have you experienced any of the following negative impacts after speaking about COVID-19 to the media, or posting on social media? (You may select multiple options.)



Little pearls

- Death threats
- Death threatening events for family members
- Hit stories from journalists (e.g. “What is the most unethical and unbelievable attack you have received?”)
- Alluded conflicts of interest (e.g. “$5,000 for 200+ volunteer contributors)
- “Fact-checking” versus “investigation”
- Cancel campaigns (“Stalin”, “fire him”, “YouTube”, professional attacks)
- Left-wing or right-wing attacks? (“Send immigrants to barren islands”)
- Mis-characherization (“dubbed in Italian”)


OUTTOSEE
Home
JOHN IOANNIDIS
CHILDREN'S RIGHTS
SUPPORT
FILMTEAM
LANGUAGE
A Film with John Ioannidis

DOI: 10.1111/jep.13776

EDITORIAL

Journal of Evaluation in Clinical Practice
International Journal of Public Health Policy and Health Services Research
WILEY

An acceptance speech

John P. A. Ioannidis MD, DSc

Departments of Medicine, Epidemiology and Population Health, Biomedical Data Science and Statistics, Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, California, USA

Correspondence: John P. A. Ioannidis, MD, DSc, Departments of Medicine, Epidemiology and Population Health, Biomedical Data Science and Statistics, Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, CA, USA.
Email: jioannid@stanford.edu

KEYWORDS
evidence-based medicine, medical ethics, science

I am greatly honoured to receive the inaugural Harwood Prize for Intellectual Courage. Edward C. Harwood was a remarkable intellectual, thriving at Rensselaer and MIT before founding the American Institute for Economic Research (AIER). A West Point graduate, during World War II he was decorated for his courage with the Legion of Merit and a Bronze Star medal. I cautiously hope this information is correct, because I found it in Wikipedia. Conversely, personally I doubt my qualifications as an intellectual. For many years now, I have admitted in my Stanford webpage that I know next to nothing and the pandemic only made me even more aware that I know next to nothing. As for courage, every winter I fear the flu and struggle (often in vain) to avoid

Several years ago in an essay honouring my late mentor David Sackett I described myself as a failure, acknowledging my inability to counter the ongoing hijacking of evidence-based medicine. During the pandemic, hijacking escalated. Evidence became politicized, polarized, misinformed, disinformed beyond imaginable limits. Countering the devastation of evidence-based medicine almost became a mission impossible. I applaud the many scientists who worked dispassionately under unfavourable circumstances. Their brilliance and commitment saved lives and illuminated understanding.

Over the years I have received anonymous, pseudonymous and eponymous attacks and threats from Big Tobacco, spurious entre-

Massive publications are the norm now

March 20, 2023

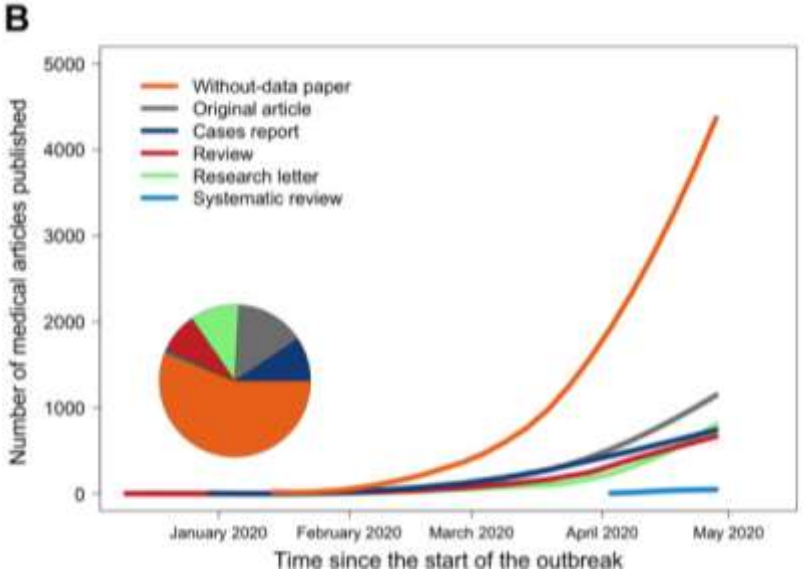
The Rapid Growth of Mega-Journals
Threats and Opportunities

John P. A. Ioannidis, MD, DSc[1,2]; Angelo Maria Pezzullo, MD, MSc[3]; Stefania Boccia, MSc, DSc, PhD[3,4]

» Author Affiliations

JAMA. 2023;329(15):1253-1254. doi:10.1001/jama.2023.3212

COVID-19 articles: most had no data (but they had strong opinions, even urging mandates)



Raynaud M et al. BMC Med Res Methodol 2021;21:1

Science and decision-making from models – models tramped evidence

European Journal of Epidemiology (2020) 35:733–742
https://doi.org/10.1007/s10654-020-00669-6

COVID-19

Check for updates

A case study in model failure? COVID-19 daily deaths and ICU bed utilisation predictions in New York state

Vincent Chin[1,2] · Noelle I. Samia[3] · Roman Marchant[1,2] · Ori Rosen[4] · John P. A. Ioannidis[5,6,7,8,9,10] · Martin A. Tanner[3] · Sally Cripps[1,2]

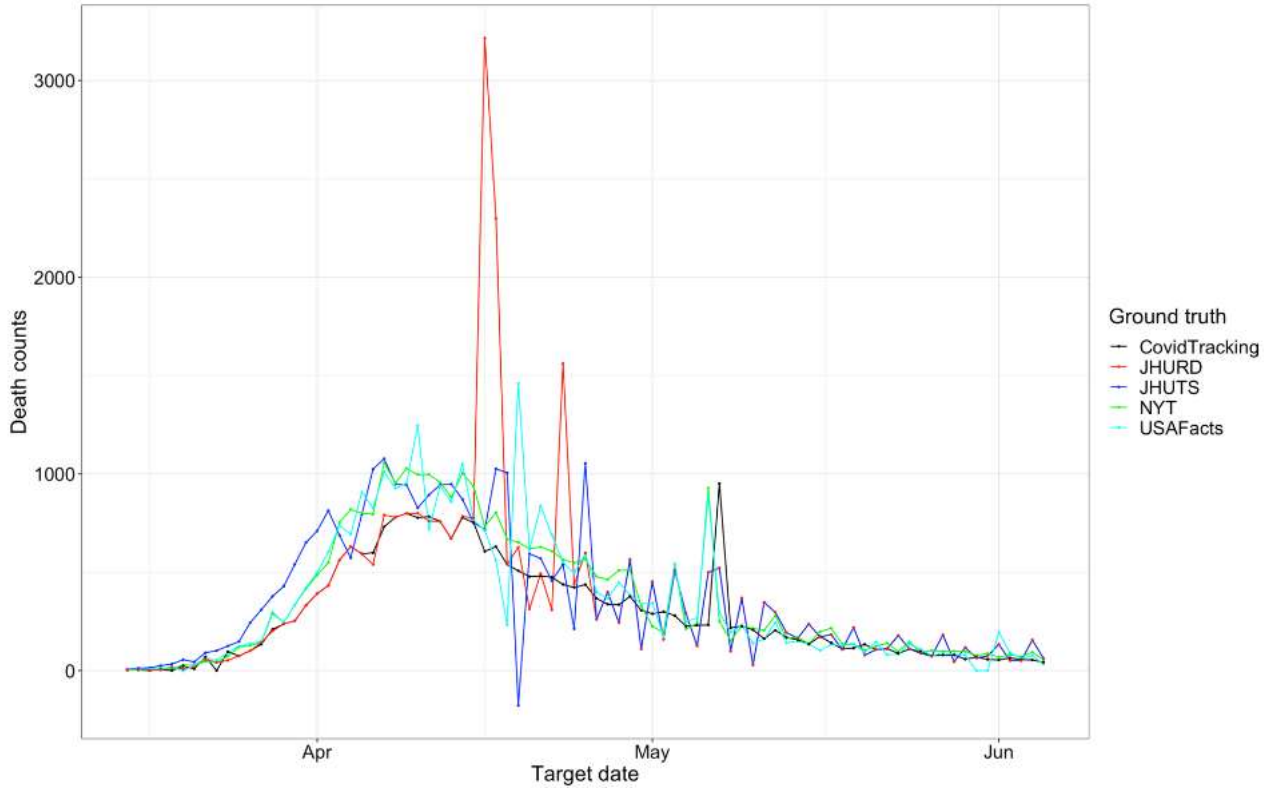


Fig. 1 A comparison of the daily death counts ground truth from CovidTracking (black), JHURD (red), JHUTS (dark blue), NYT (green) and USAFacts (light blue) for the period March 15–June 5 for NY

Broader considerations for failed forecasting

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Forecasting for COVID-19 has failed

John P.A. Ioannidis [a,*], Sally Cripps [b], Martin A. Tanner [c]

[a] Stanford Prevention Research Center, Department of Medicine, and Departments of Epidemiology and Population Health, of Biomedical Data Science, and of Statistics, Stanford University, and Meta-Research Innovation Center at Stanford (METRICS), Stanford, CA, USA

[b] School of Mathematics and Statistics, The University of Sydney and Data Analytics for Resources and Environments (DARE) Australian Research Council, Sydney, Australia

[c] Department of Statistics, Northwestern University, Evanston, IL, USA

Table 3
Potential reasons for the failure of COVID-19 forecasting along with examples and extent of potential amendments.

