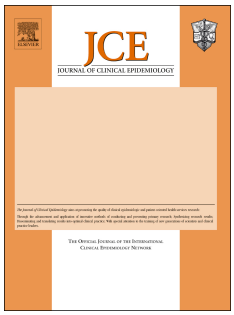


Minimal reporting improvement after peer review in reports of covid-19 prediction models: systematic review

Mohammed T. Hudda, Lucinda Archer, Maarten van Smeden, Karel G.M. Moons, Gary S. Collins, Ewout W. Steyerberg, Charlotte Wahlich, Johannes B. Reitsma, Richard D. Riley, Ben Van Calster, Laure Wynants

PII: S0895-4356(22)00323-7

DOI: https://doi.org/10.1016/j.jclinepi.2022.12.005

Reference: JCE 10974

To appear in: *Journal of Clinical Epidemiology*

Received Date: 13 September 2022

Revised Date: 29 November 2022

Accepted Date: 7 December 2022

Please cite this article as: Hudda MT, Archer L, van Smeden M, Moons KGM, Collins GS, Steyerberg EW, Wahlich C, Reitsma JB, Riley RD, Van Calster B, Wynants L, Minimal reporting improvement after peer review in reports of covid-19 prediction models: systematic review, *Journal of Clinical Epidemiology* (2023), doi: https://doi.org/10.1016/j.jclinepi.2022.12.005.

Minimal reporting improvement after peer review in reports of covid-19 prediction models: systematic review

Mohammed T Hudda*[1], Lucinda Archer*[2 3], Maarten van Smeden[4], Karel G M Moons[4 5], Gary S Collins[6 7], Ewout W Steyerberg[8], Charlotte Wahlich[1], Johannes B Reitsma[4], Richard D Riley[3], Ben Van Calster[8 9], Laure Wynants[8 10]

*Joint first authors

1 Population Health Research Institute, St George's University of London, Cranmer Terrace, London, England, SW17 0RE

2 Centre for Prognosis Research, School of Medicine, Keele University, Staffordshire, UK

3 Institute of Applied Health Research, University of Birmingham, Edgbaston, UK

4 Julius Center for Health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht University, Utrecht, Netherlands

5 Cochrane Netherlands, University Medical Centre Utrecht, Utrecht University, Utrecht, Netherlands

6 Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Musculoskeletal Sciences, University of Oxford, Oxford, UK

7 NIHR Oxford Biomedical Research Centre, John Radcliffe Hospital, Oxford, UK

8 Department of Biomedical Data Sciences, Leiden University Medical Centre, Leiden, Netherlands

9 Department of Development and Regeneration, KU Leuven, Leuven, Belgium

10 Department of Epidemiology, CAPHRI Care and Public Health Research Institute, Maastricht University, Peter Debyeplein 1, 6229 HA Maastricht, Netherlands

Correspondence to:
Dr Mohammed T Hudda
Population Health Research Institute
St George's University of London
Cranmer Terrace
London
SW17 0RE

Conflict of Interests: None

AbstractObjective: To assess improvement in the completeness of reporting COVID-19 prediction models after the peer review process.

Study Design and Setting: Studies included in a living systematic review of COVID-19 prediction models, with both pre-print and peer-reviewed published versions available, were assessed. The primary outcome was the change in percentage adherence to the TRIPOD reporting guidelines between pre-print and published manuscripts.

Results: 19 studies were identified including seven (37%) model development studies, two external validations of existing models (11%), and 10 (53%) papers reporting on both development and external validation of the same model. Median percentage adherence amongst pre-print versions was 33% (min-max: 10 to 68%). The percentage adherence of TRIPOD components increased from pre-print to publication in 11/19 studies (58%), with adherence unchanged in the remaining eight studies. The median change in adherence was just 3 percentage points (pp, min-max: 0-14pp) across all studies. No association was observed between the change in percentage adherence and pre-print score, journal impact factor, or time between journal submission and acceptance.

Conclusions: Pre-print reporting quality of COVID-19 prediction modelling studies is poor and did not improve much after peer review, suggesting peer review had a trivial effect on the completeness of reporting during the pandemic.

Key words: Peer review, reporting guidelines, prediction modelling, COVID-19, TRIPOD, adherence, prognosis and diagnosis

Running Title: Impact of peer review on reporting guideline adherence for COVID-19 prediction model reports

What's New?

In this study, we compared adherence to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement for studies developing or validating COVID-19 prediction models with a pre-print version available, before and after the peer review process

The findings of this report demonstrate a poor quality of reporting of COVID-19 prediction modelling studies amongst pre-print versions, which did not improve much following peer review.

Most TRIPOD items saw no change in the frequency of their reporting, with only the coverage of discussion items being substantially improved in the published version.

While it has previously been reported that the adherence to reporting guidelines for prediction modelling studies was poor, our findings also suggest that the peer review process had little impact on this adherence during the pandemic.

The implication of these findings is that a greater focus is needed on the importance of adhering to reporting guidelines by authors as well as checking during the editorial process.

Introduction

The coronavirus (COVID-19) pandemic presents a serious and imminent challenge to global health. Since the outbreak of COVID-19 in December 2019, in excess of 529 million cases have been confirmed in over 200 countries with over 6 million deaths [1]. Given the burden this pandemic has placed on health care systems around the world, efficient risk stratification of patients is crucial.

Diagnostic and prognostic prediction models combine multiple variables in order to estimate the risk that a specific event (e.g., disease or condition) is present (diagnostic) or will occur in the future (prognostic) in an individual [2]. In 2015, in an effort to improve the completeness of reporting of studies developing or validating prediction models, the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) Statement was developed [2, 3] and subsequently endorsed by several journals and included in the

library of the Enhancing the QUAlity and Transparency Of health Research (EQUATOR) network, an international initiative that seeks to promote transparent and accurate publishing via the use of reporting guidelines.

However, previous studies have demonstrated incomplete reporting of studies developing or validating prediction models [4, 5]. One study assessing 170 prediction model studies, published just prior to the introduction of the TRIPOD statement in articles across a wide range of clinical domains, found the median percentage of TRIPOD items reported to be just 44% [4]. More recently, a living systematic review of all COVID-19 related prediction models critically evaluated over 230 models, and concluded that almost all of the models were poorly reported and were subject to a high risk of bias [5].

