Author's response to reviews

Title: Specification of Implementation Interventions to Address the Cascade of HIV Care and Treatment in Resource Limited Settings: a Systematic Review

Authors:

Matthew Hickey (Matt.Hickey@ucsf.edu)
Thomas Odeny (deny@uw.edu)
Maya Petersen (mayaliv@berkeley.edu)
Torsten Neilands (Torsten.Neilands@ucsf.edu)
Nancy Padian (npadian@berkeley.edu)
Nathan Ford (fordn@who.int)
Zachary Matthay (zachary.matthay@ucsf.edu)
David Hoos (dh39@columbia.edu)
Meg Doherty (dohertym@who.int)
Chris Breyer (cbeyer1@jhu.edu)
Stefan Baral (sbaral@jhu.edu)
Elvin Geng (elvin.geng@ucsf.edu)

Version: 1 Date: 17 May 2017

Author’s response to reviews:

Responses are below. Please see separate attachment for formatted responses that are easier to read.

Reviewer #1: This article addresses an important topic, how to increase intervention implementation to improve HIV care. The literature reviewed is extensive, a large amount of data has been extracted, and the authors have described several limitations with this literature that will be of interest to clinicians and groups working in this area. However, there are a number of substantial issues with the manuscript in its current form which would need to be addressed before publication.
1) The review is focused on ‘implementation interventions’ however it is unclear to me how the authors defined ‘implementation interventions’ and differentiated this from trials of interventions. In its current form, the review does not explicitly give a definition of implementation interventions and in places the term implementation intervention appears to describe any intervention which is implemented. This means that some of the identified trials are focused on the efficacy of an intervention, rather than of different ways of implementing the intervention. As reviews of interventions to increase adherence to ART or increase uptake of HIV testing are already available, the current review is less novel than if it had focused tightly on implementation interventions. In addition, because the inclusion criteria mean that some interventions that reported outcomes such as CD4 count were excluded, it is arguably less complete than some of the existing reviews that directly focus on particular parts of HIV care rather than on implementation. Moreover the inclusion criteria states that interventions targeted at system, community or organisational levels were excluded on the grounds that these are not modifiable. I’m unclear from the description how the authors define these terms, however, it would seem likely to me that some system level interventions e.g. increasing time available for training, modification of financial incentives for healthcare providers, could be implementation interventions. From the information given it therefore seems likely to me that there may be missing relevant implementation interventions. The authors should give a definition of implementation intervention, and use this to select studies, or should reframe the paper in broader terms.

Thank you for these important points. We have added our definition of implementation interventions to the introduction (line 79) to further clarify our intent. Overall, our goal is to describe the quality and comprehensiveness of reporting for interventions designed to improve implementation of antiretroviral therapy (ART). In a sense, ART is the cornerstone evidence-based intervention for which all studies included in the review seek to improve effectiveness. We use the ‘cascade of care’ to further delineate set of steps required to effectively implement ART across the natural history of HIV treatment within an individual. By eliminating studies that evaluate only clinical outcomes (e.g. mortality, virologic suppression, CD4 count) without also measuring impact on at least one cascade step, our review focuses on studies that attempt to impact process outcomes with in ART delivery. Some of these studies compared a particular intervention to standard of care, while others compared different interventions or different ways of implementing a particular type of intervention.