Reasons	Examples	How to fix: extent of potential amendments
Poor data input on key features of the pandemic that go into theory-based forecasting (e.g. SIR models)	Early data providing estimates for case fatality rate, infection fatality rate, basic reproductive number, and other key numbers that are essential in modeling were inflated.	May be unavoidable early in the course of the pandemic when limited data are available; should be possible to correct when additional evidence accrues about true spread of the infection, proportion of asymptomatic and non-detected cases, and risk-stratification. Investment should be made in the collection, cleaning, and curation of data.
Poor data input for data-based forecasting (e.g. time series)	Lack of consensus as to what is the 'ground truth" even for seemingly hard-core data such as the daily the number of deaths. They may vary because of reporting delays, changing definitions, data errors, etc. Different models were trained on different and possibly highly inconsistent versions of the data.	As above: investment should be made in the collection, cleaning, and curation of data.
Wrong assumptions in the modeling	Many models assume homogeneity, i.e. all people having equal chances of mixing with each other and infecting each other. This is an untenable assumption and, in reality, tremendous heterogeneity of exposures and mixing is likely to be the norm. Unless this heterogeneity is recognized, estimates of the proportion of people eventually infected before reaching herd immunity can be markedly inflated	Need to build probabilistic models that allow for more realistic assumptions; quantify uncertainty and continuously re-adjust models based on accruing evidence
High sensitivity of estimates	For models that use exponentiated variables, small errors may result in major deviations from reality	Inherently impossible to fix; can only acknowledge that uncertainty in calculations may be much larger than it seems

Lack of incorporation of epidemiological features	Almost all COVID-19 mortality models focused on number of deaths, without considering age structure and comorbidities. This can give very misleading inferences about the burden of disease in terms of quality-adjusted life-years lost, which is far more important than simple death count. For example, the Spanish flu killed young people with average age of 28 and its burden in terms of number of quality-adjusted person-years lost was about 1000-fold higher than the COVID-19 (at least as of June 8, 2020).	Incorporate best epidemiological estimates of age structure and comorbidities in the modeling; focus on quality-adjusted life-years rather than deaths
Poor past evidence on effects of available interventions	The core evidence to support "flatten-the-curve" efforts was based on observational data from the 1918 Spanish flu pandemic on 43 US cites. These data are >100-years old, of questionable quality, unadjusted for confounders, based on ecological reasoning, and pertaining to an entirely different (influenza) pathogen that had ~100-fold higher infection fatality rate than SARS-CoV-2. Even thus, the impact on reduction of total deaths was of borderline significance and very small (10%–20% relative risk reduction); conversely, many models have assumed a 25-fold reduction in deaths (e.g. from 510,000 deaths to 20,000 deaths in the Imperial College model) with adopted measures	While some interventions in the broader package of lockdown measures are likely to have beneficial effects, assuming huge benefits is incongruent with past (weak) evidence and should be avoided. Large benefits may be feasible from precise, focused measures (e.g. early, intensive testing with thorough contact tracing for the early detected cases, so as not to allow the epidemic wave to escalate [e.g. Taiwan or Singapore]; or draconian hygiene measures and thorough testing in nursing homes) rather than from blind lockdown of whole populations.

Reasons	Examples	How to fix: extent of potential amendments
Lack of transparency	The methods of many models used by policy makers were not disclosed; most models were never formally peer-reviewed, and the vast majority have not appeared in the peer-reviewed literature even many months after they shaped major policy actions	While formal peer-review and publication may unavoidably take more time, full transparency about the methods and sharing of the code and data that inform these models is indispensable. Even with peer-review, many papers may still be glaringly wrong, even in the best journals.
Errors	Complex code can be error-prone, and errors can happen even by experienced modelers; using old-fashioned software or languages can make things worse; lack of sharing code and data (or sharing them late) does not allow detecting and correcting errors	Promote data and code sharing; use up-to-date and well-vetted tools and processes that minimize the potential for error through auditing loops in the software and code
Lack of determinacy	Many models are stochastic and need to have a large number of iterations run, perhaps also with appropriate burn-in periods; superficial use may lead to different estimates	Promote data and code sharing to allow checking the use of stochastic processes and their stability
Looking at only one or a few dimensions of the problem at hand	Almost all models that had a prominent role in decision-making focused on COVID-19 outcomes, often just a single outcome or a few outcomes (e.g. deaths or hospital needs). Models prime for decision-making need to take into account the impact on multiple fronts (e.g. other aspects of health care, other diseases, dimensions of the economy, etc.)	Interdisciplinarity is desperately needed; as it is unlikely that single scientists or even teams can cover all this space, it is important for modelers from diverse ways of life to sit at the same table. Major pandemics happen rarely, and what is needed are models which combine information from a variety of sources. Information from data, from experts in the field, and from past pandemics, need to combined in a logically consistent fashion if we wish to get any sensible predictions.

Lack of expertise in crucial disciplines	The credentials of modelers are sometimes undisclosed; when they have been disclosed, these teams are led by scientists who may have strengths in some quantitative fields, but these fields may be remote from infectious diseases and clinical epidemiology; modelers may operate in subject matter vacuum	Make sure that the modelers' team is diversified and solidly grounded in terms of subject matter expertise
Groupthink and bandwagon effects	Models can be tuned to get desirable results and predictions; e.g. by changing the input of what are deemed to be plausible values for key variables. This is especially true for models that depend on theory and speculation, but even data-driven forecasting can do the same, depending on how the modeling is performed. In the presence of strong groupthink and bandwagon effects, modelers may consciously fit their predictions to what is the dominant thinking and expectations – or they may be forced to do so.	Maintain an open-minded approach; unfortunately, models are very difficult, if not impossible, to pre-register, so subjectivity is largely unavoidable and should be taken into account in deciding how much forecasting predictions can be trusted
Selective reporting	Forecasts may be more likely to be published or disseminated if they are more extreme	Very difficult to diminish, especially in charged environments; needs to be taken into account in appraising the credibility of extreme forecasts

Realizing that transparency matters

RESEARCH ARTICLE

A meta-epidemiological assessment of transparency indicators of infectious disease models

Emmanuel A. Zavalis[1,2☯], **John P. A. Ioannidis**[1,3☯*]

1 Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, California, United States of America, **2** Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, Solna, Stockholm, Sweden, **3** Departments of Medicine, of Epidemiology and Population Health, of Biomedical Data Science, and of Statistics, Stanford University, Stanford, California, United States of America

