**Methodological and reporting quality of COVID-19 and other research – focussed literature review**

Protocol v1.0 13.05.20

Protocol v2.0 31.05.20

Current Protocol V2.3 02.06.20

**Summary**

**Objective:** To collate and assess published research in four major clinical journals, comparing the methodological and reporting quality of coronavirus disease 2019 (covid-19) research papers and non-COVID-19 research papers.

**Design:** Focussed literature review and critical appraisal.

**Data Sources:** Research publications (print and online) over three months from: The BMJ, Journal of the American Medical Association, The Lancet, New England Journal of Medicine.

**Study selection:** All clinical research from first publication of a covid-19 clinical research paper to May 2020.

**Data Extraction:** Paired authors will independently extract data with a focus on methodological quality (risk of bias) and reporting quality. The following tools will be used:Cochrane Risk of Bias version one (ROB1), Consolidated Standards of Reporting Trials (CONSORT) for randomised controlled trials; National Heart Lung and Blood Institute quality assessment, Strengthening the Reporting of Observational studies in Epidemiology (STROBE) for observational research; Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2), Standards for Reporting Diagnostic Accuracy (STARD 2015); A Measurement Tool to Assess systematic Reviews version 2 (AMSTAR 2), Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). For methodological quality an overall assessment of high or low risk of bias will be made, for reporting quality the sum of reported items will be used. We will perform quantitative comparison of COVID-19 and non-COVID-19 research reporting and quality adjusted for data of publication, journal, publication format and primary research method.

**Results:**

**Conclusion:**

**Review question**

In high profile biomedical journals, is there a difference in methodological or reporting quality between published COVID-19 research and published research pertaining to other topics.

**Team**

Suzanne C Freeman – Department of Health Sciences, University of Leicester

0000-0001-8045-4405

Clareece R Nevill – Department of Health Sciences, University of Leicester

0000-0001-8305-2516

Kris McGill – NMAHP Research Unit, Glasgow Caledonian University

0000-0002-0307-1440

Kerry Dwan – Cochrane Methods Support Unit

0000-0001-6918-1215

Jennifer K Burton - Institute of Cardiovascular and Medical Sciences, University of Glasgow

0000-0002-4752-6988

Martin Taylor Rowan – Institute of Cardiovascular and Medical Sciences, University of Glasgow

Terence J Quinn - Institute of Cardiovascular and Medical Sciences, University of Glasgow

0000-0003-1401-0181

Ryan Field – Health Economics and Health Technology Assessment, University of Glasgow

0000-0002-4424-9890

Ping-Hsuan Hsieh - Health Economics and Health Technology Assessment, University of Glasgow; Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan.

0000- 0002- 0430- 260X

Claudia Geue – Health Economics and Health Technology Assessment, University of Glasgow

0000-0003-2243-0733

Dikshyanta Rana - Health Economics and Health Technology Assessment, University of Glasgow

0000-0001-9133-3094.

Yiqiao Xin – Health Economics and Health Technology Assessment, University of Glasgow

0000-0001-5856-3103

**Statistician:** Ben Carter – Institute of Psychiatry, Psychology and Neuroscience Kings College London

0000-0003-0318-8865

**Rationale**

The severe acute respiratory syndrome coronavirus 2 (SARS COV-2), and the resulting clinical manifestations of coronavirus disease (COVID-19) have disrupted all aspects of healthcare. Biomedical research has not been spared. In the context of a novel virus, and a global pandemic of a severity not seen for generations, there is an urgent need for clinical research to inform practice and policy. The resulting, concerted global effort to conduct COVID-19 research has been impressive in scope and in speed. As the pathogen is new, research is needed in various scientific spheres including observational epidemiology, assessment of the accuracy of test strategies, trials of interventions and others. All of this research activity is incredibly time sensitive as healthcare systems need guidance, tests and treatments as soon as possible.

The main platform for sharing results of research remains the peer reviewed, biomedical journal. Although the viral pandemic has brought many challenges for academic publishing, journals have risen to the challenge creating processes to allow for rapid sharing of relevant research. Just as the research community has had to work quickly to produce research, so publishers and journal teams have had to adapt to a new urgency and demand. Many biomedical publishers have reported a substantial increase in submissions in 2020, with most of the content related to COVID-19. Journals have responded by reducing time from submission to acceptance. However, there has been concern that in the rush to share clinical knowledge and data, some good practice aspects of research design, conduct and interpretation may not have been adhered to.