Regarding the concern that our study included “interventions targeted at system, community or organizational levels”, the statement in the original manuscript was misleading. The original statement read “We excluded studies of system, organization, community or patient characteristics that may influence outcomes, but which are, in general, not directly modifiable.” We have edited this statement to clarify our intended meaning – that we excluded studies that were primarily “risk factor” studies, seeking to identify factors associated with a cascade step but that are not modifiable. For example, if a study sought to identify patient factors associated with poor adherence, but did not evaluate an intervention aimed at modifying adherence, it was excluded. Similarly, a study of provider incentive pay to increase ART initiation would be included; a study evaluating associations between provider socioeconomic status and ART initiation rates would not.
2) Although the PRISMA checklist has been completed, some of the information referred to is difficult to locate as it is only in supplements or absent from the paper. I strongly recommend that the authors restructure the Method and Results section (and the rest of the paper) considering PRISMA recommendations. For example, the search strategy should include a list of databases, sample search strategy, date of search. There should be clear statement in the paper of the exact inclusion and exclusion criteria, and a rationale for this provided. I would recommend that the authors consider framing this in terms of PICOS broadly to ensure that it is clear which sorts of interventions, outcomes and participants they focused on. Basic information, such as search dates, inclusion of unpublished studies e.g. dissertations, government reports, conference abstracts, whether authors were contacted for further information, how many people extracted data and whether they agreed on coding etc are missing from the paper. Likewise, using PRISMA to structure the Results section would enhance clarity e.g. the first section of the Results is headed 'Study Characteristics’ but initially describes the search. Some of this information is included in the supplements but more should be included in the main body of the paper. Even with the addition of supplements, more information e.g. dates of search, reviews used to identify other papers, experts consulted, should be provided.

We have added information from our Supplemental Appendix 1 into the body of paper to provide additional information on our search strategy. Additionally, we have updated the searches through February 28, 2017 per the reviewer’s request in #9 below. We have updated the PRIMSA statement to reflect these changes; it now appears that the relevant methods information can all be found in the paper, instead of requiring the reader to look in the supplemental material. The remaining information in the PRISMA statement that is marked ‘n/a’ is not relevant for this review as it refers to data synthesis, which is not undertaken in this review due to the vast heterogeneity of outcomes reported.

3) The authors refer throughout to using Proctor's headings to extract and code the papers. However, they do not extract information on all the points that Proctor refers to, and operationalize some of the terms recommended in the Proctor paper differently from the original recommendations. The authors should explain these differences and justify them as otherwise they are misrepresenting the original recommendations. For example, the Proctor paper recommends that 'Justification' is described. However, no mention of this is made in the current paper. Likewise, the 'action target' is described in Proctor's paper as being both the person whose behaviour the implementation seeks to change and the motivations of this person. In the current paper, the authors only focus on the later. The authors code for 'behavioural target' - I'm not clear whether this maps to the 'action' or the 'implementation outcome addressed' referred to in the original Proctor paper.

We have included additional text in both the introduction and the methods that describe the original components of the Proctor paper and how we have adapted them for this review. We clarify that we excluded ‘justification’ because it was ultimately too subjective in practice to qualitatively assess whether or not a paper included justification for their implementation strategy. Regarding the reviewers comments on the ‘action target’, we found it necessary to separate the compound concepts of both the actor and the motivations of that actor into separate measures. Thus, we defined ‘action target’ as the immediate expected outcome of the action on the actor for whom behavior change was intended. Specifically, we used Susan Michie’s COM-B
framework that describes capability, opportunity and/or motivation as the determinants of behavior change. If a study specifies some aspect of this framework as a target of the action, and thus, mediator of the action on the behavior change of interest, we considered this study to have specified the action target. We separately ascertained whether the study specified whether the behavior change was intended at a patient, provider, or organizational level; data on the distribution of the level of intervention can be found in Table 2.

4) There is information in the Results section (E.g. whether the study reported a positive effect) which is not mentioned in the Method. This should be rectified so that the reader can fully understand the Results. For example, how did the authors judge whether there was an intervention effect? i.e. effect size, significance level, which outcome was considered when multiple outcomes were reported? Likewise, for studies that reported, for example, an intervention delivered by community workers to patients, were patients/community workers/both included in the sample size? I recommend that the Method describes how this information was coded and extracted, and also that more details are reported in the Results. For example, what outcomes were the effects positive, what type of sample was used. It would be easier to understand the sample size data if it was perhaps split into the patients, action target, and actors rather than grouped together.

Thank you for this pickup. We have added additional text to the methods section (under the measurements heading) to clarify these procedures. Regarding the sample size, we included the number of patients for whom cascade outcomes were assessed in both the intervention and control arm of the study. Though we agree that further disaggregation of the sample size into patients, action targets and actors would be optimal, the substantial heterogeneity in study design, level at which the intervention was targeted, and quality of reporting limits our ability to report this level of detail in aggregate form.