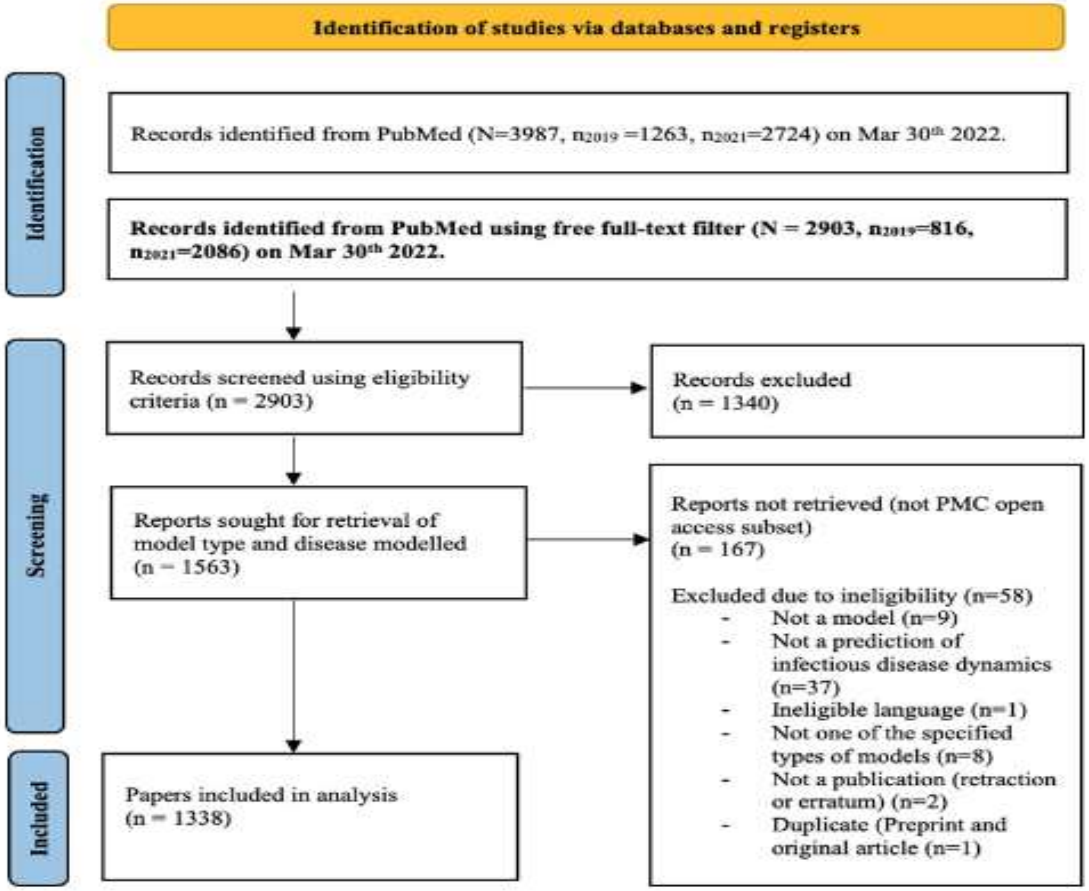


Fig 1. Flow chart for study selection.

https://doi.org/10.1371/journal.pone.0275380.g001

Table 2. Key transparency indicators overall and per year/COVID-19 focus.

N = 1338	Code sharing	Data sharing	Registration	COI	Funding
	N (%)	N (%)	N (%)	N (%)	N (%)
Overall	288 (21.5)	332 (24.8)	6 (0.4)	1197 (89.5)	1109 (82.9)
2019	38 (17.6)	59 (27.3)	3 (1.4)	197 (91.2)	202 (93.5)
2021	250 (22.3)	273 (24.3)	3 (0.3)	1000 (89.2)	907 (80.8)
COVID-19	207 (25.3)	199 (24.3)	0	730 (89.2)	635 (77.6)
non-COVID-19	43 (14.1)	74 (24.3)	3 (1)	270 (88.8)	272 (89.5)
Fisher's exact test (p-values)					
2019 vs 2021	0.15	0.35	0.06	0.45	1.0×10^{-6}
2019 vs 2021 non-COVID-19	0.33	0.48	0.70	0.46	0.12
2021 non-COVID-19 vs. COVID-19	5.1×10^{-5}	1	0.02	0.83	3.5×10^{-5}

COI: conflicts of interest

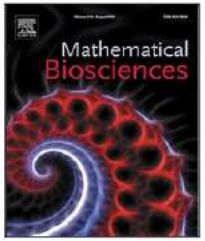
https://doi.org/10.1371/journal.pone.0275380.t002

Mathematical Biosciences 345 (2022) 108782

Contents lists available at ScienceDirect

Mathematical Biosciences

journal homepage: www.elsevier.com/locate/mbs



Perspective

Pre-registration of mathematical models

John P.A. Ioannidis *

Departments of Medicine, of Epidemiology and Population Health, of Biomedical Data Science, and of Statistics, and Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, CA, USA

ARTICLE INFO

Keywords:
Mathematical modeling
Pre-registration
Bias
Reproducibility

ABSTRACT

Pre-registration is a research practice where a protocol is deposited in a repository before a scientific project is performed. The protocol may be publicly visible immediately upon deposition or it may remain hidden until the work is completed/published. It may include the analysis plan, outcomes, and/or information about how evaluation of performance (e.g. forecasting ability) will be made, Pre-registration aims to enhance the trust one can put on scientific work. Deviations from the original plan, may still often be desirable, but pre-registration makes them transparent. While pre-registration has been advocated and used to variable extent in diverse types of research, there has been relatively little attention given to the possibility of pre-registration for mathematical modeling studies. Feasibility of pre-registration depends on the type of modeling and the ability

Table 1
Conditions that may favor or disfavor pre-registration.

Pre-registration favored	Pre-registration disfavored
Rigorous design thought in advance	Design to be fine-tuned iteratively
Standardized procedures preconceived	Procedures to be discovered
Optimal choices conceived in advance	Optimal choices unknown
Confirmatory research	Exploratory discovery research
Outcome/performance evaluation, e.g. forecasting	No outcome/performance evaluation
Projects can be separated into specific steps	Projects too chaotic even to specify steps
Data are to be collected prospectively	Existing data are used

Table 2
Potential advantages and disadvantages of pre-registration.

Potential advantages
Increased trust in research work
More objective assessment of model performance
Decrease in the possibility of bias/manipulation of results and inferences
Making research visible in public earlier
Reduction of redundancy in research efforts, better overall research agenda
Allowing to claim early credit for scientific work and ideas[a]
Potential disadvantages
Extra work needed
Fake pre-registration (registration has happened after the study was done)
Over-optimism that quality and efficiency of research would improve

Selective reporting (sadly) matters

ELSEVIER

Journal of Clinical Epidemiology 136 (2021) 96–132

ORIGINAL ARTICLE

Effect estimates of COVID-19 non-pharmaceutical interventions are non-robust and highly model-dependent

Vincent Chin[a,b], John P.A. Ioannidis[d,e,f,g,h,*], Martin A. Tanner[c], Sally Cripps[a,b]

Abstract

Objective: To compare the inference regarding the effectiveness of the various non-pharmaceutical interventions (NPIs) for COVID-19 obtained from different SIR models.

Study design and setting: We explored two models developed by Imperial College that considered only NPIs without accounting for mobility (model 1) or only mobility (model 2), and a model accounting for the combination of mobility and NPIs (model 3). Imperial College applied models 1 and 2 to 11 European countries and to the USA, respectively. We applied these models to 14 European countries (original 11 plus another 3), over two different time horizons.

Results: While model 1 found that lockdown was the most effective measure in the original 11 countries, model 2 showed that lockdown had little or no benefit as it was typically introduced at a point when the time-varying reproduction number was already very low. Model 3 found that the simple banning of public events was beneficial, while lockdown had no consistent impact. Based on Bayesian metrics, model 2 was better supported by the data than either model 1 or model 3 for both time horizons.

Conclusion: Inferences on effects of NPIs are non-robust and highly sensitive to model specification. In the SIR modeling framework, the impacts of lockdown are uncertain and highly model-dependent.