Poor reporting is considered to be a contributing factor to the high proportion of avoidable clinical research waste [6, 7], an issue that was estimated to affect up to 85% of all research prior to the COVID-19 pandemic [8]. It has been suggested that research waste increased during the pandemic for a number of reasons, including the rise in the widespread use of pre-print servers to communicate COVID-19 related research findings, prior to journal peer review [9]. As COVID-19 swept the globe, the use of pre-print servers to disseminate research findings skyrocketed in an effort to mitigate against delays in the dissemination of research that traditional publications faced. However, it was suggested that pre-prints prior to peer review led to irresponsible dissemination of flawed research, and to such poor reporting that a critical appraisal of the methodology and trustworthiness of results became difficult [9]. The vast number of pre-prints made public during the pandemic provides a unique opportunity to investigate whether these concerns are justified. It is of interest to compare the completeness of reporting of COVID-19 prediction models between pre-print and peer reviewed publications to identify areas related to the development, validation, and subsequent reporting of prediction models which were improved upon following the peer review process and crucially, areas which were not improved.

We therefore investigated whether peer reviewed articles are different in their completeness of reporting of COVID-19 prediction models as compared to non-peer reviewed versions of the same article. In particular, we aim to explore the adherence to the TRIPOD Statement for studies developing diagnostic or prognostic COVID-19 prediction models when released as a pre-print and compare this to the published article following peer review.

Methods

This study is based on articles identified and included in a living systematic review of all COVID-19 related prediction models by Wynants et al. (3rd update, published January 2021) [5]. Studies were included in the living systematic review if they developed or validated a multivariable model or scoring system, based on individual participant level data, to predict any COVID-19 related outcome. These models included three types of prediction models: diagnostic models to predict the presence or severity of COVID-19 in patients with suspected infection; prognostic models to predict the course of infection in patients with COVID-19; and prediction models to identify people in the general population at risk of COVID-19 infection or at risk of being admitted to hospital with the disease [5]. The latest update of the living systematic review included all publications identified from database searches repeatedly conducted up to 1st July 2020. Further details of the databases, search terms, and inclusion criteria used for the living review have been published previously [5]. Aspects of both the PRISMA (preferred reporting items for systematic reviews and meta-analyses) [10] and TRIPOD [2, 3] guidelines were considered when reporting our study.

Inclusion criteria

For this study, we included all published articles (post peer review) from the latest version of the living systematic review [5], with an earlier pre-print version of the same article available on either the arXiv, medRxiv, or bioRxiv servers. Additionally, we included all pre-print articles which were included in the living systematic review and were subsequently published in a peer

reviewed journal (search dates for published versions: July-September 2021). If more than one pre-print version of a report was available on the respective server, the most recent version was reviewed. An exception was made for pre-print versions posted after the date of submission to the journal to ensure the pre-print version being included was drafted prior to peer review. We included pre-prints published a maximum of two days following the date of the first submission to the journal (to allow a window for pre-prints to become available following submission to the server). We excluded reports with: missing information on the pre-print upload date, manuscript first submission date, prediction models based upon only imaging/audio recording data, and substantial changes to the aims/objectives of the study between pre-print and published reports.

Data extraction and calculating adherence to TRIPOD

The data extraction form (www.tripod-statement.org/wp-content/uploads/2020/01/TRIPOD-Adherence-assessment-form_V-2018_12.xlsx) consists of 22 main components which appear on the TRIPOD checklist [2, 3]. They relate to items that should be reported in the title and abstract, introduction, methods, result and discussion of a prediction modelling study, as well as additional information on the role of funding and supplementary material. Several of the 22 components are further split into items, that all need to be adhered to in order to achieve adherence to that specific component. For studies reporting on both model development and validation, there are a maximum of 36 relevant items of adherence. For studies focusing on model development only and for studies solely focussed on external validation only, there are a maximum of 30 relevant items of adherence [2, 3].

Data were extracted by one reviewer (MTH), from both the pre-print and published version of the article. A second reviewer (LA) independently extracted data of both versions of nine articles (five randomly selected articles, plus all studies with increases in adherence from pre-print to published version of >10% (three), and the one study which appeared to have a poorer adherence in the published version than the pre-print). Minor discrepancies in data extraction of three of the pre-print versions were observed; these were discussed and resolved between

the two reviewers without the need for reference to a third opinion. Given the consistency in extraction between the two reviewers, it was concluded that single extraction on the remaining 10 articles was sufficient.

Based on the data extraction for the 21 main components of the TRIPOD checklist (excluding Q21 relating to Supplementary Information which is not included in the TRIPOD scoring calculation), both versions of each article were given an adherence score as detailed by Heus et al [11]. A score of 1 was assigned if an item was adhered to in the article and a score of 0 was assigned for non-adherence. An article's overall TRIPOD adherence score was calculated by dividing the sum of the adhered TRIPOD items by the total number of applicable TRIPOD items for that article. Percentage adherence (%adherence) was then calculated. The above process was carried out for the pre-print and the final published versions of each article.

Outcome measures and statistical analysis

The primary outcome of this study was the change in percentage adherence score between the pre-print version and the published version of each article.

We summarised and described the percentage adherence from pre-print and published versions of the same article and, in particular, the change in adherence (in percentage points, pp) between these two versions for each study. We further summarised the percentage adherence across all studies, by type (pre-print and published manuscript). The proportions of pre-prints and published articles reporting individual components of the TRIPOD checklist were described, and items where reporting changed substantially between the two versions were identified. We assessed whether there was an association between the published version adherence score and the following factors of interest, after adjustment for pre-print adherence score: (i) impact factor of the journal which published the article, (ii) time between the first submission to the journal and acceptance date, (iii) number of pre-print updates made prior to journal submission, (iv) number of days between the start of the pandemic (11th March 2020 as per WHO declaration) and manuscript acceptance, or (v) the presence of a statement

relating to the use of the TRIPOD reporting guidelines in the journal's instructions to authors. To quantify these associations, we fit a series of linear regression models with the adherence score of the published version (re-scaled to be a decimal between 0 and 1) as the dependent variable and each of the above factors of interest (one at a time) as the independent variable. Models were adjusted for the baseline adherence score from the pre-print version (re-scaled to be a decimal between 0 and 1) in the regression model.

For reports that were published in a journal with an open peer review process, we also explored the: (i) the background of the reviewers, (ii) the number of reviewers / peer review rounds, (iii) the explicit mention of TRIPOD guidelines/checklist within the peer review document, and (iv) whether key reporting items/elements on the TRIPOD checklist were raised by reviewers (even if TRIPOD was not explicitly mentioned).