5) More information should be given regarding how the intervention types were arrived at and how these were coded. For example, did each of the listed experts read a description of each intervention and then code it using pre-specified intervention labels? Were a subsample of papers used to derive these codes and then the codes used to extract information from the other papers? Was coding independent, and if so, what was the level of agreement between coders? For some of the codes, quite disparate interventions e.g. food disbursement and behavioural economics are labelled as being part of the same type of intervention approach. How were these judgements made?

The intervention types were arrived at through an iterative process with the study authors. We began by brainstorming general intervention types that the three listed authors were aware of (MDH, TAO, EHG). These three authors then independently reviewed a subset of 10 articles and independently collected data on the Proctor-derived reporting domains, cascade steps, and intervention types. We discussed these articles and updated our list of intervention types with new types that emerged from these papers and our discussion. We did not do formal testing of level of agreement, but through this process came to consensus about how to handle ‘grey’ areas.

6) The authors state that they just coded for presence/absence of the Proctor criteria, but do not describe how they judged this and how disagreements or difficult cases were dealt with. It seems
likely to me that there were some 'grey' areas, for example for text message reminders, who would be the intervention 'actor' as they are presumably delivered by an IT system. More information should be given on how these judgements were made to enable the reader to understand the meaning of the information in Table 3.

There generally were grey areas for a number of included articles. Our approach to areas of uncertainty was a bias toward giving studies credit for reporting. Our handling of the example of text message reminder systems was to give credit if any mention was made of the company used to send the messages, server system used to automatically send messages, IT/engineering staff who programmed the system, or actions taken by clinic staff to initiate the SMS messages.

7) I have doubts regarding whether the authors have sufficient power to conduct a linear regression to detect the effect of the different intervention approaches on intervention reporting when some of the intervention approaches were only used by 3 or 5 interventions.

While we agree that our numbers may be small, ultimately we do not think that this undermines the value of the observed effect sizes and confidence intervals seen in this study.[1] Furthermore, we have mitigated concerns about small numbers in two ways. First, we have eliminated the PCA categories for intervention approach in favor of a more understandable set of six intervention approaches (see #8 below). Second, we have updated our search through 2/28/2017, including an additional 51 studies. Thus our intervention approaches now have a minimum of 27 included studies.

8) The section on the PCA is quite unclear and I have doubts regarding the validity of this approach. My understanding is that, to conduct PCA, the interpretation can become controversial if continuous variables are not used. Were the relevant assumptions for this analysis checked? Was there a sufficiently large sample size (I've seen 150 quoted for continuous samples) to conduct PCA? The information presented in the supplementary material shows negative relationships between some of the intervention types and the components they were grouped under- are these valid groupings under a data driven approach such as PCA? The footnote states that some components were formed 'empirically'- what information was used to make these judgements- presumably this means that PCA was not used as the final arbiter of the classifications? A scree plot value of 0.2 is given as a cut off- my understanding is that 1 is a more typical value to use- the authors should explain why they used such a low value. More globally, I'm unsure of the value of using this type of analysis to categorize interventions, as different intervention strategies e.g. food supplements and counselling could plausibly co-occur in multiple studies but would not necessarily be the same intervention type, whereas similar intervention methods e.g. two different types of patient counselling may tend to not be given in the same intervention. The authors should discuss the usefulness of classifying such a wide range of intervention strategies into a small number of categories in this way in the Introduction and Discussion.

The reviewer’s concerns are valid and raise important points about the limitation of PCA for combining intervention types in this manuscript. Though we were hoping to combine intervention types into a smaller number of categories for analytic purposes through a data-driven process, we are ultimately limited by small numbers, lack of continuous variables and the
fact that different types of interventions are often included in the same study. Our scree plot values were also very low, as the reviewer mentioned. For all of these reasons, we have elected to forgo our PCA-based approach to combining intervention types into approaches. We have instead elected to combine intervention types into approaches empirically through consensus discussion with the authors. Though not a data-driven process, we believe these empiric categories reflect generally recognizable broad categories of interventions. The manuscript text has been updated to reflect this change.

9) From the abstract, it appears that the search was last updated in March 2014. I realise it is likely to be very time consuming but would strongly suggest that the review is updated.