Different models, different inferences

Chin, Ioannidis, Tanner, Cripps.
J Clin Epidemiol 2021

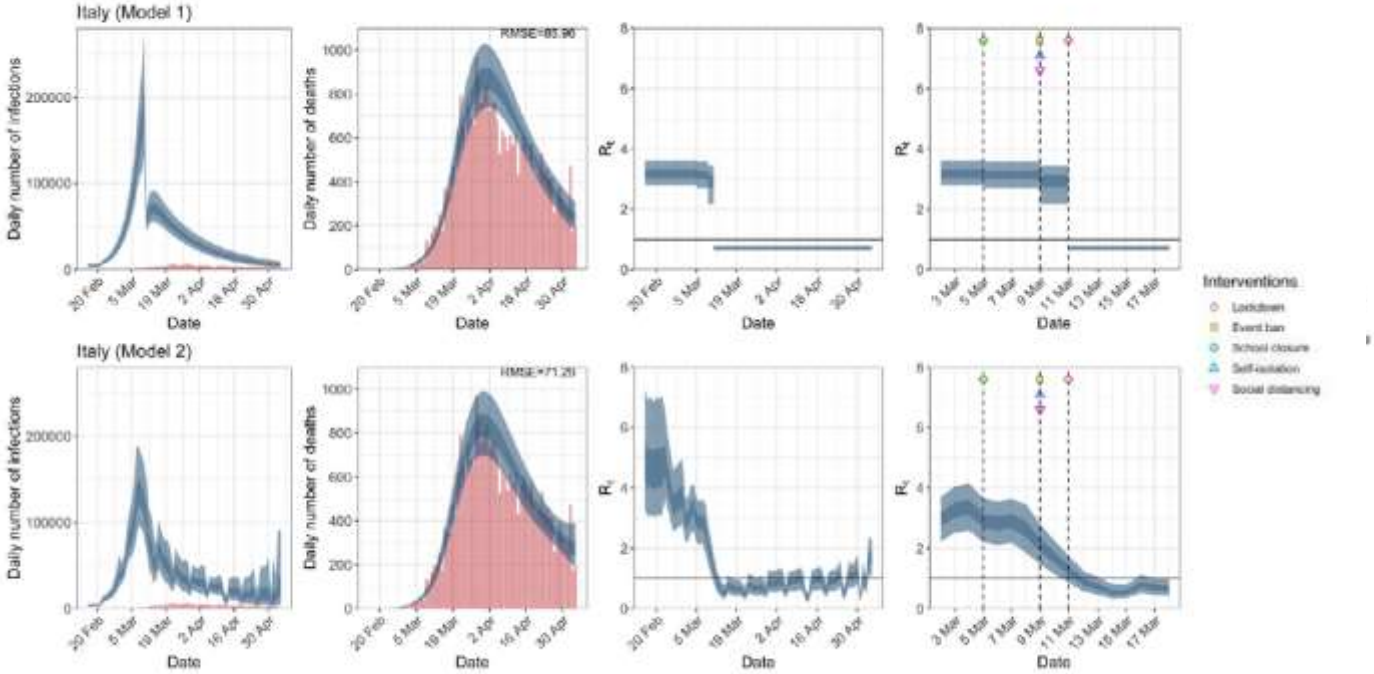


Table 2. Estimates and standard errors of the differences of various information criteria against model 1; the Watanabe-Akaike information criterion, $WAIC1 = -2\text{lppd} + 2p_{WAIC1}$ and $WAIC2 = -2\text{lppd} + 2p_{WAIC2}$ which uses lppd as a measure of fit with p_{WAIC1} and p_{WAIC2} as the effective number of parameters to penalize the fit respectively; the Deviance information criterion $\text{DIC} = -2\log p(\mathbf{y}|\hat{\theta}_{Bayes}) + 2p_{DIC}$ which uses $\log p(\mathbf{y}|\hat{\theta}_{Bayes})$, as the measure of fit, and p_{DIC} as the penalty. Note that a negative value implies a better predictive model compared to model 1, and the preferred model for each criteria and time period is shown in bold. See Appendix B for computational details.

Model	Time period	Δ_{WAIC1}	Δ_{WAIC2}	Δ_{DIC}
2	Up to May 5th	-31.21 ± 0.30	-29.95 ± 0.34	-30.46 ± 0.28
3	Up to May 5th	-24.03 ± 0.31	-22.49 ± 0.36	-23.29 ± 0.29
2	Up to July 12th	-54.27 ± 1.78	-49.93 ± 3.42	-51.95 ± 0.37
3	Up to July 12th	-36.74 ± 1.30	-32.24 ± 3.22	-34.97 ± 0.37

Table A.4. RMSE of daily death counts for models 1 and 2 for the data up to May 5th and July 12th. A lower RMSE between models 1 and 2 for each country is shown in bold.

Country	Up to May 5th		Up to July 12th	
	Model 1	Model 2	Model 1	Model 2
UK	**145.41**	145.64	134.26	**129.68**
Austria	**5.88**	**5.88**	**4.48**	4.57
Belgium	71.16	**52.91**	25.20	**15.84**
Denmark	3.27	**3.08**	2.42	**2.39**
France	242.07	**227.22**	187.33	**168.34**
Germany	**48.62**	48.75	37.04	**36.32**
Italy	85.96	**71.29**	63.47	**57.42**
Norway	**3.06**	3.07	**2.21**	2.22
Spain	95.23	**92.43**	143.82	**135.03**
Sweden	35.82	**35.55**	33.12	**33.09**
Switzerland	14.61	**14.34**	10.37	**10.31**
Greece			1.72	**1.51**
Netherlands			21.48	**21.01**
Portugal			6.29	**5.75**

Composite hard outcome: excess deaths

Environmental Research 213 (2022) 113754

Contents lists available at ScienceDirect

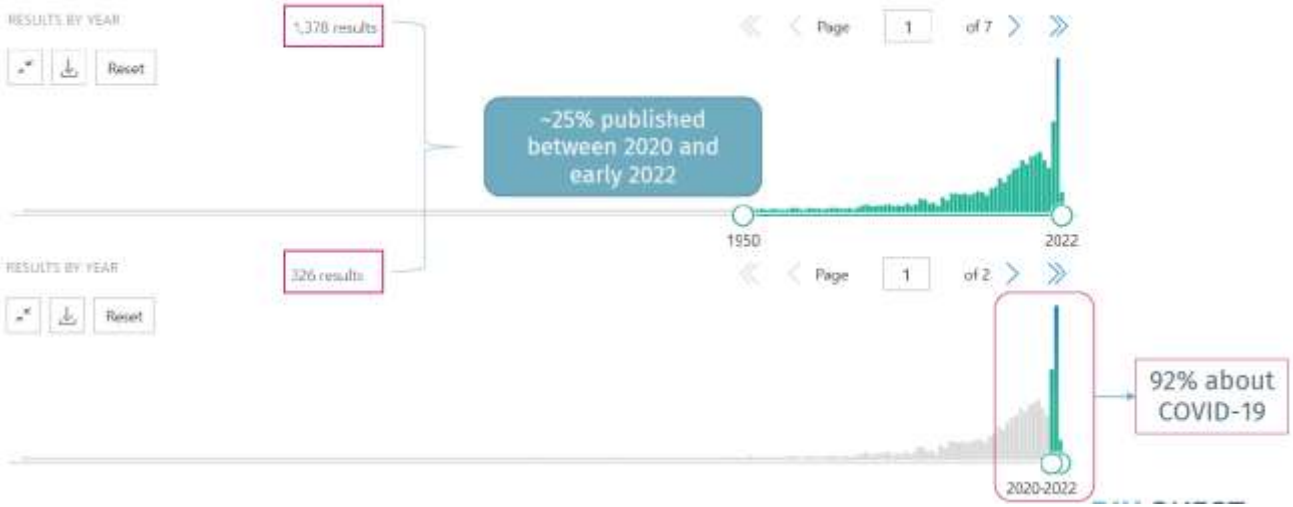
Environmental Research

journal homepage: www.elsevier.com/locate/envres

Comparison of pandemic excess mortality in 2020–2021 across different empirical calculations

Michael Levitt [a], Francesco Zonta [b], John P.A. Ioannidis [c,d,e,f,g,*]

An excess of excess mortality estimates



Excess mortality estimates

Eligible studies

- 160 studies assessing the excess mortality
    - Random sample of 60 studies

Expected mortality

- 6 (10%) studies used a single year
- 29 studies (48%) estimated the average across multiple years
- 25 studies (42%) applied a regression model or time-series analysis

Benchmarking of methods

- Only 2 studies (3%) compared excess mortality estimates based on different methods

Subgroup analyses

- age-specific estimates (n = 30, 50%)
- sex-specific estimates (n = 21, 35%)

Transparency

- Raw data availability (n = 25, 42%)
- Code availability (n = 5, 8%)

Beaney T et al. J R Soc Med 2020;113(9):329-334

Flaws and uncertainties in pandemic global excess death calculations

John P. A. Ioannidis[1] | Francesco Zonta[2] | Michael Levitt[3]

TABLE 1 Key issues and potential for correction or improvement in global estimates of excess deaths.

Issues	Potential for correction or improvement
Adjusting for changing population structure	Detailed adjustment of all excess death calculations for narrow age bins in the countries where these are available, so as to account for fine change in population structure over time
Adjusting for changes in other high-risk indicators	Capture and adjustment for other variables that affect mortality risk, in particular residence in long-term facilities (rates may have changed over time in various countries, with different patterns for elderly and for young residents)
Completeness corrections	Allowance for uncertainty in completeness corrections; consideration that completeness may have changed during pandemic years
Sensitivity to modelling choice	Consideration of different options regarding choice of pre-pandemic reference period and regarding imposed models; showing full range of results rather than single spuriously accurate average or weighted average.
Post hoc corrections in specific countries	Avoidance of post hoc corrections that are not based on pre-specified rules; pre-specification of objective, unambiguous criteria for any required post-modelling corrections
All-cause mortality modelling	Ensuring full transparency of model and model performance, including variance explained; exploration of transportability; acknowledgement of measurement errors and biases in included variables, consideration of alternative variables and models
Underestimation of uncertainty	Incorporation of uncertainty from each step in the modelling and from each of the variables considered; cautious interpretation since uncertainty may still be underestimated.
Excess death estimates per risk strata	Routine provision of excess death estimates per age group and according to other major risk strata (e.g. separately for community vs. long-term care resident population and per ethnic/racial strata)
Causal (mis)interpretation	Avoidance of causal statements of excess deaths attributed directly to SARS-CoV-2; consideration of direct and indirect effects of the pandemic and of the measures taken; in-depth assessment of causes and attribution will require other types of studies