Results

Of the 169 studies in the Wynants et al review [5], 58 (34%) were identified as having both published and pre-print versions. Of these, 26 papers related to diagnosis or prognosis through imaging, one concerned prediction via audio recording, and eight had either missing information on their initial submission date or uploaded the pre-print version more than two days (min-max: 11 to 63 days) after submission of the published version and thus were excluded from our analyses. A further four studies made substantial changes in aims between pre-print and publication, three of which contained an external validation in the published version which was previously not present in the pre-print. These three authors were contacted in order to ascertain whether these amendments were pre-planned; two of these studies replied and confirmed that the addition of the external validation was pre-planned and added due to external datasets becoming available at a later date. All four of these studies were excluded from this study. This left 19 papers for our analyses: seven (37%) model development studies (possibly with internal validation), two external validation of an existing

model (11%), and 10 (53%) papers reporting on both model development and external validation (see figure 1).

The majority of studies (17, 89%) uploaded only one pre-print prior to publication, with one paper (5%) uploading two and one uploading three versions of their pre-print (Table 1). The median number of days between first submission to the publishing journal and acceptance was 80 (Lower Quartile (LQ) – Upper Quartile (UQ): 40 to 187) and ranged from 22 to 259 days. The journal impact factor of published articles ranged from 2.7 to 39.9 with an average of 4.2 (LQ-UQ: 3.6 to 6.8), with most studies being published in journals with no reference to TRIPOD or EQUATOR on their websites (11 studies, 58%).

Changes in completeness of reporting following peer review

Adherence to TRIPOD was highest on average in the published manuscripts, with a median adherence of 42% (LQ-UQ: 31% to 57%), ranging from a low of 10% to a high of 71%. In contrast, median adherence across pre-prints was 33% (LQ-UQ: 30% to 50%), ranging from 10% to 68%. Overall adherence was low, with only 4 pre-prints and 6 published manuscripts adhering to more than 50% of the items in the TRIPOD checklist (see figure 2). Adherence increased from pre-print to publication in 11 (58%) of the studies assessed, with a median change in adherence of 3 percentage points (pp) across all 19 studies (see Figure 3; LQ-UQ: 0 to 7). While no study had a lower adherence overall in their published manuscript than in their pre-print, 8 (42%) of the studies assessed showed no change in total %adherence following peer review.

Unsurprisingly, there was a strong positive correlation between pre-print and published version adherence scores (Pearson's correlation coefficient=0.96). After adjusting for the pre-print adherence score, there was no evidence of meaningful or statistically significant associations between published article adherence score and; journal impact factor, time between journal submission and acceptance, the number of pre-prints uploaded prior to journal submission, the number of days between the start of the pandemic and manuscript

acceptance, or the presence of a statement relating to the use of the TRIPOD in the journal's instructions to authors (table 2).

The median adherence of published articles was similar across the 6 journals that mentioned TRIPOD on their website, and the 13 journals that did not mention TRIPOD on their website (44% vs 42% respectively). Median adherence was 49% in the 2 papers published in journals mentioning alternative reporting guidelines (EQUATOR [12] and STARD [13]), and 40% for 11 papers in journals with no reference to any reporting guidelines.

Reporting of individual TRIPOD items

At pre-print, only 5/36 TRIPOD adherence items were reported in more than 75% of the included studies (6/36 items for published manuscripts), with nine items reported in less than 20% (fewer than three) of either published or pre-print manuscripts (see supplementary table S2). Five items were reported less frequently in publications than in pre-prints (supplementary table S3), including those surrounding details on how predictors were measured and reporting of the full model equation. Additional details are given in Appendix 1 of Supplementary Material.

Assessment of open peer review

Open peer review was available for only 5/19 manuscripts [14-18]. Peer reviewers asked for a median of 4 TRIPOD items (min-max: 2 to 8), but only one reviewer mentioned TRIPOD explicitly. Additional results are given in Appendix 2 of Supplementary Material.

Discussion

In this study, we have compared adherence to the TRIPOD statement for studies developing or validating COVID-19 prediction models with a pre-print version available, before and after the peer review process. Peer reviewed articles showed low adherence and a modest

improvement in adherence to TRIPOD reporting items compared to pre-print versions of the same article.

Most TRIPOD items saw no change in the frequency of their reporting, with only the coverage of discussion items (including implications of research, and the availability of supplementary resources) being substantially improved in the published version. This highlights the need for a greater focus on the importance of adherence to reporting guidelines by authors and checking during the editorial process.

Comparison to other studies

Previous studies have shown that reporting of prediction model research has generally been poor across medical disciplines [6, 19, 20]. This contributes to a vast amount of avoidable research waste [7, 8], as the lack of transparency may obscure study biases and hinder reproducibility of the reported results. The results of the current analysis, where we observed a median adherence of 42% (LQ-UQ: 31 to 57%) amongst published articles, were in line with findings from an earlier study by Heus et al. [4] which reported a median adherence of 44% (LQ-UQ: 35 to 52) in 170 models across 37 clinical domains, in articles published prior to the publishing of the TRIPOD statement. Since the introduction of the TRIPOD statement [2, 3], continued poor reporting has been observed [21-23], with the present analysis suggesting that current reporting quality in the field of COVID-19 prediction is no different to the situation prior to publication of the TRIPOD statement. This is likely due to the less rigorous peer review and editorial process for COVID-19 articles during the pandemic. When considering the reporting of individual TRIPOD items, our findings are consistent with similar reviews in different clinical areas. For example, Jiang et al 2020, when discussing reporting completeness of published models for melanoma prediction, also concluded that titles and abstracts were among the worst reported items from the TRIPOD checklist, and that the discussion items (such as giving an overall interpretation of the results, item 19b) were most commonly reported [24].

Only 2 studies included in the present analysis were focused specifically on an external validation of an existing model; one of which had the best adherence and the other being amongst the reports with the poorest adherence. These findings are in agreement with previous research suggesting that external validation of multivariable prediction models is reported as equally poorly as model development studies [25].

Regarding the change in completeness of reporting between pre-print and published manuscripts of the same article, our analyses showed an improvement in reporting between pre-print and published versions in 11/19 (58%) studies with a median improvement across all included studies of just 3pp. While this small magnitude of improvement in completeness of reporting is in line with findings by Carneiro et al., where a difference in %reporting score of 4.7pp was observed between pre-prints and published articles in 2016 [26], it is in contrast to findings from other pre-pandemic studies which concluded substantially improved quality of reporting in the published manuscript [27, 28].

Furthermore, while the main analysis focussed on comparing overall adherence between pre-print and published reports, several individual TRIPOD items were reported less frequently in the published version than the pre-print, with other items being reported more frequently.