Arguments around the integrity and quality of COVID-19 research have been rehearsed in the lay and scientific press. While these important issues have generated substantial copy, there has been little scientific description of quality as a basis for this discourse. It seems paradoxical to criticise the quality of published science based on a non-scientific selection of this literature. What is needed is a systematic, comprehensive and objective assessment of the high profile, high visibility published research literature.

There are many aspects of a scientific paper that contribute towards the overall ‘quality’ of the published research. Two complementary factors that are often assessed during critical appraisal are the design and conduct of the research (methodological quality) and the way the study and results are communicated (reporting quality). Both are important. Methodological quality is important to ensure that the results presented are robust and free from bias, while reporting quality is important to ensure transparency and aid interpretation.

We wish to assess methodological quality (risk of bias) and reporting quality for COVID-19 research published in the highest impact biomedical journals and compare with contemporaneous research on non COVID-19 topics. Our hypothesis is that even though COVID-19 research is being produced at an unprecedented volume and speed, methodological and reporting quality in the highest profile journals will be maintained.

**Search strategy**

We will include the following journals, chosen as representing the highest impact clinical research titles in the category Medicine, General and Internal (based on Journal citation reports 2018 category (Clarivate Analytics)): The BMJ (British Medical Association); The Journal of the American Medical Association (JAMA – American Medical Association) The Lancet (Elsevier); The New England Journal of Medicine (NEJM - Massachusetts Medical Society). We favoured these four titles as they publish weekly print editions, represent international publishing, editorial bases and readerships and are generally considered to have the highest standards of publication.

We will assess all articles labelled as original clinical research by the parent journal, including systematic reviews, case reports, brief reports. We will not include editorials, comment, non-systematic literature review or pre-clinical science articles. Each print journal will be hand searched by two authors to ensure no relevant papers are missed. As an internal validity check, the journals’ online search facility and/or any specific COVID-19 resource hubs will be checked to ensure no relevant print or electronic content has been missed during the search period.

Sampling will begin in the first week that the relevant journal publishes a COVID-19 related clinical research paper (late February for most journals) and will conclude on the second week of May inclusive. On a single date in mid-May (17th May 2020), the journal websites will be searched for all papers accepted for publication and available online prior to print. These dates have been chosen to allow three months of original research content to be assessed.

**Papers to be included**

Inclusion will be based on classification as original research piece by the parent journal. We will use the most recent version of the paper available, but note whether the paper had to be updated due to an issue in the original publication.

Within the original research remit, we will exclude pre-clinical science papers, as the tools we have chosen for assessment are not suited to this scientific method (indeed for many pre-clinical science approaches there may not be equivalent methodology and reporting guidelines). Other potential papers that may not suit our predefined approach include but are not limited to papers with a purely methodological focus; papers with a focus on psychometrics and test development (rather than a specific focus on test accuracy); genetic profiling studies and data modelling studies. Given the broad clinical readership of the chosen journals, we do not anticipate that many papers will be excluded on this basis.

If an article is otherwise eligible but is based on a research method that is not aligned with our suite of quality and reporting tools (see below) the article will not undergo quality assessment but will still be included in any aggregate descriptions of journal content. If there is more than one such paper, the team will decide whether to try and include in this review and find an appropriate quality and reporting tool.

We will classify included papers based on the study method using the system described below, this classification will inform the choice of quality assessment tools that will be used.

We have pre-defined six categories that should encompass most clinical research.

* (Randomised) Controlled Trials (RCT) – where the aim is to assesses the effect of an intervention made by the research team against a comparator. Where a specific trial design is used, such as cluster methods, then this will be noted. We will include controlled trials with no clear randomisation. Pseudo-RCTs, where there has been no intervention by the research team, but rather existing data are used to make inferences about an intervention and comparator will be treated as observational.
* Diagnostic test accuracy (DTA) – where the aim is to assess the properties of a test, or measurement or clinical classification scheme.
* Observational studies (Observational) – further subdivided into case-study/case-series (authors own description), case-control, cohort and cross sectional, where the aim is to use existing data to look for patterns or associations, with no intervention by the research team. Within this rubric we will include case studies and case series.
* Qualitative (Qualitative) - where the aim is to interpret unstructured, non-numerical data to create meaning.
* Prognosis (Prognosis) – further subdivided into fundamental prognosis, prognostic factor, prognostic model or rule, where the aim is to assess for association between baseline states and future outcomes.
* Systematic review (SR) – where the aim is to use existing data from all relevant sources, usually aggregate published data in biomedical journals, to answer a research question. Other forms of meta-research will be included in this rubric.