Excess deaths during 2020-2021 in Germany

- Our age-adjusted estimate is **55,000** excess deaths
- Without age-adjustment we calculated **125,000** excess deaths
- Lancet calculated **203,000** excess deaths
- eLife calculated **88,000** excess deaths
- Economist calculated **113,000** excess deaths
- Baum (2022) calculated **22,000** excess deaths after age adjustment
- Koenig et al (2022) calculated ~**130,000** excess deaths without age adjustment
- The recorded COVID-19 deaths were **111,000**

- **In Germany, the number of people aged >80 years increased from 4.8 million in 2016 to 5.8 million in 2020, so consideration of age is crucial.**

Excess death estimates from multiverse analysis in 2009-2021

Michael Levitt,[a*] Francesco Zonta,[b] John P.A. Ioannidis[c]

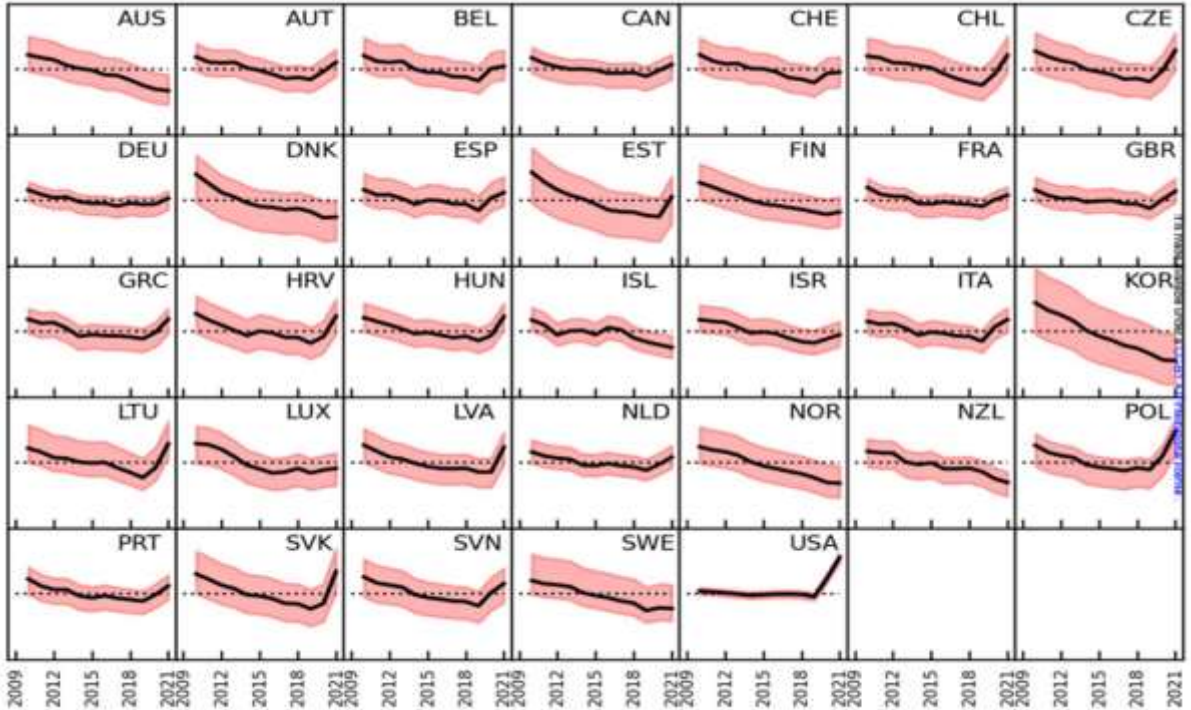


Figure 1

Randomized trials matter

Preserving equipoise and performing randomised trials for COVID-19 social distancing interventions

Ioana Alina Cristea[1], Florian Naudet[2] and John P. A. Ioannidis[3]

[1]Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy; [2]University Rennes, CHU Rennes, Inserm, CIC 1414 (Centre d'Investigation Clinique de Rennes), F-35000, Rennes, France and [3]Departments of Medicine, of Epidemiology and Population Health of Biomedical Data Science, and of Statistics, and Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, California, USA

Abstract

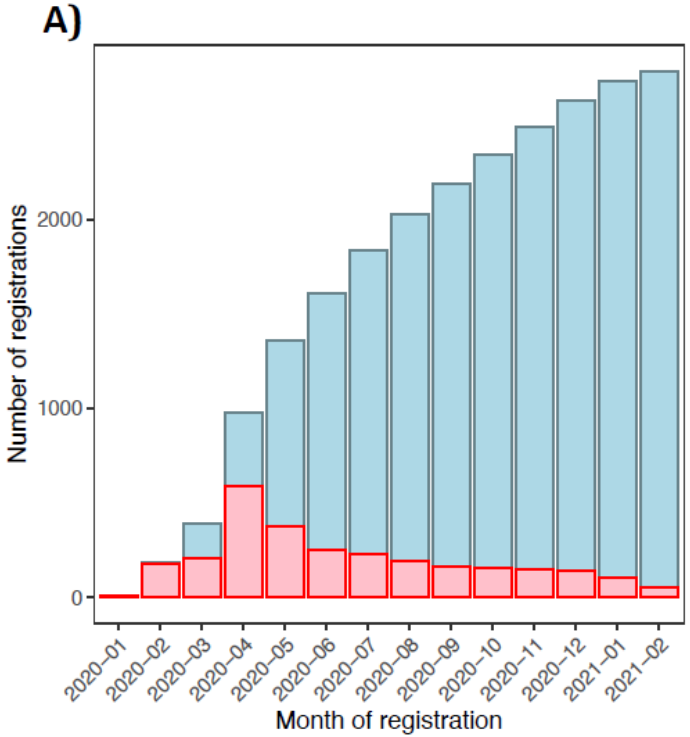
In the coronavirus disease 2019 (COVID-19) pandemic, a large number of non-pharmaceutical measures that pertain to the wider group of social distancing interventions (e.g. public gathering bans, closures of schools, workplaces and all but essential business, mandatory stay-at-home policies, travel restrictions, border closures and others) have been deployed. Their urgent deployment was defended with modelling and observational data of spurious credibility. There is major debate on whether these measures are effective and there is also uncertainty about the magnitude of the harms that these measures might induce. Given that there is equipoise for how, when and if specific social distancing interventions for COVID-19 should be applied and removed/modified during reopening, we argue that informative randomised-controlled trials are needed. Only a few such randomised trials have already been conducted, but the ones done to-date demonstrate that a randomised trials agenda is feasible. We discuss here issues of study design choice, selection of comparators (intervention and controls), choice of outcomes and additional considerations for the conduct of such trials. We also discuss and refute common counter-arguments against the conduct of such trials.

Large randomized trials were the greatest success story of pandemic research (even when they failed to show benefits)

RECOVERY
SOLIDARITY

>3,000 randomized trials registered on COVID-19 treatments in 1 year

>10,000 entries in clinicaltrials.gov as of 10/2023



Janiaud, Hemkens, Ioannidis, CJC 2021

How many trials materialized?

JAMA Network | Open

Research Letter | Statistics and Research Methods

Recruitment and Results Reporting of COVID-19 Randomized Clinical Trials Registered in the First 100 Days of the Pandemic

Perrine Janiaud, PhD; Cathrine Axfors, MD, PhD; John P. A. Ioannidis, MD, DSc; Lars G. Hemkens, MD, MPH

ELSEVIER

Journal of Clinical Epidemiology ■ (2022) ■

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Highly cited favorable studies for coronavirus disease 2019 treatments ineffective in large trials

John P.A. Ioannidis*

Departments of Medicine, of Epidemiology and Population Health, of Biomedical Data Science, and of Statistics, and Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, CA, USA

Table 1. Clinical studies with more than 150 Scopus citations that assess COVID-19 treatments that have shown no benefit in large trials (RECOVERY and SOLIDARITY)