Of the studies included in this review, 12/19 were assessed for risk of bias using PROBAST at both pre-print and publication, as a part of the COVID-19 living systematic review by Wynants et al [5]. It was established that all 12 studies were at high risk of bias at pre-print, while 11/12 were at high risk of bias at publication. Only Carr et al 2021 [15] improved in overall risk of bias category between pre-print and publication, due to the inclusion of an external validation and the resolution of some unclear reporting in the published report. While this study was among the better reported studies included in our analysis, the percentage adherence in fact did not change between pre-print and publication. Given the strong overlap between the PROBAST risk of bias assessment and the TRIPOD checklist, with many of the items in the TRIPOD checklist being a requirement to conclude low risk of bias, it is unsurprising that our assessment of TRIPOD adherence was highly consistent with the risk of bias assessment.

Given the limited number of open peer reviews available for our included studies, we were unable to draw any meaningful conclusions about the impact of the background of the reviewers, or whether reviewers identified key reporting issues, on changes to TRIPOD adherence. Previously, reviewers from statistical backgrounds have been seen to improve reporting in biomedical articles, over-and-above using solely field experts [29], and targeted reviews to identify missing items from reporting guidelines were found to improve manuscript quality [30]. However, the available open reviews suggest that statistical review of prediction models seems to be rare in the field of COVID-19 prediction, and just one of the peer reviewers' comments included mention of the TRIPOD checklist.

Strengths and limitations of this study

The current study has several strengths. Firstly, we obtained an objective assessment of reporting quality, using the TRIPOD adherence checklist (www.tripod-statement.org/wp-content/uploads/2020/01/TRIPOD-Adherence-assessment-form_V-2018_12.xlsx) to assess adherence for all included articles. Furthermore, extraction of adherence information was performed by medical statisticians with familiarity of the TRIPOD checklist and extensive experience in the field of prediction modelling research, to minimise the risk of incorrect assessment. Finally, this study included all studies from the latest version of a living systematic review of COVID-19 prediction models with an earlier pre-print version of the same article available, and thus gives a complete view of state of the current literature.

There are also a number of limitations to the current study, in particular given the TRIPOD adherence checklist can be quite unforgiving in certain aspects. When calculating adherence, each of the 22 TRIPOD components are given the same weighting and importance, for example a clearly reported title is given equal importance to clearly reported statistical methodology. Additionally, despite some of 22 items requiring only a single statement to achieve adherence, others required reporting of up to 12 separate sub-items to successfully adhere. While poor reporting for some TRIPOD components may hide possible sources of bias, or make the research question unclear on first reading, others could lead to research

being impossible to replicate. For example, a study with inadequate reporting of the statistical methods and results sections would likely prevent the study from being replicated or the developed model from being externally validated, whereas poor reporting of the title and/or abstract would likely not have such severe consequences.

Secondly, this study reports on the differences in reporting quality between pre-prints and published peer-reviewed articles for a relatively small set, limited to COVID-19 prediction models published at the beginning of the pandemic where time pressures may have played more of a role on reporting than other studies before the pandemic, reducing our ability to generalise our conclusions. The (lack of) detectable improvement is not necessarily attributable to the peer review process. Authors or editors may initiate changes without probes from peer reviewers, and authors or editors may ignore or overrule peer review requests (for example, to adhere to word limits). Furthermore, recent studies have discussed the increasing importance of pre-prints in the dissemination of research during the COVID-19 pandemic, in particular addressing concerns around the lower quality of pre-prints when compared to peer reviewed manuscripts and highlighting the importance of social media platforms and comments sections of pre-print servers in the informal peer review of research [31, 32]. With open peer review available for only 4 of the 19 studies, we were limited in our ability to ascertain whether changes in reporting completeness were a result of formal peer review or a result of journal requirements, editorial processes, or authors' reformatting or editing of their report for submission. Nonetheless, our assessment of the available open peer review reports indicates that peer reviewers generally requested a small number of TRIPOD items in studies with considerable shortcomings in reporting, indicating that peer reviewers might not detect all omissions.

Implications for practice and areas for future research

All prediction model studies included in this review showed incompleteness in their reporting, with only some improvement evident after publication. This suggests the importance of

reporting quality is currently largely overlooked during the publishing process for prediction models.

While it is primarily the author's responsibility to ensure that their prediction modelling study is reported transparently, including all important aspects of model development and/or validation, it is currently unclear whether this responsibility extends beyond the authors onto journal editors and peer reviewers. While many journals do include some form of reporting checklists as a submission requirement, these are often unspecific or even irrelevant to the research questions posed in prediction modelling. Inclusion of relevant reporting guidelines as a submission requirement for journals (and pre-print servers) publishing prediction research could improve this state of affairs considerably. In addition to this inclusion within the submission guidelines for authors, journals should make it explicit whether the checking of completeness of reporting within a manuscript being reviewed is a task for the editors or peer reviewers. If the latter, they should engage individuals with prediction modelling expertise in the peer review process and/or remind invited peer reviewers of prediction model papers of the available reporting checklists such as TRIPOD to check the completeness of reporting within the manuscript being reviewed.

Conclusions

Pre-print and published prediction model studies for COVID-19 are poorly reported, and there is little improvement in the completeness of reporting after peer review and subsequent publication. These findings suggest limited impact of peer review on the completeness of reporting during the COVID-19 pandemic in the field of prediction research. This highlights the need for journals to make adherence to the TRIPOD checklist a submission requirement for prediction modelling studies in this field, and to engage more risk prediction methodology experts in the peer review process.

Declarations

Ethics approval and consent to participate: Not Applicable

Consent for publication: Not Applicable

Availability of data and materials: The datasets generated and analysed during the current study are available from the corresponding author on reasonable request

Conflict of Interests: The authors confirm we have no competing interests

Funding: MTH is supported by St. George's, University of London. LA, RDR, and GSC were supported by funding from the MRC Better Methods Better Research panel (grant reference: MR/V038168/1)**.** GSC was also supported by Cancer Research UK (programme grant: C49297/A27294). The funders had no role in considering the study design or in the collection, analysis, interpretation of data, writing of the report, or decision to submit the article for publication.

Acknowledgements: We would like to thank the peer reviewers and editors for their helpful feedback and comments which helped improve the article.