We have now updated the searches through 2/28/2017 to incorporate more recent literature.

Reviewer #2: Thank you for the opportunity to review this interesting and novel systematic review. Whilst the aim of the review differed somewhat to the standard systematic review, it should be of interest to an audience involved in implementation science as well as to those interested in implementation of HIV care.

Major comments

- It seems that you used principal component analysis to group intervention types into broader intervention approaches. As this is quite novel, it would be beneficial if the authors went into further detail as to how they did this and why. The section on analysis in the main paper is not very clear. How were the intervention types mathematically transformed or determined in order to enable PCA? What decisions did the authors make and how, to name the particular components? When they mention collapsing 2 factors into the 10 main component intervention approaches, how did they decide this?

As described above in response to Reviewer #1’s comments on our PCA approach, we have eliminated the use of PCA to combine the intervention types and instead elected to use an empiric approach to defining intervention approaches. Please see discussion above in response to reviewer #1 for further discussion.

- Connected to this, the authors have decided to analyse at the level of intervention types and approaches, but they do not present anywhere in the paper on the individual studies (n=106) and how they fell into these two categories, and where the overlap may have been. It is likely that a number of intervention types or approaches might have been taken in one study (as they acknowledge), but the reader is not able to get a sense of the full variety and complexity of the studies drawn on here, apart from a brief mention on lines 174-182 of the Franke and Igumbor studies. I am not sure how meaningful it is to solely focus at the level of intervention approaches when many studies will have combined multiple approaches.

We certainly recognize the limitations presented by the fact that many studies combine multiple types of interventions into a single complex intervention. While in an ideal world, we would be
able to parse these complex interventions into their fundamental parts, this is generally not possible in most studies where sequential randomized trials or factorial designs are the exception, rather than the rule. For less robust analyses, we are forced to accept estimates of the entire complex intervention en bloc. We have strengthened our description of how we categorized studies into the various intervention types. We also further describe the process of classifying these types into the broader intervention approaches. Though this latter grouping process is empiric, the groupings are available in Table 2, allowing readers the opportunity to judge the appropriateness of the groupings. We do agree that a more granular approach would be helpful, though as can be seen from Table 2, the numbers of studies that evaluate many of the intervention types are small. With further iterations of this project, more granular analysis will be possible, though for now we are limited to grosser clustering of similar intervention approaches.

- Line 131, the authors say that they report on the 6 dimensions of Proctor but not on "quality of reporting". The reasoning for this should be made explicit. It would also be helpful, given the aim of the study is to look at reporting of implementation interventions, if the authors did some quality appraisal of the 106 studies beyond the Proctor framework. Is the body of work flawed in general, or is it more specifically to do with intervention reporting?

The reason that we elected to not assess quality of reporting is twofold. First, there is no clear framework for determining the quality of reporting for each of the domains we evaluated. Such an assessment is likely to be controversial and unlikely to be repeatable between different reviewers. We feel that our approach to count a domain as present if any description is included, however limited. Despite this very low threshold for counting a domain as present/reported, we still saw substantial reporting gaps. The second reason for deferring reporting on quality is that even if this were possible in a semi-objective way, the amount of effort required for such assessments in such a large number of articles was prohibitive. We have further clarified this in the text.

- Table 1: This table is not completely clear. Why have they named the columns "n or median" and "% or IQR"? Is column 2 not reporting number of studies by design, region, level of behavioural target, positive effect, and publication year? Should each of these categories add up to 106, given n of studies? Only design and positive effect groupings currently do.

- IQR - if this is referring to interquartile range, it makes little sense that the authors report a full range (392-5425) on line 169 and in table 1. Do they simply mean to report the range of sample sizes?

We apologize for the confusion in this table. We intended to convey that all the variables in the first column of data represented the n or number of studies with that particular characteristic. The exception to this is the sample size variable, which is reported as the median and interquartile range (IQR). The numbers reported under the IQR represent the range between the 25th and 75th percentile of the sample sizes within the included studies, rather than a “full range” as the reviewer may have misinterpreted. We have modified the headings of this table to add clarity.