Papers that do not fit any of these categories will be given a label of ‘other’.

**Definition of a COVID-19 study**

Within our dataset of published papers, we will create two groups to facilitate comparisons, ‘COVID-19 research’ and ‘other’.

The COVID-19 label will be used where the exposure or intervention or test relates to COVID-19 / SARS COV-2 or the outcomes relate to COVID-19 / SARS COV-2. We recognise that other research areas may also be prioritised by journals as they are relevant to the COVID-19 pandemic, for example studies of other viruses. However, we will only include research directly related to SARS COV2 and its clinical consequences within our COVID-19 group.

**Outcomes of interest**

Our primary outcome of interest is the ‘quality’ of the published paper. We will assess methodological quality and assign papers a status of ‘low risk of bias’ using appropriate tools. We will also assess reporting quality and describe compliance with reporting guidelines.

For our quality assessment we will use validated tools specific to our pre-specified study designs. We have chosen a tool for each of the study methods described above. There is no consensus on the optimal tool for quality or reporting assessment. Choice of tools for this study was based on validation, availability of training materials and guidance and familiarity within the team. We have favoured those tools used by Cochrane where possible but note that Cochrane do not have a preferred tool for non-interventional observational studies. We will use the NHLBI tool for observational research, as it meets our criteria and has the flexibility to allow scoring across various observational methods (cross-sectional, case-control and cohort).

For quality assessment, we have chosen a series of validated assessment checklists that assess risk of bias or internal validity. Each checklist assesses various aspects of the study and can be graded as ‘low’ or ‘high’ risk of bias or not applicable. An option of ‘uncertain’ or ‘unclear’ risk of bias is allowed, but in comparative quantitative analyses, uncertain will be recoded as high risk. For each assessment the option of recording a concern that is not captured under the existing headings will be available (other). An example, of ‘other’ risk of bias could be non-availability of a protocol or statistical analysis plan (SAP). Risk of bias will be assessed at the level of pre-specified domains and also at the level of the complete paper. Scoring of risk of bias will take account of all relevant materials and should (where available) include assessment of protocols, trial registration details, supplementary materials etc.

For reporting quality, we will use checklists chosen from the EQUATOR (Enhancing the quality and transparency of health research) resource. We will favour the tools recommended by the included journals (where the journal states a preference). Reporting will be scored as ‘reported’ (where reporting was deemed adequate), ‘not reported’ or ‘not applicable’. Scoring of reporting will be described by checklist item. Where an item has more than one question, for example a part A and part B, all of these need to be reported for the parent item to be recorded as adequately reported. Reporting will be based on what is presented in the paper or is signposted and easily available to the reader, for example online supplementary materials.

|  |  |  |  |
| --- | --- | --- | --- |
| **Design** | **Quality****(risk of bias)** | **Domains assessed** | **Reporting** |
| Controlled trial | Cochrane RoB | randomisation, allocation, blinding (participants), blinding (outcomes), incomplete outcomes, selective reporting, other  | CONSORT  |
| Observational  | NHLBI | question, population, exposure, outcomes, confounding, other  | STROBE |
| Test accuracy  | QUADAS2 | patient selection, index test, reference standard, flow and timing, generalisability  | STARD |
| Systematic review | AMSTAR2 | design & protocol, search strategy, paired extraction, inclusion/exclusion, risk of bias, meta-analysis, conflicts of interest  | PRISMA |
| Qualitative | CASP | design, recruitment, data collection, relationships, analysis  | COREQ |
| Prognosis  | PROBAST\* | participants, predictors, outcomes, analysis, generalisability | TRIPOD |

**RoB**=Risk of bias (version one); (Higgins Julian P T, Altman Douglas G, Gøtzsche Peter C, Jüni Peter, Moher David, Oxman Andrew D et al. The Cochrane Collaboration’s tool for assessing risk of bias in randomised trials BMJ 2011; 343 :d5928)

**CONSORT**=Consolidated Standards of Reporting Trials (Schulz Kenneth F, Altman Douglas G, Moher David. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials BMJ 2010; 340 :c332)

**NHLBI**=National Heart Lung and Blood Institute (https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools)

**STROBE**=Strengthening the Reporting of Observational studies in Epidemiology (von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)statement: guidelines for reporting observational studies. BMJ. 2007 Oct 20;335(7624):806-8)

**QUADAS** (version 2)=Quality Assessment of Diagnostic Accuracy Studies. (Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JAC, and Bossuyt PMM. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. Ann Intern Med 155 (8):529-536, 2011.)