Author contributions:

Conceptualization - MTH, MvS, GSC, EWS, JBR, RDR, BVC, LW

Data Curation - MTH, LA, CW, LW

Formal Analysis – MTH, LA, LW

Investigation – MTH, LA, LW

Methodology - MTH, LA, MvS, GSC, EWS, JBR, RDR, BVC, LW

Project Administration – MTH

Writing – Original Draft – MTH, LA , LW

Writing – Review&Editing – All Authors

Tables & Figures

Table 1: Descriptive statistics of the reports included within the analyses

	Median (Lower Quartile-Upper quartile) / [Minimum to Maximum]
Pre-print version %adherence	33 (30 - 50) / [10 to 68]
Published version %adherence	42 (31 - 57) / [10 to 71]
Change in %adherence	3 (0 - 7) / [0 to 14]
Journal Impact Factor	4 (4 - 7) / [3 to 40]
Number of days between first submission to the journal and acceptance date	80 (40 - 187) / [22 to 259]
Number of pre-print versions prior to journal submission	1 (1 - 1) / [1 to 3]
Days between manuscript first submission and start of pandemic*	67 (35 - 141) / [12 to 202]
	n (%)
Presence of statement relating to TRIPOD on journal website	
Yes	6 (32)
No*	13 (68)
Pre-print server	
arXiv	2 (11)
medRxiv	16 (84)
bioRxiv	1 (5)

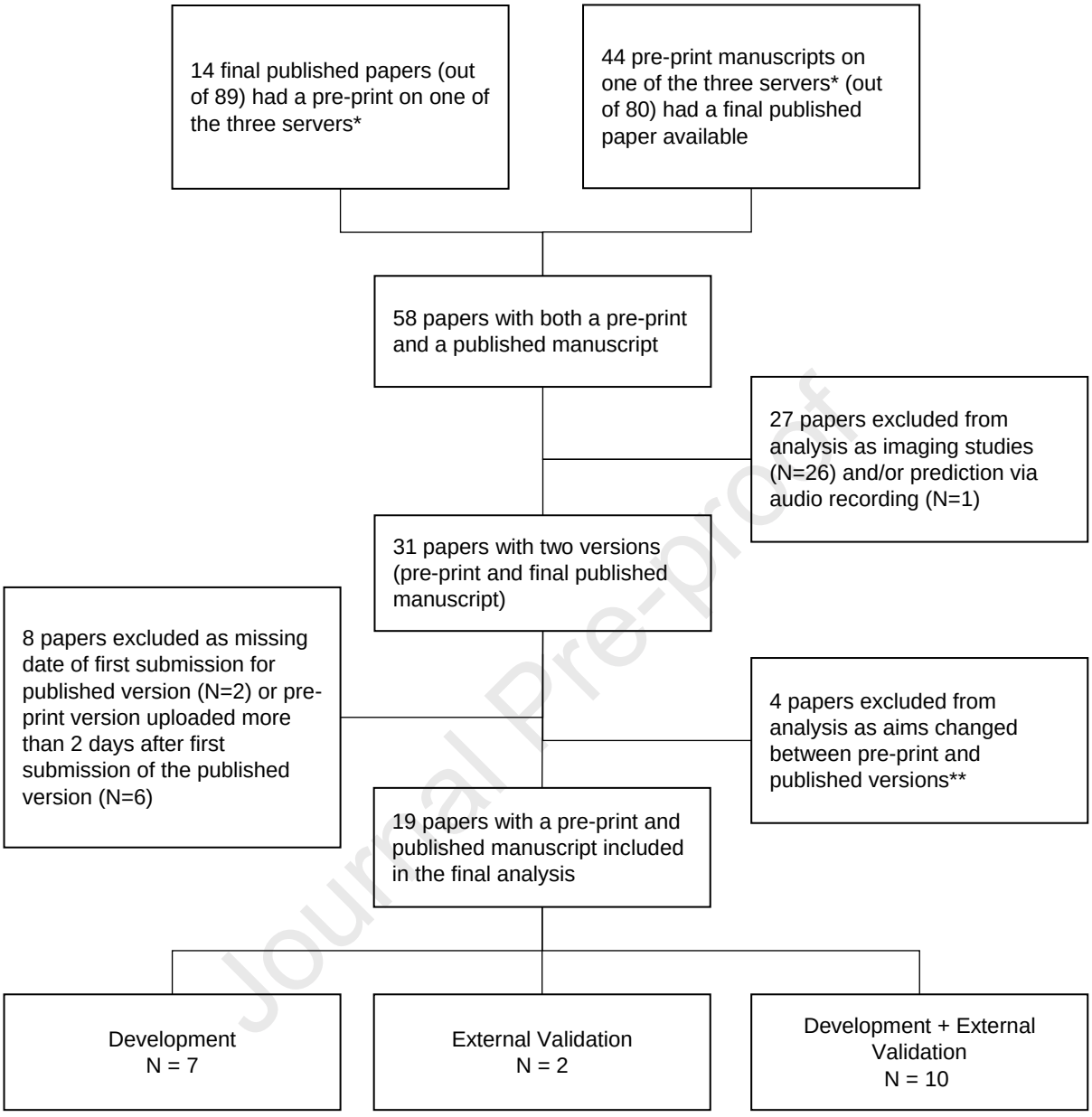
Footnote: *Includes 2 papers published in journals mentioning alternative reporting guidelines (EQUATOR and STARD)

Table 2: Regression coefficients from models assessing associations between published article adherence scores and a range of variables

	Covariate	Coefficient (95% CI)
Model 1	Pre-print score	1.01 (0.87 to 1.16)
	Intercept	0.03 (-0.03 to 0.10)
Model 2	Journal Impact Factor	0.00 (-0.003 to 0.003)
	Pre-print score	1.01 (0.83 to 1.19)
	Intercept	0.04 (-0.03 to 0.11)
Model 3	Time between journal submission and acceptance (per 28-day period)	-0.005 (-0.01 to 0.004)
	Pre-print score	1.01 (0.86 to 1.15)
	Intercept	0.06 (-0.02 to 0.13)
Model 4	Number of pre-prints uploaded prior to journal submission	-0.03 (-0.07 to 0.02)
	Pre-print score	1.03 (0.88 to 1.18)
	Intercept	0.06 (-0.01 to 0.13)
Model 5	Days between manuscript first submission and start of pandemic* (per 28-day period)	-0.008 (-0.02 to 0.003)
	Pre-print score	1.06 (0.90 to 1.21)
	Intercept	0.04 (-0.02 to 0.10)
Model 6	*Statement on the use of the TRIPOD in the journal's instructions to authors*	
	No	(reference group)
	Yes	-0.002 (-0.05 to 0.05)
	Pre-print score	1.01 (0.86 to 1.17)
	Intercept	0.03 (-0.03 to 0.10)

Footnote: *Start of pandemic was defined as 11/03/2020, corresponding to the date the WHO declared COVID-19 as a pandemic

Figure 1: Flowchart of article screening and inclusion/exclusion criteria



FOOTNOTE: *Pre-print servers included were the arXiv, medRxiv, and bioRxiv servers. **Three of the excluded papers whose aims changed between pre-print and published versions were due to the pre-print version concentrating solely on model development whereas the published version included an external validation in addition.

