APPRAISAL OF THE SYSTEMATIC REVIEW BY BYRANT *et al*. ON USE OF IVERMECTIN FOR TREATMENT AND PREVENTION OF COVID-19

Date: 2 July 2021

Background:

An updated NEMLC COVID-19 rapid review of ivermectin (18 June 2021) for the management of COVID-19[1] was published on the National Department of Health website in June 2021. A meta-analysis and systematic review of randomised controlled trials (RCTs) for ivermectin by Byrant *et al.* had been published in the American Journal of Therapeutics on 17 June 2021[2]. This study was not included in the rapid review, and thus an appraisal of this review follows:

Overview:

Rosenthal[3] on meta-analysis: combining apples and oranges makes sense if your goal is to produce a fruit salad.

In the last few decades, reaching conclusions about the efficacy and safety of medical interventions has moved from reliance on expert opinion and narrative reviews to a more transparent and formalized collaborative process of searching, quality appraisal, and synthesis of all relevant evidence. The conclusions reached are critically dependent on unbiased adherence to all steps, and on the quality of the underlying evidence. A critical final process entails transforming conclusions about strength and direction of evidence into clinically useful recommendations, often by groups independent of the review process. A key principle is that decisions can and should be made using the best available evidence, even when this is imperfect.

Considerable time and effort goes into conducting high quality systematic reviews, and when done well, they are a valuable resource. Like any human endeavor, they still have vulnerabilities. The more obvious issues can be detected using quality appraisal tools such as AMSTAR[4] which evaluate whether a review meets the main reporting requirements, however the tool does not address the content of the review. There are other more subtle ways in which bias can occur rendering results less reliable. The rigour of the Cochrane process, and formal collaborative use of software such as RevMan[5] are specifically designed to address many of these issues.

Issues which may render the conclusions of a systematic review unreliable include undeclared intellectual conflicts of interest (where reviewers may not approach a research question entirely objectively), inconsistent rigour in risk of bias assessment (where studies supporting a particular viewpoint may be reviewed more leniently), inclusion of studies of low reliability, and issues with meta-analytic methods. This last point is particularly problematic in an era where software allows almost instantaneous iterative data analysis, which makes it difficult to determine whether a submitted data analysis plan is truly based on *a priori* scientific considerations or *post hoc* adoption of the model found to yield preferred results. Other issues in meta-analytic technique, such as the handling of studies that observed no outcome events in either arm, weighting methodologies, and the handling of heterogeneity and potential small study effects, engender vigorous debate, as in many other evolving areas of statistics.

The Bryant *et al.* review raises a number of concerning methodological issues. Some of these are described in more detail below, but the key issue is that no matter how rigorous and detailed the review and statistical analysis, the evidence pool is currently too small for reliable decision making. This review focuses only on mortality as findings for

[1] South African National Department of Health. Rapid review of Ivermectin for COVID-19 Update – 18 June 2021. http://www.health.gov.za/covid-19-rapid-reviews/

[2] Bryant A, Lawrie TA, Dowswell T, et al. Ivermectin for Prevention and Treatment of COVID-19 Infection: A Systematic Review, Meta-analysis, and Trial Sequential Analysis to Inform Clinical Guidelines. Am J Ther. 2021 Jun 17. https://pubmed.ncbi.nlm.nih.gov/34145166/

[3] Introduction to Meta-Analysis. Michael Borenstein, L. V. Hedges, J. P. T. Higgins and H. R. Rothstein © 2009 John Wiley & Sons, Ltd. ISBN: 978-0-470-05724-7 Chapter 40

[4] 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.

[5] Review Manager (RevMan) [Computer program]. Version 5.4, The Cochrane Collaboration, 2020.

all other endpoints were listed by the authors as based on low or very low quality evidence. The mortality endpoint was the only endpoint considered by the authors to be based on moderate quality evidence. For mild or moderate COVID-19, despite 11 trials, information on mortality was only available in five trials with a total of 13 deaths, and for severe COVID-19, on 5 trials, with a total of 539 patients, 200 of which were contributed by Elgazzar *et al*.'s study - reviewed below. The Naiee *et al.* study, in COVID-19 of undifferentiated severity, was not included in these two subgroup analyses, but contributed to the total analysis.