Author (reference)	Interventions	*n*	RCT	>200 deaths	Favorable for index treatment	Citations
Cao [15]	LPV/r vs. SOC	199	Yes	No	Equivocal (benefit in MITT analysis)	2,859
Gautret [16]	HCQ ± AZ	38	No	No	Yes	2,839
Beigel [17]	Remdesivir vs. placebo	1,062	Yes	No	Yes	2,562
Wang [18]	Remdesivir vs. placebo	237	Yes	No	Equivocal (nonsignificant trend)	1,612
Grein [19]	Remdesivir	53	No	No	Yes	1,444
Shen [20]	Convalescent plasma	5	No	No	Yes	1,331
Duan [21]	Convalescent plasma	10	No	No	Yes	1,034
Geleris [22]	HCQ	1,446	No	No	No	931
Hung [23]	LPV/r + ribavirin + interferon vs. LPV/r	127	Yes	No	Yes	772
Boulware [24]	HCQ prophylaxis vs. placebo	821	Yes	No	No	688
Pan [14],[a]	Four active interventions (HCQ, remdesivir, lopinavir, and interferon) vs. control	11,330	Yes	Yes	No	646
Rosenberg [25]	HCQ, AZ, both, neither	1,438	No	Yes	No	625
Li [26]	Convalescent plasma vs. SOC	103	Yes	No	Equivocal (benefit in severe disease and for PCR conversion)	615
Goldman [27]	Remdesivir five vs. 10 days	397	Yes	No	No	562
Tang [28]	HCQ vs. SOC	150	Yes	No	No	552
Cavalcanti [29]	HCQ vs. HCQ + AZ vs. SOC	667	Yes	No	No	510
Molina [30]	HCQ + AZ	11	No	No	No	448
Horby [10],[a]	HCQ vs. SOC	4,716	Yes	Yes	No	430
Spinner [31]	Remdesivir vs. SOC	596	Yes	No	Equivocal (uncertain clinical value)	428
Gautret [32]	HCQ + AZ	80	No	No	Yes	396
Chen [33]	HCQ vs. control	30	Yes	No	No	322
Simonovich [34]	Convalescent plasma vs. placebo	228	Yes	No	No	311
Libster [35]	Convalescent plasma vs. placebo	160	Yes	No	Yes	276
Arshad [36]	HCQ, HCQ + AZ, AZ, neither	2,541	No	Yes	Yes	267
Agarwal [37]	Convalescent plasma vs. SOC	464	Yes	No	No	262
Million [38]	HCQ + AZ	1,061	No	No	Yes	246
Horby [9],[a]	LPV/r vs. SOC	5,040	Yes	Yes	No	236
Mahevas [39]	HCQ, control	181	No	No	No	235
Skipper [40]	HCQ vs. placebo	491	Yes	No	No	232
Zhang [41]	Convalescent plasma	4	No	No	Yes	231
Ye [42]	Convalescent plasma	6	No	No	Yes	222
Magagnoli [43]	HCQ+/-AZ, control	807	No	No	No	203
Zhou [44]	Interferon or interferon + Arbidol	77	No	No	Yes	198
Liu [45]	Convalescent plasma, control	39	No	No	Yes	195
Ahn [46]	Convalescent plasma	2	No	No	Yes	192
Joyner [47]	Convalescent plasma	5,000	No	No	Yes	191
Joyner [48]	Convalescent plasma	20,000	No	Yes	Yes	186
Zeng [49]	Convalescent plasma	6	No	No	Yes	180
Deftereos [50]	Colchicine vs. SOC	105	Yes	No	Yes	177
Saleh [51]	(HCQ or chloroquine) ±AZ	201	No	No	Yes	162

Table 2. Qualitative analysis of recent citations to the most highly cited article for each index treatment that reached favorable or equivocal conclusions

Intervention	Highly cited article	Critical citations (among 10 recent sampled citing articles)	RECOVERY/SOLIDARITY trials cited[a]
LPV/r	Cao, NEJM [15]	1/10 ("So far none of these drugs have been found to be an appropriate drug for COVID-19")	0/10
HCQ±AZ	Gautret, Intern J Antimicrob Agents [16]	3/10 ("So far none of these drugs have been found to be an appropriate drug for COVID-19" and "The excitement surrounding hydroxychloroquine was fueled early on by excessive media attention after a nonrandomized study (with questionable, hotly debated reliability) was released" and "The WHO announced the failure of the the solidarity trial, which means that hydroxychloroquine did not achieve the desired effect in the treatment of COVID-19")	2/10 (SOLIDARITY)
Remdesivir	Beigel, NEJM [17]	2/10 ("So far none of these drugs have been found to be an appropriate drug for COVID-19... WHO have made a conditional recommendation against the use of Remdesivir for hospitalized COVID patients, regardless of the disease's severity, because of a lack of evidence showing that it improves survival rate" and "had little or no effect on overall mortality, initiation of ventilation, or duration of hospital stay")	2/10 (SOLIDARITY)
Convalescent plasma	Shen, JAMA [20]	2/10 ("Small studies using convalescent serum for SARS-CoV-2 patients suggested that treatment was well tolerated, reduced viraemia and clinical symptoms [Shen et al., 2020, Duan et al., 2020], whereas the larger RECOVERY Collaborative Group [2021c], testing convalescent plasma as a treatment in life-threatening COVID-19 did not result in significant improvement and was discontinued early" and "the clinical effect of this CP intervention has not yet been determined, because patients could have recovered due to other treatments administrated in parallel")	1/10 (RECOVERY)
Interferon	Hung, Lancet [23]	0/10	2/10 (SOLIDARITY)
Colchicine	Deftereos, JAMA Network Open [50]	1/10 ("in this large, well powered trial, we found no evidence of a benefit from colchicine")[b]	1/10 (RECOVERY)

Examples of high-quality research

D1534–D1540 Nucleic Acids Research, 2021, Vol. 49, Database issue
doi: 10.1093/nar/gkaa952
Published online 9 November 2020

LitCovid: an open database of COVID-19 literature

Qingyu Chen †, Alexis Allot† and Zhiyong Lu *

National Center for Biotechnology Information (NCBI), National Library of Medicine (NLM), National Institutes of Health (NIH), Bethesda, MD 20892, USA

Received August 15, 2020; Revised October 02, 2020; Editorial Decision October 07, 2020; Accepted October 08, 2020

OPEN ACCESS

Check for updates

FAST TRACK

Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal

Laure Wynants,[1,2] Ben Van Calster,[2,3] Gary S Collins,[4,5] Richard D Riley,[6] Georg Heinze,[7] Ewoud Schuit,[8,9] Elena Albu,[2] Banafsheh Arshi,[1] Vanesa Bellou,[10] Marc M J Bonten,[9,11] Darren L Dahly,[12,13] Johanna A Damen,[8,9] Thomas P A Debray,[8,14] Valentijn M T de Jong,[8,9] Maarten De Vos,[2,15] Paula Dhiman,[4,5] Joie Ensor,[6] Shan Gao,[2] Maria C Haller,[7,16] Michael O Harhay,[17,18] Liesbet Henckaerts,[19,20] Pauline Heus,[8,9] Jeroen Hoogland,[8] Mohammed Hudda,[21] Kevin Jenniskens,[8,9] Michael Kammer,[7,22] Nina Kreuzberger,[23] Anna Lohmann,[24] Brooke Levis,[6] Kim Luijken,[10] Jie Ma,[5] Glen P Martin,[25] David J McLernon,[26] Constanza L Andaur Navarro,[8,9] Johannes B Reitsma,[8,9] Jamie C Sergeant,[27,28] Chunhu Shi,[29] Nicole Skoetz,[23] Luc J M Smits,[1] Kym I E Snell,[6] Matthew Sperrin,[30] René Spijker,[8,9,31] Ewout W Steyerberg,[2] Toshihiko Takada,[8,32] Ioanna Tzoulaki,[10,33] Sander M J van Kuijk,[34] Bas C T van Bussel,[1,35] Iwan C C van der Horst,[35] Kelly Reeve,[36] Florien S van Royen,[8] Jan Y Verbakel,[37,38] Christine Wallisch,[7,39,40] Jack Wilkinson,[25] Robert Wolff,[41] Lotty Hooft,[8,9] Karel G M Moons,[8,9] Maarten van Smeden[8]