Figure 2: TRIPOD adherence % score for pre-print and published versions of the 19 included studies

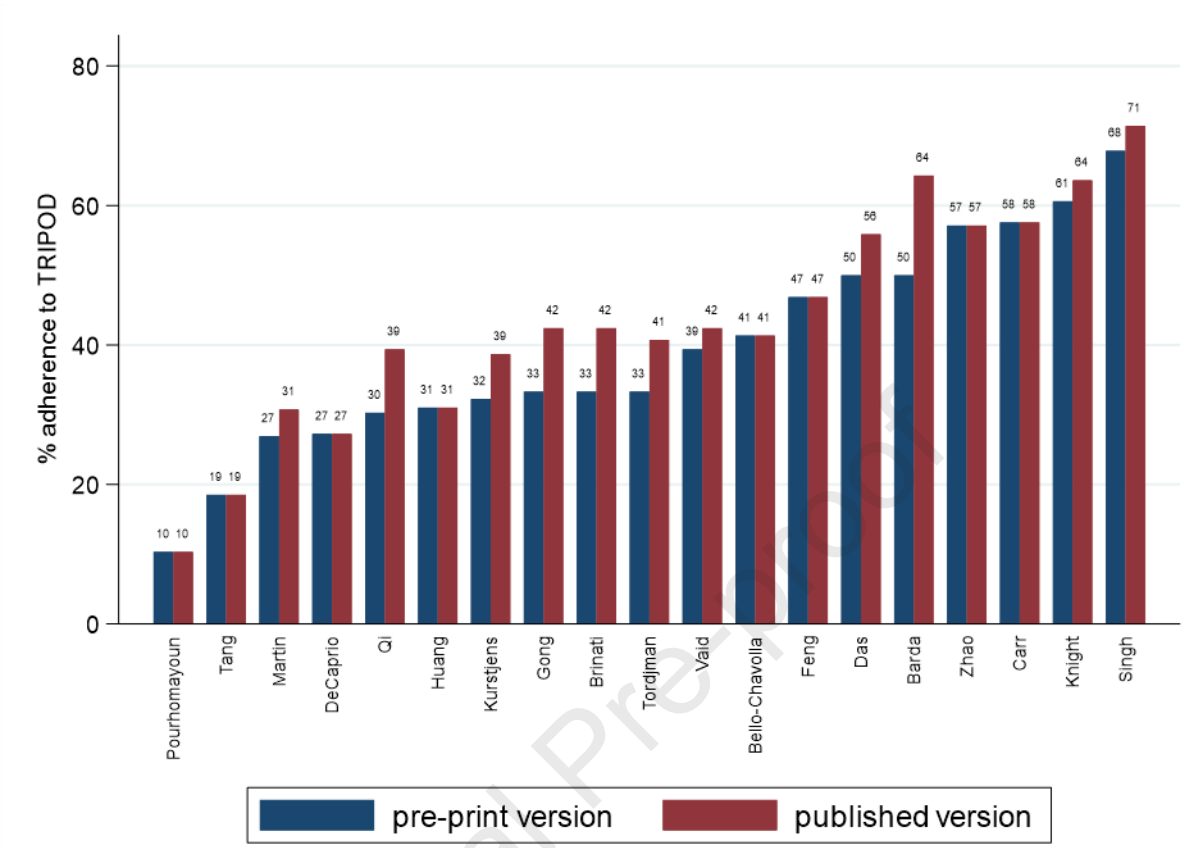


Figure 3: Change in TRIPOD adherence score across publications

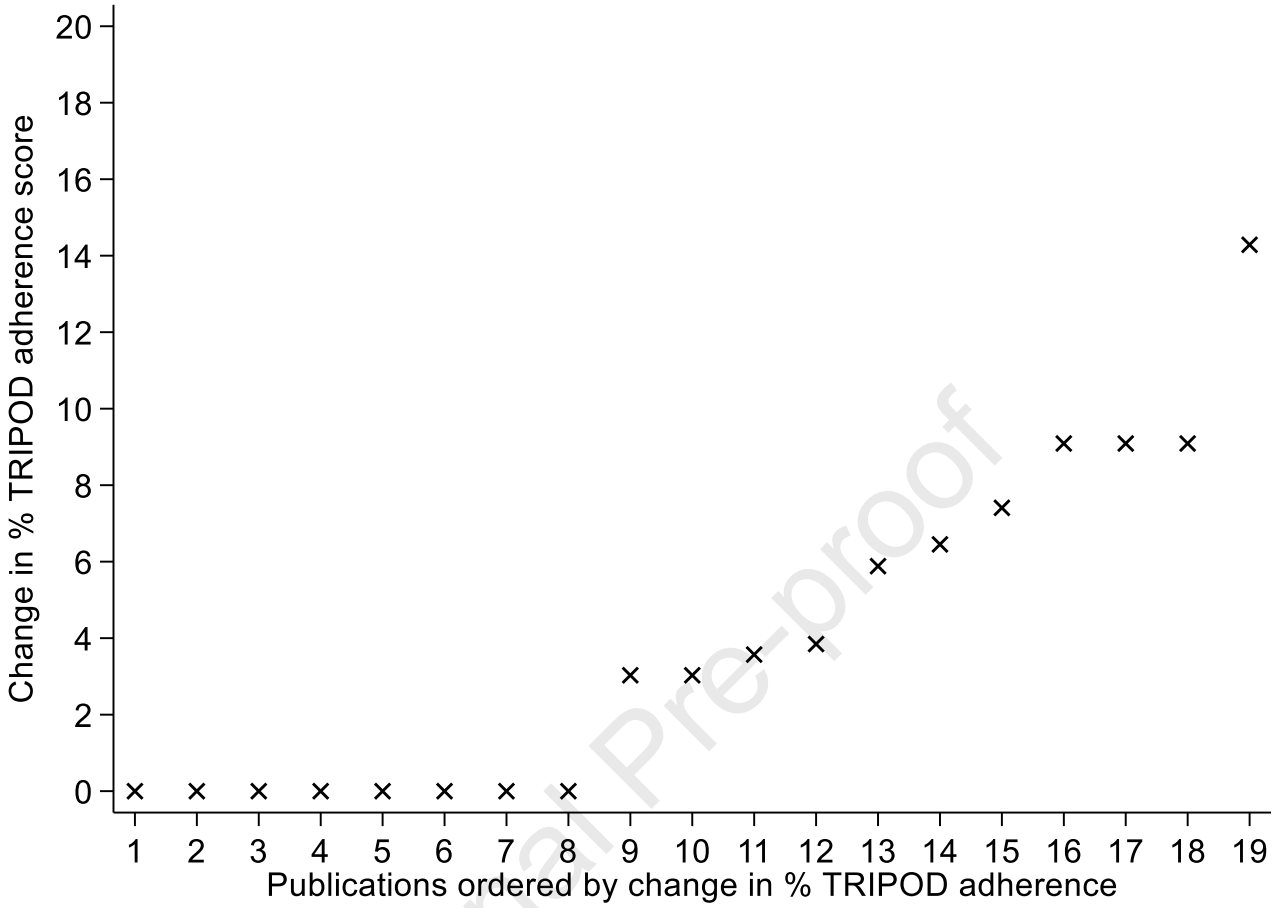
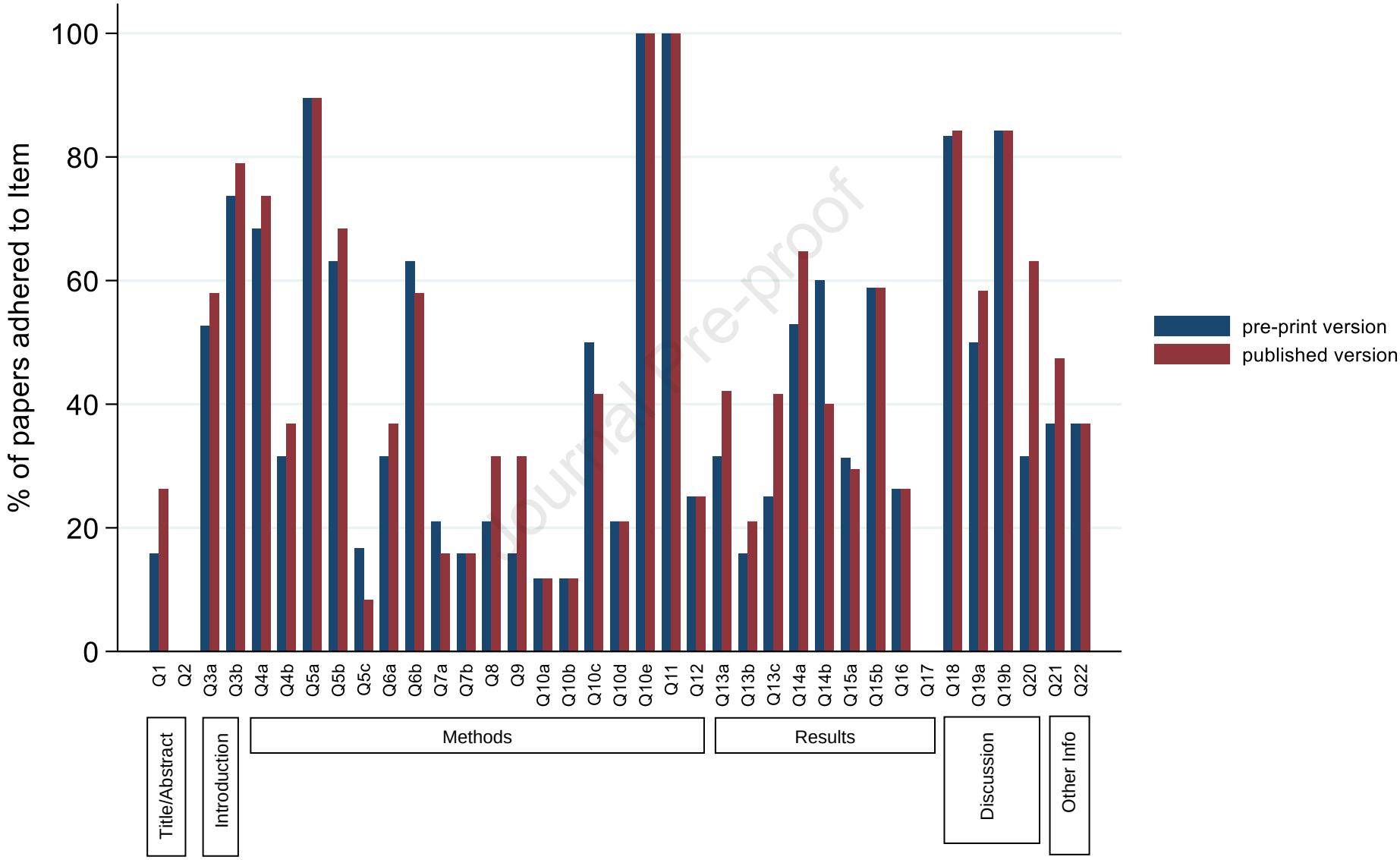


Figure 4: Proportion of papers which adhered to the individual components of adherence in the pre-print and published versions



References

1. **COVID-19 Coronavirus Pandemic** [https://www.worldometers.info/coronavirus/?utm_campaign=homeAdvegas1]
2. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF, Collins GS: **Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration**. *Ann Intern Med* 2015, **162**(1):W1-73.
3. Collins GS, Reitsma JB, Altman DG, Moons KGM: **Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement**. *BMC medicine* 2015, **13**(1):1.
4. Heus P, Damen JA, Pajouheshnia R, Scholten RJ, Reitsma JB, Collins GS, Altman DG, Moons KG, Hooft L: **Poor reporting of multivariable prediction model studies: towards a targeted implementation strategy of the TRIPOD statement**. *BMC medicine* 2018, **16**(1):1-12.
5. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, Bonten MMJ, Dahly DL, Damen JAA, Debray TPA *et al*: **Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal**. *BMJ* 2020, **369**:m1328.
6. Collins GS, Omar O, Shanyinde M, Yu LM: **A systematic review finds prediction models for chronic kidney disease were poorly reported and often developed using inappropriate methods**. *J Clin Epidemiol* 2013, **66**(3):268-277.
7. Glasziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, Michie S, Moher D, Wager E: **Reducing waste from incomplete or unusable reports of biomedical research**. *Lancet* 2014, **383**(9913):267-276.
8. Chalmers I, Glasziou P: **Avoidable waste in the production and reporting of research evidence**. *Lancet* 2009, **374**(9683):86-89.
9. Glasziou PP, Sanders S, Hoffmann T: **Waste in covid-19 research**. *Bmj* 2020, **369**:m1847.
10. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE *et al*: **The PRISMA 2020 statement: An updated guideline for reporting systematic reviews**. *J Clin Epidemiol* 2021, **134**:178-189.
11. Heus P, Damen JAAG, Pajouheshnia R, Scholten RJPM, Reitsma JB, Collins GS, Altman DG, Moons KGM, Hooft L: **Uniformity in measuring adherence to reporting guidelines: the example of TRIPOD for assessing completeness of reporting of prediction model studies**. *BMJ Open* 2019, **9**(4):e025611.
12. **Enhancing the QUAlity and Transparency Of health Research** [https://www.equator-network.org/]
13. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, Lijmer JG, Moher D, Rennie D, de Vet HC *et al*: **STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies**. *BMJ* 2015, **351**:h5527.
14. Das AK, Mishra S, Saraswathy Gopalan S: **Predicting CoVID-19 community mortality risk using machine learning and development of an online prognostic tool**. *PeerJ* 2020, **8**:e10083.