Authors of reviews can draw their own conclusions from their analysis, but the aim of scientific scrutiny is to allow others to look at the same information and potentially reach different interpretations. A responsible interpretation is not that this data is irrefutable proof of efficacy, but simply that information of this quality renders efficacy conclusions highly vulnerable to change as further data becomes available.

A few specific points:

1. The data search section states that that Kory and Malik were consulted as 'experts in the field'. As members of Front Line COVID-19 Critical Care Alliance (FLCCC), a group with previously demonstrated views supporting ivermectin use, they have taken a partisan and potentially biased, position as evident in their own narrative review in the same journal. There seems little evidence of a search for experts who might hold equivocal or negative views about ivermectin.
2. The table of included studies contain several situations where 'prepublication data/manuscript in progress/ obtained via email' was stated as the origin of the data. From the perspective of scientific method, this information is not currently available for public scrutiny and has not completed a peer-review process. (Some information listed in this way in the table is now published.) This leaves the reader with little opportunity to check validity. Including all available evidence is, in principal, a good practice. However the authors specifically state that they have not considered these data as adding potential risk of bias or decreasing certainty in the findings, a position that that would not be consistently held by reviewers.
3. The Elgazzar *et al.* study remains in the analysis despite some other studies at high risk of bias having been removed. Elgazzar *et al*. studied the effect of ivermectin vs hydroxychloroquine in a 6-arm trial that included both patients and contacts. The two arms that received ivermectin had deaths in 0/100 and 2/100, whereas those that received hydroxychloroquine had deaths in 4/100 and 20/100. Both arms received azithromycin as part of standard of care, so effectively the comparison was ivermectin and azithromycin versus hydroxychloroquine and azithromycin. Both of the latter agents are associated with QT prolongation. In addition, allocation concealment was unclear and randomisation procedures were not described in sufficient detail, it is unclear whether any blinding occurred, and the outcomes reported in the preprint differ from those in the trial registry. Studies with an active comparator may reduce apparent efficacy if the comparator is also active against the disease, or may flatter the trial medication if the comparator causes harm. Combining such studies with studies having a placebo control may introduce uncertainty.
4. A sub-analysis of studies was done removing studies at high risk of bias. This means that the primary analysis contained such studies. It is difficult to reconcile this with a statement that this constitutes moderate quality evidence.
5. The confidence interval for ivermectin's effect on mortality in mild to moderate COVID-19 ranges from 0.06 to 0.94, reflecting the paucity of events (1 death in the intervention arm and 12 in the control, out of 11 included studies, 6 of which (55%) observed no deaths in either arm). The confidence interval for use in severe COVID-19 includes 1, and thus is not statistically significant, even when including data from Elgazzar *et al.* Most of the other endpoints were contributed by the Fonseca study, one of only three considered at low risk of bias. Overall, one of the challenges with reviews of small trials is recognizing the 'fragility' of the results. When the number of deaths is so low, shifting one or two events from the ivermectin group to the control would change the result substantially from statistically significant to not[6].

[6] Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence--imprecision. J Clin Epidemiol. 2011;64(12):1283-93.

6. Another way of demonstrating the frailty of the evidence is using the authors' own study assessments. In the main forest plot, they include trials they indicate are at high risk of bias. In sensitivity analysis, these are removed. Another sensitivity analysis removes trials with active comparators. If both are done together (removing studies at high risk of bias and those with active comparators), no studies on severe COVID-19 remain, and the three remaining studies in mild COVID-19 together with the single study on mixed severity have a total of 24 events, with two thirds of the weight then provided by the Niaee *et al*. study.

Conclusion

Using evidence in clinical decision making requires meticulous attention to assessing both the quality of individual trials and how the information is pooled in a meta-analysis. Trials can be considered potentially misleading if their design, conduct, or reporting raise concerns; there is sound empiric evidence that failure to exercise caution in the face of these warning quality signs makes it highly likely that any conclusions drawn will be overturned by subsequent evidence.

As Guyatt *et al*.[6] stated, "Early trials addressing a particular question will, particularly if small, substantially overestimate the treatment effect. A systematic review of these early trials will also generate a spuriously large effect estimate. These considerations argue for skepticism regarding evidence summaries that generate apparent benefits, or harms, of therapy with what appear to be satisfactorily narrow CIs on the basis of small trials with relatively few events."