OPEN ACCESS

Check for updates

Drug treatments for covid-19: living systematic review and network meta-analysis

Reed AC Siemieniuk,[1,2,*] Jessica J Bartoszko,[1,*] Dena Zeraatkar,[1,*] Elena Kum,[1,*] Anila Qasim,[1] Juan Pablo Díaz Martinez,[1] Ariel Izcovich,[1] Bram Rochwerg,[1,3] Francois Lamontagne,[4] Mi Ah Han,[5] Arnav Agarwal,[1,2] Thomas Agoritsas,[1,6] Maria Azab,[1] Gonzalo Bravo,[1] Derek K Chu,[1,7] Rachel Couban,[1] Ellen Cusano,[8] Tahira Devji,[1] Zaira Escamilla,[1] Farid Foroutan,[1,9] Ya Gao,[1] Long Ge,[10] Maryam Ghadimi,[1] Diane Heels-Ansdell,[1] Kimia Honarmand,[11] Liangying Hou,[1] Sara Ibrahim,[1] Assem Khamis,[12] Bonnie Lam,[1] Cristian Mansilla,[1] Mark Loeb,[1,7] Anna Miroshnychenko,[1] Maura Marcucci,[1,2] Shelley L McLeod,[13,14] Shahnaz Motaghi,[1] Srinivas Murthy,[15] Reem A Mustafa,[1,16] Hector Pardo-Hernandez,[17,18] Gabriel Rada,[19] Yamna Rizwan,[1] Pakeezah Saadat,[20] Charlotte Switzer,[1] Lehana Thabane,[1] George Tomlinson,[21] Per O Vandvik,[22,23] Robin WM Vernooij,[24,25] Andrés Vitesi-Garcia,[25,26] Ying Wang,[1] Liang Yao,[1] Yunli Zhao,[1] Gordon H Guyatt,[1,2] Romina Brignardello-Petersen[1]

REVIEW

Annals of Internal Medicine

Ventilation Techniques and Risk for Transmission of Coronavirus Disease, Including COVID-19

A Living Systematic Review of Multiple Streams of Evidence

Holger J. Schünemann, MD, PhD, MSc; Joanne Khabsa, MPH*; Karla Solo, MSc*; Assem M. Khamis, MD; Romina Brignardello-Petersen, DDM; Amena El-Harakeh, MPH; Andrea Darzi, MD, MPH; Anisa Hajizadeh, MPH; Antonio Bognanni, MD; Anna Bak, PharmD; Ariel Izcovich, MD; Carlos A. Cuello-Garcia, MD, PhD; Chen Chen, MM; Ewa Borowiack, MSc; Fatimah Chamseddine, MD; Finn Schünemann, MD; Gian Paolo Morgano, MSc; Giovanna E.U. Muti-Schünemann, Cand. Med; Guang Chen, MD, PhD; Hong Zhao, PhD; Ignacio Neumann, MD, PhD; Jan Brozek, MD; Joel Schmidt, MD; Loyal Hneiny, MPH, MLIS; Laila Harrison, MPH; Margo Reinap, MA; Matz Jurek, MD; Nancy Santesso, PhD, MLIS; Rayane El-Khoury, MPH; Rebecca Thomas, MPH, MBChB; Robby Nieuwlaat, PhD; Rosa Stalteri, BSHc; Sally Yaacoub, MPH; Tamara Lotfi, MD, MPH; Tejan Baldeh, MPH; Thomas Piggott, MD, MSc; Yuan Zhang, PhD, MSc; Zahra Saed, MSc; Brum Rochwerg, MD, MSc; Dan Perri, MD; Eddy Fan, MD; Florian Stehling, MD; Imad Bou Akl, MD; Mark Loeb, MD, MSc; Paul Garner, MD; Stephen Aston, MD; Waleed Alhazzani, MD, MSc; Wojciech Szczeklik, MD; Derek K. Chu, MD, PhD; and Elie A. Akl, MD, MPH, PhD

Meta-analyses: “negative” results are highly informative for evidence reversals

JAMA | Original Investigation

Association of Convalescent Plasma Treatment With Clinical Outcomes in Patients With COVID-19

A Systematic Review and Meta-analysis

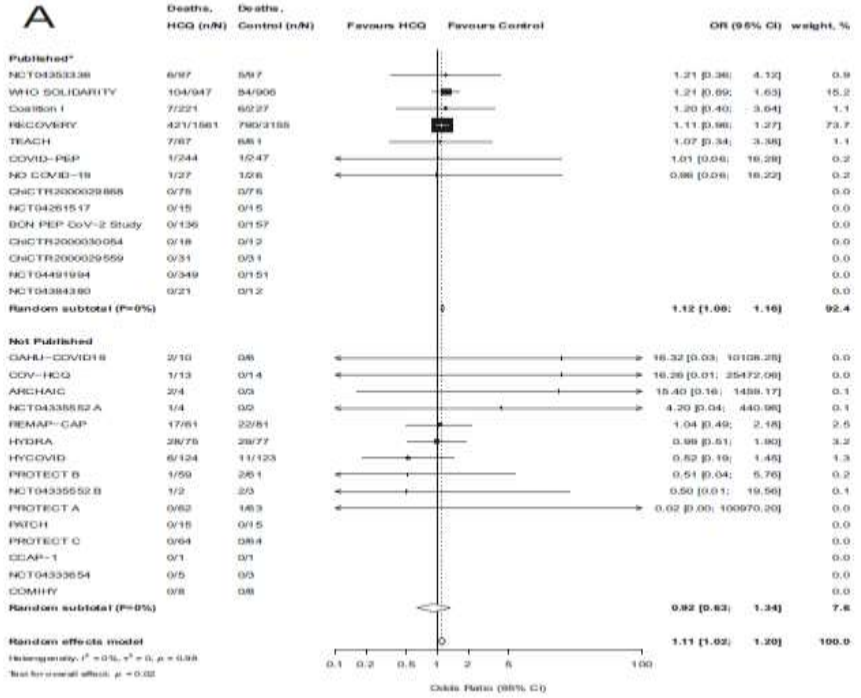
Perrine Janiaud, PhD; Cathrine Axfors, MD, PhD; Andreas M. Schmitt, MD; Viktoria Gloy, PhD; Fahim Ebrahimi, MD, MSc; Matthias Hepprich, MD; Emily R. Smith, ScD, MPH; Noah A. Haber, ScD; Nina Khanna, MD; David Moher, PhD; Steven N. Goodman, MD, PhD; John P. A. Ioannidis, MD, DSc; Lars G. Hemkens, MD, MPH

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-020-22446-z OPEN

Mortality outcomes with hydroxychloroquine and chloroquine in COVID-19 from an international collaborative meta-analysis of randomized trials



Eventually HCQ kills?

The triumph and disaster of observational data

EBM analysis

Factors influencing estimated effectiveness of COVID-19 vaccines in non-randomised studies

John P A Ioannidis

10.1136/bmjebm-2021-111901

Stanford Prevention Research Center, Department of Medicine and Department of Epidemiology and Population Health, and Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, California, USA

Correspondence to:
Dr John P A Ioannidis,
SPRC, Medical School

Abstract

Non-randomised studies assessing COVID-19 vaccine effectiveness need to consider multiple factors that may generate spurious estimates due to bias or genuinely modify effectiveness. These include pre-existing immunity, vaccination misclassification, exposure differences, testing, disease risk factor confounding, hospital admission decision, treatment use differences, and death attribution. It is useful to separate whether the impact of each factor admission decision, treatment use differences, and death attribution. Steps and measures to consider for improving vaccine effectiveness estimation include registration of studies and of analysis plans; sharing of raw

symptomatic, severe or any documented (including asymptomatic)), hospitalisations and deaths.

Factors influencing vaccine effectiveness estimates

Pre-existing immunity

Vaccine effectiveness may be adding only a small absolute benefit in people with some pre-existing immunity, while the benefit may be substantially larger in those without pre-existing immunity. The typical reason for pre-existing immunity is prior infection. Prior infection may or may not have been documented, since most infections remain undocumented.[7] The literature on the additional benefits of hybrid immunity (prior infection plus

Table 1 Factors influencing estimated COVID-19 vaccine effectiveness: outcomes affected and type of influence

Factors	Outcomes affected*	Influence on effectiveness estimate†
Pre-existing immunity		
Same in V and UNV	I, H, D	Genuine
Different in V and UNV	I, H, D	Spurious (selection bias)
Vaccination misclassification	I, H, D	Spurious (misclassification)
Exposure difference		
Induced by perceived vaccine protection	I, H, D	Genuine
Pre-existing, carried forward	I, H, D	Spurious (selection bias)
Testing		
Typical diagnosis bias	I, H, D	Spurious (selection bias, misclassification)
Affecting treatment	H, D	Genuine
Disease risk factor confounding	H, D	Spurious (confounding bias)
Hospital admission decision		
Induced by perceived vaccine protection	H, (D)‡	Genuine
Other reasons	H, (D)‡	Spurious (selection bias, confounding)
Treatment use difference		
Induced by perceived vaccine protection	D, (H)§	Genuine
Other reasons	D, (H)§	Spurious (selection bias, confounding)
Death attribution	D	Spurious (misclassification)

Table 2 Some measures to consider for improving the reliability of evidence from non-randomised studies of COVID-19 vaccine effectiveness