15. Carr E, Bendayan R, Bean D, Stammers M, Wang W, Zhang H, Searle T, Kraljevic Z, Shek A, Phan HTT *et al*: **Evaluation and improvement of the National Early Warning Score (NEWS2) for COVID-19: a multi-hospital study**. *BMC medicine* 2021, **19**(1):23.
16. DeCaprio D, Gartner J, McCall CJ, Burgess T, Garcia K, Kothari S, Sayed S: **Building a COVID-19 vulnerability index**. *Journal of Medical Artificial Intelligence* 2020, **3**.
17. Knight SR, Ho A, Pius R, Buchan I, Carson G, Drake TM, Dunning J, Fairfield CJ, Gamble C, Green CA *et al*: **Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score**. *BMJ* 2020, **370**:m3339.
18. Barda N, Riesel D, Akriv A, Levy J, Finkel U, Yona G, Greenfeld D, Sheiba S, Somer J, Bachmat E *et al*: **Developing a COVID-19 mortality risk prediction model when individual-level data are not available**. *Nature communications* 2020, **11**(1):4439.
19. Collins GS, Mallett S, Omar O, Yu LM: **Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting**. *BMC medicine* 2011, **9**:103.
20. Bouwmeester W, Zuithoff NP, Mallett S, Geerlings MI, Vergouwe Y, Steyerberg EW, Altman DG, Moons KG: **Reporting and methods in clinical prediction research: a systematic review**. *PLoS Med* 2012, **9**(5):1-12.
21. Damen JA, Hooft L, Schuit E, Debray TP, Collins GS, Tzoulaki I, Lassale CM, Siontis GC, Chiocchia V, Roberts C *et al*: **Prediction models for cardiovascular disease risk in the general population: systematic review**. *Bmj* 2016, **353**:i2416.
22. Wen Z, Guo Y, Xu B, Xiao K, Peng T, Peng M: **Developing Risk Prediction Models for Postoperative Pancreatic Fistula: a Systematic Review of Methodology and Reporting Quality**. *The Indian journal of surgery* 2016, **78**(2):136-143.
23. Zamanipoor Najafabadi AH, Ramspek CL, Dekker FW, Heus P, Hooft L, Moons KGM, Peul WC, Collins GS, Steyerberg EW, van Diepen M: **TRIPOD statement: a preliminary pre-post analysis of reporting and methods of prediction models**. *BMJ Open* 2020, **10**(9):e041537.
24. Jiang MY, Dragnev NC, Wong SL: **Evaluating the quality of reporting of melanoma prediction models**. *Surgery* 2020, **168**(1):173-177.
25. Collins GS, de Groot JA, Dutton S, Omar O, Shanyinde M, Tajar A, Voysey M, Wharton R, Yu LM, Moons KG *et al*: **External validation of multivariable prediction models: a systematic review of methodological conduct and reporting**. *BMC Med Res Methodol* 2014, **14**:40.
26. Carneiro CFD, Queiroz VGS, Moulin TC, Carvalho CAM, Haas CB, Rayêe D, Henshall DE, De-Souza EA, Amorim FE, Boos FZ *et al*: **Comparing quality of reporting between preprints and peer-reviewed articles in the biomedical literature**. *Research Integrity and Peer Review* 2020, **5**(1):16.
27. Pierie J-PEN, Walvoort HC, Overbeke AJPM: **Readers' evaluation of effect of peer review and editing on quality of articles in the Nederlands Tijdschrift voor Geneeskunde**. *The Lancet* 1996, **348**(9040):1480-1483.
28. Goodman S, Berlin J, Fletcher S, Fletcher R: **Manuscript Quality before and after Peer Review and Editing at Annals of Internal Medicine**. *Annals of Internal Medicine* 1994, **121**(1):11-21.

29. Cobo E, Selva-O'Callagham A, Ribera J-M, Cardellach F, Dominguez R, Vilardell M: **Statistical Reviewers Improve Reporting in Biomedical Articles: A Randomized Trial**. *PLOS ONE* 2007, **2**(3):e332.
30. Cobo E, Cortés J, Ribera JM, Cardellach F, Selva-O'Callaghan A, Kostov B, García L, Cirugeda L, Altman DG, González JA *et al*: **Effect of using reporting guidelines during peer review on quality of final manuscripts submitted to a biomedical journal: masked randomised trial**. *BMJ* 2011, **343**:d6783.
31. Fraser N, Brierley L, Dey G, Polka JK, Pálfy M, Nanni F, Coates JA: **The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape**. *PLoS biology* 2021, **19**(4):e3000959.
32. Brierley L, Nanni F, Polka JK, Dey G, Pálfy M, Fraser N, Coates JA: **Tracking changes between preprint posting and journal publication during a pandemic**. *PLoS biology* 2022, **20**(2):e3001285.

What's New?

In this study, we compared adherence to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement for studies developing or validating COVID-19 prediction models with a pre-print version available, before and after the peer review process

The findings of this report demonstrate a poor quality of reporting of COVID-19 prediction modelling studies amongst pre-print versions, which did not improve much following peer review.

Most TRIPOD items saw no change in the frequency of their reporting, with only the coverage of discussion items being substantially improved in the published version.

While it has previously been reported that the adherence to reporting guidelines for prediction modelling studies was poor, our findings also suggest that the peer review process had little impact on this adherence during the pandemic.

The implication of these findings is that a greater focus is needed on the importance of adhering to reporting guidelines by authors as well as checking during the editorial process.

Author contributions:

Conceptualization - MTH, MvS, GSC, EWS, JBR, RDR, BVC, LW

Data Curation - MTH, LA, CW, LW

Formal Analysis – MTH, LA, LW

Investigation – MTH, LA, LW

Methodology - MTH, LA, MvS, GSC, EWS, JBR, RDR, BVC, LW

Project Administration – MTH

Writing – Original Draft – MTH, LA , LW

Writing – Review&Editing – All Authors