**STARD**=Standards for Reporting Diagnostic Accuracy (Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ. 2015;351:h5527. Published 2015 Oct 28. doi:10.1136/bmj.h5527)

**AMSTAR**=A Measurement Tool to Assess systematic Reviews (Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both BMJ 2017; 358 :j4008)

**PRISMA**=Preferred Reporting Items for Systematic Reviews and Meta-Analyses (Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA **statement. BMJ. 2009;339:b2535.**

**CASP**=Critical Appraisal Skills Program (https://casp-uk.net/casp-tools-checklists/)

**COREQ**= Consolidated criteria for reporting qualitative research (Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care. 2007;19(6):349-357.)

**PROBAST**=Prediction model Risk Of Bias Assessment Tool.Wolff RF, Moons KGM, Riley RD, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Ann Intern Med. 2019;170(1):51‐58. doi:10.7326/M18-1376

**TRIPOD**=Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ. 2015;350:g7594. Published 2015 Jan 7. doi:10.1136/bmj.g7594)

\* depending on design of prognostic study, the Quality in Prognosis Studies (QUIPS) tool may be used

**Training and calibration**

The team were selected from volunteers, from various academic centres, with interest and experience in research methodology or meta-research. To ensure consistency in approach, we will hold a (virtual) kick-off meeting and then weekly team meetings until all papers are assessed and analysis complete. Between meetings, discussion through email and ad hoc calls will be encouraged. Written training materials for each of the tools to be used will be shared with all team members.

Within the team we have six researchers who self-identify as highly experienced in assessing methodological and reporting quality. We will pair each of these researchers with a researcher who is less experienced in use of the chosen tools. Allocating of pairs will be random using the online random.org resource.

Assessor pairs will be randomly allocated a test set of six papers for review that will be discussed within their pairs and if necessary, with the complete team, discrepancies in scoring resolved before commencing the main review.

If an allocated paper uses a method with which the reviewer team are not familiar, it will be reallocated.

We will collect data on item level disagreement requiring discussion for each assessor pair and present this as percentage agreement.

**Data Extraction**

Two people will independently extract the following study level data to describe included studies:

Journal name, date of publication, study design (as defined above), whether the paper was identified as a ‘brief report’ or equivalent, the exposure of interest (or intervention, or index test), the primary outcome, the topic area of the paper, the total ‘n’ included in the study at baseline (for a systematic review this was taken as the number of included papers), the length of follow-up (time from first measure to last measure for primary outcome, quantified in weeks) and the funding source (academic, industry, mixed). We will also assess for whether the paper had an accompanying editorial, or a printed retraction, serious correction.

We will also determine whether the study is framed as positive (hypothesis is proven) or neutral. This will be based on the interpretation of the authors of the original paper as given in the abstract and main text conclusions. For studies that have no underlying hypothesis, a category of not applicable will be used.

Reviewers will use standardised data extraction forms that have been piloted on two studies (one RCT, one observational). Data will be collected by reviewer pairs and then shared with a coordinating researcher who will keep a master file.

We have created a data dictionary that operationalises all of the variables that we will collect at included paper level.

**Process of data extraction and outcomes assessment**

A long list of eligible papers will be created. Assessor pairs will be randomly allocated papers to assess, randomisation will use the random.org platform. Each team will receive weekly blocks of ten papers to review.

Each member of the team will assess methodological and reporting quality independently and then compare scores and reach consensus. Where there is no consensus, papers will be discussed with a third reviewer independent of the original pair. The final consensus scores will be submitted and collated.

As a further validity check, a random selection of 10% of the included papers will be selected for second review by a third reviewer independent of the original pair with any discrepancies discussed with the original scoring team. Again, random.org will inform the randomisation process.

**Strategy for data synthesis**

In our reporting we will adhere to the relevant sections of the PRISMA preferred reporting items for systematic reviews and meta-analyses. We will create a PRISMA flow chart to illustrate the included studies and relative contribution from each of the target journals.