Measures to consider	Rationale	Challenges
Overarching measures		
Registration of studies and analysis plans	Allows to know what studies and analysis plans were preconceived and adhered to original plans and reduces degrees of freedom for data dredging	Most observational studies are non-registered or are registered after the analyses are done; there is debate on whether retrospective designs should/could be meaningfully registered; analytical plans are rarely registered in sufficient detail
Sharing of raw data and code	Allows independent validation of analyses and optimises the use of the data in overarching syntheses of data from multiple studies	Sharing has been limited for various reasons (privacy, consent and legal issues, as well as reluctance of primary investigators)
Better data collection		
Background collection of reliable information on seroprevalence, exposures, testing, disease risk factors, risk profiles on hospital admission and use of treatments	Allows for better adjustments and exploration of effect modification	Some of this information may be biased or very difficult to collect reliably
Blinded assessment of outcomes, for example, death causes	Allows removing some outcome misclassification biases	Blinding records requires time and resources and a committed effort
Better designs		
Use of maximal/best information in properly matched studies, multivariable analyses, propensity analyses and other models	Designs that consider and hopefully address more biases are better	Observational studies are unlikely to ever eliminate all possible biases
Performing randomised trials, whenever possible, for suitable questions (eg, use of booster doses, comparative effectiveness of different vaccination strategies)	Removes many of the biases	Reluctance to perform randomised trials when data suggest large efficacy (but this may be less of a concern for comparative effectiveness), randomised trials also have biases
Systematic review		
Living reviews and meta-analyses	Provide bird's eye view of evolving evidence	Meta-analyses have their own, long list of biases
Better communication		
Use of both relative and absolute metrics of risk reduction and presentation of uncertainty	Allows better comprehension of the magnitude of the benefit	Poor ability of many/most people to understand risks and other quantitative metrics
Avoidance of exaggeration in communicating results to the general public	Minimises misconceptions, confusion, panic (eg, from misleading claims of loss of vaccine effectiveness) or dangerous behaviour changes (eg, from misleading claims of retaining high effectiveness even with high exposures)	There is an avid market seeking immediate information on what is new on the pandemic and vaccines and sensationalism is prominent; the anti-vax movement makes confusion worse by adding extra misinformation

Decision-making (personal and public)

must be multi-dimensional

Cause of excess death	Reason/comments	Possible time horizon for excess deaths
People with AMI and other acute disease not given proper hospital care	Patients afraid to go to hospital and hospitals reducing admissions afraid of overload	Acute, during pandemic
People with cancer having delayed treatment	Postponement of cancer treatment in anticipation of COVID-19 overload	Next 5 y
Disrupted cancer prevention	Inability to offer cancer prevention services under aggressive measures	Next 20 y
Other healthcare disruption	Postponement or cancellation of elective procedures and regular care	Variable for different medical conditions
Suicides	Mental health disruption	Both acute and long-term
Violence (domestic, homicide)	Mental health disruption	Acute, possibly long-term
Starvation	Disruption in food production and transport	Acute, and possibly worse over next several years
Tuberculosis	Disruption of tuberculosis management programmes	Next 5 y
Childhood diseases	Disruption of vaccination programmes	Next 5 y
Alcoholism and other diseases of despair	Mental health disruption, unemployment	Next 10 y
Multiple chronic diseases	Unemployment, lack of health insurance and poverty	Next 20 y
Lack of proper medical care	Disruption of healthcare, as hospitals and health programmes get financially disrupted, furlough personnel or even shut down services	Next 20 y

Ioannidis, Eur J Clin Invest 2020

Can we affect factors underlying COVID-19 deaths?

- Social injustice, inequalities, racism, poverty
- Smoking
- Other modifiable risk factors/lifestyle (e.g., see obesity)
- Poor protection of nursing homes
- Poor adoption of effective public health measures
- Adoption of harmful, pro-contagion public health measures (e.g., blind draconian lockdowns)
- Suboptimal and harmful treatments and medical care
- Lack of effective vaccination, inefficient vaccination strategies

Slow data public health

Arnaud Chiolero[1,2,3] · Stefano Tancredi[1] · John P. A. Ioannidis[4]

Table 1 Three problems magnified by the pandemic and hampering the application of evidence-based and rigorous data-driven health decision-making

Problem	Characteristics	Solution
1. Confusion between surveillance and research	Poor knowledge by researchers of surveillance activities and of policymakers' needs	Increase surveillance culture among researchers
	Weak health data literacy of decision-makers	Foster collaboration between policymakers, surveillance experts, and researchers
		Rely for surveillance on independent and scientific institutions with expertise in epidemiology and surveillance methods
2. Big data do not speak by themselves	Poor quality of (organic) data	Train public health experts in measurement issues (methods, type of error, performance)
	Difficulty to characterize source population (selectivity bias)	Evaluate and document data quality systematically
	Diagnosis-based rather than population-based data	Characterize the study and target population, the sampling method, and the representativeness
		Improve data integrity, completeness, consistency, and quality
		Build surveillance systems to catch population-level data
3. Infodemic	High volume of data and information	Improve research quality and evidence synthesis production
	Multiple data sources and information producers	Track and debunk misinformation
	Misinformation spreading	Train policymakers in surveillance and health data science
	Doubts on the reliability of information and of experts, as well as on the independence of institutions producing information	Identify reliable experts and scientific institutions working in an evidence-based framework, not exposed to political influences

Research practices	Relevance for a slow data public health
Large-scale collaborative research	▪ Research and surveillance benefit from coordination of efforts and collaboration in the identification of needs with standardization in data collection methods across different sources ▪ Critical for comparisons and benchmarking
Adoption of replication culture	▪ To enhance reproducibility, especially under conditions of massive research outputs ▪ In a quality improvement framework, to provide feedback to surveillance systems for their continuous improvement
Containment of conflicted sponsors and authors	▪ To foster trust in surveillance expertise and evidence ▪ To protect surveillance activities from political influence ▪ To avoid academic militantism blurring the boundary between politics and science
More appropriate statistical methods, and standardization of definitions and analyses	▪ Highly relevant as data become more complex and error-prone and as many information producers are involved ▪ For surveillance, favor methods that are clear enough for dissemination to allow informed decision-making ▪ Give more weight to metrology training [18]
More stringent thresholds for claiming discoveries or "successes"	▪ Essential for efficient dissemination of information and to prevent wasting resources on trivial or biased information ▪ To prevent exaggerated information, excessive excitement, and eventual disappointment at the time of dissemination ▪ To enhance trust with proper and honest communication of uncertainty
Improvements in peer review, reporting, and dissemination of research	▪ For surveillance, the processes of reporting and dissemination have to be explicitly defined a priori ▪ Mediatization of surveillance and study results can create sensationalism and should be done cautiously – to avoid "medicine by press release" ▪ Requires independent and scientifically credible institutions with experts trained in epidemiology and surveillance methods [24]

Optimal research practices in slow data public health

Eur J Epi, 2023

Pandemic preparedness

- Pandemic preparedness requires evidence, reproducible evidence
- COVID-19 was a disaster for health, science, and evidence
- Hopefully we can learn from and revert the sad covidization of health, science, and evidence

Many thanks

- Cathrine Axfors, Stanford
- Ellen Kuhl, Stanford
- Lars Hemkens, U Basel
- Sally Cripps, U Sydney
- Martin Tanner, Northwestern
- Eran Bendavid, Stanford
- Prabhat Jha, U Toronto
- Jay Bhattacharya, Stanford
- Perrine Janiaud, Stanford and U Basel
- Michael Levitt, Stanford
- Ulrich Dirnagl, Charite and BIH
- Lu Tian, Stanford
- Bryan Cai, Stanford
- Vincent Chin, U Sydney
- Lazaros Belbasis, Charite BIH
- Stefan Pilz, U Graz
- Roberta Pastorino, U Rome
- Marcel Ballin, U Umea
- Peter Nordstrom, U Umea
- Ioana Cristea, U Pavia
- Michael Ewers, LMU Munich
- Michaele Schippers, Erasmus U Rotterdam
- Nicolaus Plesnila, LMU Munich
- Florian Naudet, U Rennes
- Steve Goodman, Stanford
- Janneke van't Hooft, Stanford and Amsterdam
- Despina Contopoulos-Ioannidis, Stanford
- David Grimes, Queen's U Belfast
- Stefania Boccia, U Rome
- Walter Ricciardi, U Rome
- Valentin Danchev, U Essex
- Mario Malicki, Stanford
- Neeraj Sood, U Southern California
- Matthias Peirlinck, Stanford and U Ghent
- Noah Haber, Stanford
- Michael Levitt, Stanford
- Angelo Pezzullo, U Rome
- Snezana Medic, Public Health Voivodina
- Athanasios Tsakris, U Athens
- Emmanuel Zavalis, Karolinska