The Bryant *et al.* review contains data not yet available for peer review, includes in the primary analysis studies labeled by the authors themselves as at high risk of bias, and found low or very low quality evidence for all endpoints except mortality. After removal of trials at high risk of bias or with active comparators, the few remaining studies, with very few total events, are insufficient to provide reliable information. The sensible and responsible conclusion from this review is not that ivermectin is likely to be effective, but rather that there is currently insufficient evidence to justify recommending widespread use of this agent.

Reviewers: Andy Parrish, Tamara Kredo, Jeremy Nel, Halima Dawood, Trudy Leong, Milli Reddy.

Declaration of interests: AP (Walter Sisulu University); TK (Cochrane South Africa, South African Medical Research Council; Division of Clinical Pharmacology, Stellenbosch University; South African GRADE Network); JN (Department of Internal Medicine, School of Clinical Medicine, Faculty of Health Sciences, University of the Witwatersrand); HD (Infectious diseases, Greys hospital and University of KwaZulu-Natal); TL (National Department of Health, Affordable Medicines Directorate, Essential Drugs Programme) and MR (Better Health Programme, South Africa) have no interests with regards to ivermectin.

Appendix A: Evaluating the methodological quality of the Bryant et al (2021) systematic review and meta-analysis – AMSTAR 2 tool (Shea 2017[4])

No.	Criteria	Yes/ Partial Yes/ No	Comment
1	Research questions and inclusion criteria for the review included the components of PICO	Yes	There is no PICO in the review report.
2*	Report of the review contained an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol	No	Inclusion/exclusion criteria omitted, study protocol not registered.
3	Review authors explained selection of the study designs for inclusion in the review	No	No clear explanation provided why RCTs, Quasi-RCTs and Cluster RCTs were selected.
4*	Review authors used a comprehensive literature search strategy	Yes	-
5	Review authors perform study selection in duplicate	Yes	-
6	Review authors perform data extraction in duplicate	Yes	-
7*	Review authors provided a list of excluded studies and justify the exclusions	No	Excluded studies were merely referenced (ref# *47-63*), stating that they were not RCTs. However, ref# 47, Elgazzar et al is included in the analysis.
8	Review authors described the included studies in adequate detail	Partial yes	-
9*	Review authors used a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review	Partial yes	-
10	Review authors reported on the sources of funding for the studies included in the review?	Yes	-
11*	For meta-analyses, review authors used appropriate methods for statistical combination of results	No	The authors did not sufficiently justify combining the data in the meta-analysis, and why the Quasi-RCTs were not categorized as non-RCTs.
12	For meta-analyses, review authors assessed the potential impact of RoB in individual RCTs on the results of the meta-analysis or other evidence synthesis	Yes	-
13*	Review authors accounted for RoB in individual RCTs when interpreting/ discussing the results of the review	No	This was not adequately reported in the interpretation and discussion of the results of the review.
14	Review authors provided a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review	Partial yes	-
15*	For quantitative synthesis, review authors carried out an adequate investigation of publication bias (small study bias) and discussed its likely impact on the results of the review	Yes	-
16	Review authors reported any potential sources of conflict of interest, including any funding they received for conducting the review	No	Report states that *"authors have no conflicts of interest to declare"*, but have participated in initiatives promoting ivermectin.

** Critical domains = 2, 4, 7, 9, 11, 13, 15*

Rating overall confidence in the results of the review

- *High:* No or one non-critical weakness: the systematic review provides an accurate and comprehensive summary of the results of the available studies that address the question of interest
- *Moderate:* More than one non-critical weakness*: the systematic review has more than one weakness but no critical flaws. It may provide an accurate summary of the results of the available studies that were included in the review
- *Low:* One critical flaw with or without non-critical weaknesses: the review has a critical flaw and may not provide an accurate and comprehensive summary of the available studies that address the question of interest
- *Critically low*: More than one critical flaw with or without non-critical weaknesses: the review has more than one critical flaw and should not be relied on to provide an accurate and comprehensive summary of the available studies

(*Multiple non-critical weaknesses may diminish confidence in the review and it may be appropriate to move the overall appraisal down from moderate to low confidence).

OVERALL ASSESMENT: Critically low

Rationale: Four flaws in critical domains (#2, 7, 11, 13)

Conclusion: The AMSTAR assessment suggests that the review has more than one critical flaw and should not be relied on to provide an accurate and comprehensive summary of the available studies.