If there are common issues around risk of bias and reporting these will be highlighted in a textbox and considered in the Discussion.

We will create graphical illustrations of methodological and reporting quality using the traditional traffic light system. These will be presented at paper level and in aggregate, for each of the study methods if they have more than ten relevant papers. The purpose of this project is not to single out individual journals and so no analyses will be presented at the level of the parent journal.

We will tabulate descriptive statistics for COVID-19 and non COVID-19 research variables.

To maintain transparency in all aspects of the process we will make the data collection sheets available and will have processes to allow us to share all relevant material with any interested party on reasonable request.

**Strategy for quantitative analysis**

Our primary outcome of interest is the ‘quality’ of the published paper. We will assess methodological quality and assign papers a status of ‘low risk of bias’ using appropriate tools. We will also assess reporting quality and describe compliance with reporting guidelines.

There is no consensus on how to quantitatively assess scores on assessments of methodological or reporting quality. Although we will describe domain level assessments that could be summed to give an overall quantitative quality score, creating aggregate quality scores in this way has limitations and is often discouraged in evidence-based medicine texts. In particular, the domain level summative score approach may under-estimate bias, for example a paper may have a single fundamental flaw that scores only 1 point but that completely undermines the validity of the research. Thus, we will also ask assessors to give an overall assessment of the methodological quality of the paper, using a binary high risk/low risk score. These data will be compared with COVID-19 and other research using a chi-square approach and then described as proportional agreement with corresponding confidence intervals.

For the primary analysis, we will compare proportions of ‘low risk of bias’ in COVID-19 and non-COVID-19 across all included papers. We will then run the same analyses for each research method (controlled trials, observational, prognosis, qualitative, test accuracy) if they have more than 10 included papers.

We will create multivariable models with our binary quality scores as response variable and predictors that include: COVID-19 research; journal; full or brief report style; date of publication; study design; funding; assessor pairs performing quality assessment.

We will perform similar assessments for reporting quality assessments but will limit our analyses to an ordinal scoring approach, summing each reported item. All reporting items carry similar weight and so a summative approach is appropriate in this context. We will describe total scores and also proportion of items adequately reported. Where some items are scored as ‘not applicable’ the denominator will be adjusted accordingly.

**Secondary analyses:** We will explore the effect of timing of publication, as plausibly those papers published in the early phase of the pandemic may differ from more recent papers. We will use a cut-off of second week in April (10 weeks after first paper) for these analyses.

We will discourage use of the uncertain category in primary rating. Where risk of bias is not certain in the paper, in keeping with best practice, reviewer pairs will be asked to make an informed decision about the possibility of bias. If there remain overall assessments graded as unclear they will not be included in the primary analysis but as a sensitivity analysis, we will look at the effect on quality scores if unclear assessments are changed to low and high risk of bias.

If time allows, we will use modified star plots to facilitate data visualisation across the various domains of quality, comparing COVID-19 and other research.

We will lock our database when the last review is submitted and quality control checks are complete. The database will be shared with a statistician with expertise in evidence synthesis, but independent of the main review group. COVID status will be coded so that this variable is not obvious to the statistician.

**Points to consider in interpretation**

**Context**

The UK government response to COVID-19 has been based on a mantra of ‘following the science’ (more recently revised to consider or listen to the science). This approach is laudable but assumes that the science produced at speed and at scale in a global pandemic is infallible. Is this a reasonable assumption?

Should we expect the chosen journals to offer bias free research with fully transparent reporting? There are certain standards that we should expect, for example most high impact, contemporary biomedical journals are signatories of the Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals issued by the International Committee of Medical Journal Editors (ICMJE) and to the Committee on Publication Ethics (COPE) code of conduct. However, these recommendations are designed to prevent research misconduct rather than ensure a certain methodological approach.

Most journals also mandate that the scientific writing in papers follows the guidelines for health research reporting that are collated by the Enhancing the Quality and Transparency of Health Research (EQUATOR) Network. All the journals assessed in this analysis adhere to CONSORT recommendations, although take up of other reporting guidance is less consistent.

The story of hydroxychloroquine treatment in COVID-19 gives an example of both the need for rapid communication of science and the danger of not recognising inherent biases. Based on observational studies and a modest sized RCT the drug achieved substantial visibility and entered clinical practice. Further observational studies suggested the benefits may have been overstated and large RCTs are now suggested that the drug is harmful.