- In table 2, only 32 intervention types are listed, but the authors summarise elsewhere that there were 35 types (line 247). Could they clarify?
Thank you for pointing this out. We had initially defined 35 types through the process now better described in the methods. However, there was one type, which despite identification by the authors as a potential type of intervention, did not appear in the studies selected by this review. Table 2 only lists the types that were identified in included studies. This number differs slightly from the prior manuscript because of the expanded search yielding 50% more articles than included in the previous manuscript version.

- What do the authors mean by "semi-coded or coded form" on line 111? Perhaps they could include a data extraction form as supplementary material, to make it clear what they extracted per study?

We have included our data extraction form as Supplemental Appendix 3.

- Could the authors define "resource-limited settings", as mentioned in line 100? Are they referring to the countries, or something more specific or at a different level?

We have clarified the methods to state that we included studies from low and middle-income countries as defined by the World Bank. Further details of the definition, including the countries included in the search term can be included in Supplemental Appendix 2.

- Table 3 is referred to in the text with means, and on the table heading with percentages, but is reported in decimal form in the table. It would be helpful if the authors could clarify this table and be consistent in their reporting.

We think perhaps the reviewer is referring to Tables 4 and 6. We agree that it is confusing to have proportions reported in decimal form within the table itself when the heading of the table refers to percents. We have changed the table to uniformly display percents. We also changed all references to the total score (out of 6 points) as a number ranging from 0-6. This is the way this was described in the prior text, but the table showed a proportion. We think that it is more intuitive to think of the total score in this way because it conveys the clearest information about what a given study contributed into the score (e.g. mean 3.5 components reported). We have kept the data for the individual reporting to percentages because the interpretation of this, we think, is also intuitive – e.g. 75% of Technology interventions reported a Behavioral Target.

- The discussion re-emphasises the importance of clear reporting and summarises the results further, but does not go into much depth on critically discussing why reporting might be unclear or the potential meanings of certain intervention types and certain cascade steps being better specified than others in the Proctor framework terms. Could the authors discuss this a bit further? It would also be helpful for the authors to discuss this in light of the studies themselves - with sometimes multi-faceted approaches - and the quality of the studies.

We have tried to include examples of studies with incomplete reporting to highlight key points without making the article even longer than its current length. We have also added some discussion of the possible reasons why ART retention had less complete reporting than other interventions.
- The authors could mention the purpose of process evaluations here - which tend to explore mechanisms and implementation of interventions in more depth - did they choose to exclude these types of studies? If so, why? Did any of the studies they included have process evaluations alongside their intervention studies? I think there is a critical point to be made around the specification of interventions in the types of studies included (which may be weak) and the likely enacted mechanisms of action, which may differ from what was specified.

We agree with the point about the importance of process evaluations. Process evaluations were certainly not excluded, though they were subject to the same criteria as other studies in that they must have a cascade outcome, comparator and be in a low or middle-income country. Included studies did variably assess for intervention effect on processes along the causal chain, with some more thoroughly evaluating all processes involved than others. Certainly more in-depth process evaluations would provide even more information about the potential effects of implementation, though the main argument we are trying to make is that all implementation intervention studies, whether process evaluations or not, should specify a minimum set of aspects to their intervention. From that baseline, further process evaluation would certainly deepen understanding of an intervention’s mechanism of action.

- Appendix 2 - this is almost another paper. The background part is probably unnecessary as it doubles with the paper background; however, the part on cascade steps and rationale for this being used should probably be in the main paper rather than the appendix. The Proctor framework could be shortened and made more accessible to the reader (e.g. defined in a table with examples included in the full paper)

We have added additional details from Appendix 2 into the methods to further clarify our process of assessing the implementation domains for each article. We have also added some of the examples from Appendix 2 into Table 1 for further clarity. While we agree that some of the background of Appendix 2 may be redundant from the paper, we would prefer to leave our protocol un-edited from its original form as the protocol for this study.