**Explanations**

If there are differences in methodology and reporting quality, what may drive this. It has been said that the COVID-19 crises created new incentives for publishing, including a race to be the first to document descriptions of virus presentation, investigation, management and outcomes. In a time-sensitive publishing space, quality is likely to suffer. The UK reproducibility network have commented that with COVID-19 research you can only have two of ‘fast, cheap, good’.

**What is ‘quality’**

For our assessment of methodological and reporting quality we used validated tools, chosen according to a defined process. There is a tension between having a standard quality assessment tool that allows data to be compared and pooled and having a quality assessment tool that is specific to the research under scrutiny. While our chosen tools were appropriate for all the papers assessed, within our broad categorisation of method there are sub-methods – for example, within our systematic review category we used the generic AMSTAR and PRISMA tools, but we recognise that there are specific tools for systematic review with network meta-analysis, systematic review of test accuracy data etc

Quality of science is more than valid methods and transparent reporting. Indeed, other important factors like inclusiveness and clinical urgency may seem to conflict with rigid rules on method and reporting. Even within our chosen remit of methodological quality there are many facets. For example, some of our tools have a strict focus on internal validity (bias) and others also consider external validity (generalisability). Accepting all this, we do not claim to have described a definitive measure of quality, but we have described two fundamental aspects of quality that we feel should not be sacrificed even in the midst of a viral pandemic. Finally, our quality assessment can only work with what is published and even the best quality assessment tools will not pick up those cases where there has been research fraud.

**Strengths and Limitations of included papers**

Suspect there will be a broad range of research.

Common limitations of COVID-19 research may emerge eg around case ascertainment and definitions, sufficient timing to develop outcomes, not following best practice eg pre-registering protocol or SAP

**Strengths & Limitations of our approach**

Managed team approach allowed for rapid, robust assessment, collation and analyses.

Multidisciplinary team that includes expertise in clinical research; meta-research; critical appraisal and statistics.

Multiple internal validation checks and steps to ensure consistency in assessment.

Only included four journals, not representative of majority of published research and certainly not representative of research published in pre-print repositories.

Variety of RoB and reporting tools are available with no consensus on the optimal tool. Our choice of tool was based on explicit criteria and agreed by the group, but we recognise that some authors prefer other tools and all tools have inherent strengths and limitations.

RoB versus other aspects of quality.

**Implications for research and practice**

**Contact details for further information**

Dr Terry Quinn

Institute of Cardiovascular and Medical Sciences

University of Glasgow

New Lister Building Campus, Glasgow Royal Infirmary

Glasgow, UK, G31 2ER

T: +44 (0)141 201 8510

E: terry.quinn@glasgow.ac.uk

@: @DrTerryQuinn

**Anticipated start date:**

01.05.20

(kick-off meeting 08.05.20)

**Anticipated completion date:**

June 2020

**Dissemination strategy:**

We think this analysis will be of interest to various stakeholders.

We will submit to a high impact clinical journal with a broad readership. We will also disseminate results through social media, using personal accounts and those of our organisations. Depending on results we may also involve University media teams and create a press release.

**Conflicts of interest:**

Suzanne C Freeman – Nil

Clareece R Nevill – Nil

Kris McGill – Nil

Kerry Dwan – Nil

Jennifer K Burton - Nil

Martin Taylor Rowan – Nil

Terence J Quinn - Nil

Ryan Field – Nil

Ping-Hsuan Hsieh - Nil

Claudia Geue – Nil

Dikshyanta Rana - Nil

Yiqiao Xin – Nil

Ben Carter - nil

**Funding source/sponsors**

SCF, CRN, TJQ, YX are supported by the NIHR Complex Reviews Support Unit (project number 14/178/29).. This work is independent of NIHR.

KD, TJQ are supported by Cochrane, this work is independent of Cochrane

**Contribution**

All authors will contribute to data extraction, analysis and interpretation. All authors will provide critical comment on draft manuscripts.

**References of interest**

D Kwon How swamped preprint servers are blocking bad corona virus research (news) Nature 581, 130-131 (2020) doi: 10.1038/d41586-020-01394-6

Serge P.J.M. Horbach. Pandemic Publishing: Medical journals drastically speed up their publication process for Covid-19 βιοΡχιω 2020.04.18.045963;΄ δοι:¨<https://doi.org/10.1101/2020.04.18.045963>

Glasziou Paul P, Sanders Sharon, Hoffmann Tammy. Waste in covid-19 research BMJ 2020; 369 :m1847

UK Reproducibility Network Steering Group. UKRN Position on COVID-19 Research <https://www.bristol.ac.uk/psychology/research/ukrn/news/2020/ukrn-position-on-covid-19-research.html> (accessed June 2020)

APA: Haahr, M. (2020, May 26). RANDOM.ORG: True Random Number Service. Retrieved from <https://www.random.org>

Iacobucci Gareth. Sixty seconds on . . . following the science BMJ 2020; 369 :m2014

Lancet Editors. Expression of concern: Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. Lancet.2020.DOI: [https://doi.org/10.1016/S0140-6736(20)31290-3](https://doi.org/10.1016/S0140-6736%2820%2931290-3)

Armstrong S. How a scramble for COVID-19 evidence is leaving clinicians and the public wanting. BMJ2020;369:m2045

**Potential tables and figures**

|  |  |  |
| --- | --- | --- |
|  | **COVID-19 papers****n=xx** | **Non COVID-19 papers****n=xx** |
| **Method n (%)** |  |  |
| Randomised Controlled Trial |  |  |
| Observational  |  |  |
| Systematic reviews  |  |  |
| Test accuracy  |  |  |
| Brief report format  |  |  |
| **Content**  |  |  |
| Total number included  |  |  |
| Follow-up (weeks) |  |  |
| Positive result n (%) |  |  |
| Industry funding n (%) |  |  |
| **Post-publication n (%)** |  |  |
| Editorial  |  |  |
| Correction/retraction  |  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Methodological quality COVID-19 | Methodological quality Non COVID-19  | Reporting quality COVID-19 | Reporting quality non COVID-19 |
| Randomised Controlled Trial (n=xx) |  |  |  |  |
| Observational(n=xx)  |  |  |  |  |
| Systematic reviews (n=xx)  |  |  |  |  |
| Test accuracy(n=xx)  |  |  |  |  |

****



Appendix 1

Data dictionary for ‘included studies’ master sheet

**Study ID:** unique identifier for each paper returned from search

**Journal:** Journal name. Data collection begins in the first month that the journal published eligible COVID-19 research

**Month:** Each journal has weekly print publications, so month and week are recorded using format ‘Month week x/4 [assuming a four-week month]’

**Author:** surname of first author

**Design:** The primary method used in the paper, described using the pre-defined categories of: Controlled trial, Observational, Test accuracy, Systematic review, Qualitative, Prognosis. If the paper does not fit any of these it will not be included.

**Case series:** Where the paper is described by the authors as a case-study or case series.

**Brief report:** Where the journal describes the submission as a brief report or similar, these would have less content than a full original research paper. Some journals have a regular section for brief reports and some publish such reports on an ad-hoc basis.

**Correction / retraction:** Where a substantial change to content is made following publication. This will be checked last week of May. Journals use differing approaches to highlighting corrections, but for all of our included journals, major corrections and retractions are linked to the online article. The Lancet has a ‘Department of Error’ section, where any corrections, including minor changes, are described. For the Lancet, we defined a major correction as more than one entry in the Department of Error section.

**Editorial:** Whether the paper was accompanied by an editorial in the same journal. Some of our included journals routinely publish comment on their included research, to quality as an editorial the comment must be labelled as such in the journal.

**COVID:** If the primary focus of the paper was COVID-19 / SARS COV-2 or the outcomes relate to COVID-19 / SARS COV-2.

**Topic:** The clinical discipline to which the paper belongs, using where possible MeSH subject headings.

**Exposure:** The intervention (controlled trial), the exposure (observational), the index test (test accuracy).

**Outcome:** The primary outcome (if no primary outcome is described, then coded as ‘various’)

**Total ‘N’:** The total population included at baseline or first assessment. For systematic reviews this is the number of included papers.

**Follow-up:** Time from first measure to last measure for the primary outcome (in weeks); for some study designs for example cross-sectional, systematic reiew this is coded as N/A

**Results:** In the abstract or main text conclusions, do the authors claim a 'significant result' that rejects the null, if so then label as 'positive'; if no then 'neutral'.

**Funding:** Who supported the study. Coded as academic or industry, if funding was mixed this is coded as industry.