Minor

- Keywords are not provided

- Mistake in line 32 ("are have"), abstract

- Mistake on line 96 ("practices research community")

- Mistake on line 114 ("until")

- Mistake on line 259 ("issuing")

- Mistake on line 264 ("of reporting of reporting")
- Lines 308-312 - this is unclear as they state technological interventions report action and action target better then state technology interventions report action and behavioural targets less well.

- Mistakes on line 302 ("HCW's" and "peers health workers")
- the reference on line 300 is not in keeping with referencing elsewhere
- Figure 1 - define "PMTCT"

We have corrected the above errors.

Reviewer #3: Overall a very interesting analysis which extends our understanding of intervention reporting beyond the general to the specifics of HIV and which illustrates the implications of poor reporting for implementation. Unfortunately, the manuscript is let down by intermittent spelling/typographical errors, unclear phrasing and missing words which detracted from the reading and manuscript flow.

Abstract

Remove both references to "proposed by Proctor et al" in the Abstract (Reference will be sufficient in the main body of the text).

Change "Dissemination of standards such as those proposed by Proctor et al. for reporting can promote transparency, reproducibility and scientific accumulation" to "Dissemination of standards for reporting intervention characteristics can promote transparency, reproducibility and scientific accumulation"

"in in Africa" - delete one "in"

"may be a widely generalizable strategy for supporting the ability for newly diagnosed patients to navigate the psychological as well as practical aspects of treatment" - unnecessarily wordy
Prfere: "may be a widely generalizable strategy to support newly diagnosed patients in navigating the psychological, as well as practical, aspects of treatment"

"SMS" - first time in full with abbreviation in brackets

"Accumulation of generalizable knowledge about implementation interventions addressing HIV in low and middle income countries (LMIC) depends on adequate specification and reporting interventions in implementation research - without which reproducibility, transparency, and generalizability are undermined." - Again this is unnecessarily wordy "Accumulation of generalizable knowledge about implementation interventions" and "implementation research" are essentially synonymous.
"Many studies have synthesized outcomes of implementation interventions [11-16], but none have explicitly addressed the completeness of intervention specification and reporting" - you should make it clear that you are referring here specifically to the context of HIV/AIDS because as a general statement this would be untrue.

"In the identified manuscripts, we document the completeness of intervention reporting using an approach adapted from Proctor et al." Clumsy construction. Prefer: "We document the completeness of intervention reporting in a set of published reports using an approach adapted from Proctor et al." Strictly speaking these reports are no longer "manuscripts" which typically is used to refer to submissions before publication.

"and thereby influence standard practices research community addressing HIV treatment in LMIC." Word(s) missing?

The search strategy as reported in the paper misses most of the essential ingredients of search strategy reporting (cp STARLITE - sampling strategy, type of study, approaches, range of years, limits, inclusion and exclusions, terms used, electronic sources). I know that much of this detail is contained in Supplementary Appendix 1 but within the text I can see reference only to type of study (comparative) and inclusions and exclusions. Within the text itself we particularly need to know databases searched, years and languages searched, and any supplementary approaches e.g. follow up of references etcetera. The supplement should be primarily for the search terms used.

Similarly, the description of the sifting process is an essential part of a systematic review method and should not be relegated to an appendix. Placing essential details of Methods in an Appendix is not PRISMA compliant.

Regarding these two prior points of feedback, we have now included details about our search and sifting process in the beginning of the methods section.

Correct "until of analysis"

Correct "should we specified"

Specifically refer to COM-B in the following sentence: "we operationalized the action target to be the capabilities, motivations or opportunities of an individual or organization which mediates the effect of the intervention on a subsequent desired behavior based on work by Michie et al"

"no two papers studied the exact same intervention" - Strictly speaking you have by this stage moved from papers (the unit of retrieval) to studies (the unit of analysis) so prefer "no two studies examined the exact same intervention" - of course you could retrieve multiple study reports from the same study in which case they would contradict the statement you have made.

"presented in a manuscript" - prefer "study report"

Correct spelling of "principle components analysis (PCA)."
"235 received full reviews of abstracts and full text when necessary" This a cumbersome construction partly because of the recurrence of the word "full". Prefer: "The abstracts, and where necessary full text, of 235 articles were examined by both reviewers".

"intervention in South Africa which facility based staff made home visits" Word missing.

Either delete "of" or add "the" in the following: "We additionally redistributed two of dimensions with small numbers for a total of ten final intervention groupings"

The ten intervention groups would be better in a table with numbers of occurrences of reports of each rather than in the "listing " paragraph from facility based delivery through to non-facility based delivery (p. 8)

"The actor was reported below 50% in facility-based service delivery" - Avoid this potentially ambiguous shorthand, Prefer "The actor was reported in less than 50% of reports of facility-based service delivery interventions" etcetera

"generally not optimally reported" Adding "optimally" here conlates the issues of whether they are reported at all and whether they are reported fully. Please check which of these is/are correct based on what you analysed.

"Action target reporting followed a similar pattern health care worker improvement" Should "with" be inserted after "pattern"?

"The behavioural target was reported in more than 75% of the time for six out of the ten approaches" Avoid "of the time" in all occurrences because of potential for confusion with temporality dimension. Prefer "75% of the reports".

"higher reporting" - this is not a valid concept - either "higher levels of reporting" or more probably "more complete reporting" Similarly with all occurrences of "low reporting"

"The only dimension commonly reported more than 75% across cascade steps was the behavioral target." - Again this is shorthand which obscures precise meaning.

"Overall, we found that implementation interventions addressing adult HIV care and treatment are not optimally specified." I suggest that you report completeness of reporting (across reports) first followed by specification of reporting (within individual reports). As mentioned the term "optimally" can potentially confuse these issues.

"even this dimension was only reported in only 67% of manuscripts" Replace "manuscripts" with "papers"

"67% of manuscripts - substantially below the desired 100%." This sounds rather patronising! Perhaps prefer "67% of manuscripts, falling substantially short of universal coverage." Or similar.
You only introduce the two occurrences of intervention phenotypes/phenotypic in the Discussion. It might be helpful to introduce this term earlier in the description of the classification process.

"the average completeness of reporting of reporting" No we need "Of reporting" twice?

"The absence of change" - Prefer "absence of improvement" as only uni-directional change is desired here.

"that dissemination efforts to bring fresh from outside of the field are needed." Words missing?

"Although much of the efforts in transparency have focused on analysis" Prefer "Although much effort in improving transparency has focused on analysis"

"Variation in peer based interventions along with the variable reporting" Insert "delivery of" to improve rhetorical point "Variation in the delivery of peer based interventions along with the variable reporting"

"is perhaps one reason" - No that is two reasons!

"influenced peers health workers" peer not peers

"prepped" consider either "prepared" or placing "prepped" within quotes.

"Mechanistic clarity positions results to have maximal relevance for diverse contexts." While this is admirably terse it takes re-reading for understanding. In particular, "mechanistic" seems judgemental which seems to be the first time you have acknowledged this side of the debate. Perhaps clarify as "While some commentators may criticise such mechanistic clarity we should acknowledge its critical role in positioning results so as to have maximal relevance for diverse contexts"

"First, there is no single search term that will automatically identify implementation interventions." - "automatically" is not the point here - your meaning implies "consistently identify"

References:

Some references are incomplete e.g. Volume; Issue Page Nos for "Govindasamy D, Ford N, Kranzer K. Risk factors, barriers and facilitators for linkage to art care in sub-saharan africa: a systematic review. Aids 2012." In this case having "art" in lower case obscures the abbreviation!

And:

therapy initiation among adults in Malawi: a randomized clinical trial." No Source, date pagination etcetera.

e.g. "Zeng W, Rwiyereka AK, Amico PR, Avila-Figueroa C, Shepard DS. Efficiency of HIV/AIDS Health


Numbers on PRISMA flowchart need checking e.g. 9719 + 235 ≠ 9955; and 9109 + 608 ≠ 9719

Thank you for the detailed proofreading and comments. We apologize for the numerous grammatical errors – though not an excuse, it seems that the non-proofread draft of our final manuscript was submitted, rather than the version that had undergone more extensive copyediting. Nonetheless, we believe the grammatical errors to now be addressed. We also appreciated the rewording of ambiguous or misleading statements and have made edits to the text where appropriate.

References: