BMJ Open

The effect of in-utero exposure to HIV and antiretroviral drugs on growth in HIV-exposed uninfected children: a systematic review and meta-analysis protocol

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-023937
Article Type:	Protocol
Date Submitted by the Author:	01-May-2018
Complete List of Authors:	Ekali, Gabriel; Universite de Yaounde I, Biotechnology Center; National AIDS Control Committee, Jesson, Julie; Universite Toulouse III Paul Sabatier, Inserm UMR 1027 Enok, Pascal; University of Montreal, Department of Preventive Medicine Leroy, Valériane; Universite Toulouse III Paul Sabatier
Keywords:	HIV & AIDS < INFECTIOUS DISEASES, Nutrition < TROPICAL MEDICINE, PAEDIATRICS

The effect of *in-utero* exposure to HIV and antiretroviral drugs on growth in HIV-exposed uninfected children: a systematic review and meta-analysis protocol

Gabriel Loni Ekali[1,2,3,*], Julie Jesson[2], Enok Pascal[4], Valériane Leroy[2]

1.Biotechnology Center, University of Yaounde 1, Cameroon.

2. Inserm UMR 1027, Université Toulouse 3, Toulouse, France.

3. National AIDS Control Committee, Ministry of Public Health, Cameroon.

4.Department of Preventive Medicine, University of Montreal.

***Corresponding author**: Dr Gabriel Loni Ekali

Biotechnology Center, University of Yaounde 1, Cameroon; Inserm UMR 1027, University of Toulouse 3, Toulouse, France ; National AIDS Control Committee, Ministry of Public Health, Cameroon.

Email: eloni2000@yahoo.com

Telephone: 237 - 677333023

PO Box: 3077, Yaounde, Cameroon.

Word count: 1,933

ABSTRACT

Introduction

HIV-exposed uninfected children (HEU) have higher morbidity and mortality compared to HIV unexposed uninfected children. Despite the fact that malnutrition contributes to about half of all infant deaths below 5 years of age and that growth impairment has been reported in the HEU population, the spectrum of growth disorders associated to HIV and ART exposure during the *in utero* and perinatal periods is yet to comprehensively summarized amongst the global HEU population. This protocol for a systematic review and meta-analysis aims to critically synthesize data concerning the prevalence of underweight, stunting and wasting at different ages in the global HEU population.

Methods and analysis

Medline, EMBASE, Cochrane Library, TOXLINE, WHO Global Index Medicus and the Web of Science will be searched for relevant articles published between January 1st 1989 and December 1st 2017 without language restriction. In addition, conference abstracts and reference lists of eligible papers and relevant review articles will be screened. Authors will screen and select studies, extract data, assess risk of bias as well as studies individually for heterogeneity. Study-specific estimates will be pooled through a random-effects meta-analysis model for studies that are clinically homogenous while funnel plots and Egger's test will be used to detect publication bias. Results will be presented by ART availability period, country income levels, mode of breastfeeding.

Ethics and Dissemination

Ethical approval will not be required for this study because it will be based on published data. The final report of this study will be published in a peer-reviewed journal and presented at scientific conferences. This review will summarize the evidence and quantify the problem of growth impairment in HEU infants and so shed more light on our understanding of the higher morbidity and mortality in this growing population.

PROSPERO registration number: CRD42018091762

Key words: Growth, wasting, stunting, underweight, HIV-exposed uninfected children

STRENGTHS AND LIMITATIONS OF THE STUDY

- One of the limitations of the current review could be insufficient data to quantify the growth outcomes of interest.
- To our knowledge, this will be the first study to summarize the effect of HIV and ART exposure on growth of HIV-exposed uninfected infants compared to an appropriate control group of HIV-unexposed uninfected children.

- This review will cover both the era before widespread combination ART and the era after widespread access to combination ART so as to investigate whether any differences in growth outcomes exists between these two eras.

BACKGROUND: One of the major strategies of reducing new pediatric Human Immunodeficiency Virus (HIV) infections is by preventing mother-to-child transmission (MTCT) of the virus. This transmission can occur in utero, during delivery or breastfeeding.[1] In the absence of any intervention MTCT of HIV can be as high as 45%.[2] However, with antiretroviral therapy (ART), the rate of MTCT can be reduced to less than 2%.[3] With widespread access to ART, the UNAIDS report that new pediatric HIV infections have declined by 47% between 2010 and 2016, while coverage of antiretroviral medicines– provided to pregnant women living with HIV to prevent transmission to their children rose from 47% to 76% over the same period.[4]

As MTCT rates drop, the proportion of children who are exposed in utero to HIV and ART and are born uninfected with the virus (HEU) is increasing worldwide, especially in sub-Saharan Africa where the greatest burden of the HIV pandemic is found. Indeed, HEU constitute about 30% of the Southern African population.[5] Even though HEU are not infected with HIV, a growing body of literature shows evidence of higher morbidity and mortality compared to HIV/ART unexposed uninfected children (HUU). Higher frequency of hospitalisations, more severe respiratory tract infections especially from encapsulated bacteria as well as diarrheal diseases have been described.[5-6] Children exposed to maternal HIV and ART have modest but significant impairment of development and a higher risk of growth impairment.[7] A number of recent reviews have dwelt on summarizing the available evidence of higher morbidity, mortality, immune dysfunction and its possible role in the higher susceptibility of HEU to diseases.[5-7,8-9] However, the spectrum of growth disorders found in this growing population and its possible association and role to the higher susceptibility to infectious diseases they experience is yet to be documented. This is important as it is known that impaired growth clinically represents an altered nutritional status, which is significantly associated to mortality in infancy. Indeed, nutrition-related factors contribute to about 45% of deaths in children under 5 years of age, translating into the loss of approximately 3 million young lives annually.[10] In addition, appropriate comparisons of HEU and HUU were rare before the antiretroviral era. With increasing ART access, it is methodologically challenging to disentangle effects of HIV exposure from those associated with ART exposure, particularly in regions where malnutrition and co-morbidities are common and could increase the potential effects of these exposures.

In view of the recent reports of growth disorders among HEU infants, the main goal of this paper is to review and summarize evidence on the effects of perinatal/postnatal exposure to HIV and ART on growth outcomes in HEU children until 60 months of age and -in later life.

Review question

What is the effect of in utero/postnatal exposure to HIV and ART on growth outcomes until 60 months of age in HEU compared to HUU children as documented in studies reported between January 1st, 1989 to December 1st, 2017.

Specific objectives

To answer this review question, a systematic review and meta-analysis of the existing evidence will be conducted, the objectives of which will be to determine:

1. The effect of HIV and ART exposure on weight-for-age until 60 months of age in the global HEU population compared to HUU children.
2. The effect of HIV and ART exposure on height/length-for-age until 60 months of age in the global HEU population compared to HUU children.
3. The effect of HIV and ART exposure of weight-for-height/length until 60 months of age in the global HEU population compared to HUU children.

METHODS AND ANALYSIS

Design

This will be a systematic review and meta-analysis of the published literature. The Centre for Reviews and Dissemination guidelines [11] will be used for the methodology of this review and the review will be registered in the PROSPERO International Prospective Register of systematic reviews.

Inclusion criteria:

1. Type of studies: all observational cohort studies and randomized controlled trials that have data on any growth parameter of interest, either as main or secondary outcomes. Only comparative studies with HEU and HUU children will be included in the meta-analysis.
2. Type of participants: adequately defined HEU (children exposed in utero to HIV and ART, who are declared HIV negative by either DNA PCR between 6 weeks and 18 months OR appropriate HIV serological tests above 18 months) and HUU children aged 0 to 5 years.
3. Outcome measures: prevalence of underweight, stunting, wasting; growth velocity or studies that have enough data to compute the outcomes of interest.
4. Type and period of publications: published studies and conference abstracts from 1 January 1989 to 1 December 2017.

For studies with several publications of their findings over time, we will retain in the review the one that has the best quality and the most appropriate regarding our objective. For the meta-analysis, studies with a documented measure of effect of in utero/postnatal HIV and ART exposure on growth comparing HEU and HUU groups will be included. For studies without reported measures of effect, these will be computed if the provided data is adequate.

Exclusion criteria:

We will exclude:

1. Type of studies: unpublished manuscripts, case reviews, policy reports, commentaries and editorials.
2. Studies whose data will not be sufficient to calculate the appropriate measures of effect.

Outcome measures: The following measures will be considered at different age points (6, 9,12,18,24 months etc)

- weight-for-age
- height-for-age
- weight-for-height/length
- growth velocity
- prevalence of stunting, wasting at each age point.

Search strategy to identify relevant studies

Databases: An exhaustive literature search will be carried out in Medline, EMBASE, Cochrane Library, TOXLINE, WHO Global Index Medicus and the Web of Science to identify relevant articles published on growth in HIV-exposed uninfected infants between January 1st 1989 and December 1st 2017. The proceedings of the following conferences will also be searched: International AIDS Society Conference on HIV Pathogenesis, Treatment and Prevention; International AIDS Conference; Conference of Retroviruses and Opportunistic Infections. The search strategy that will be used for online databases is presented in table 1.

Table 1: Search strategy for the review of impact of in utero exposure to HIV and ART on growth in HIV-exposed uninfected infants

Search	Search Terms
#1	HIV OR HIV-1 OR HIV-2 OR HIV1 OR HIV2 OR HIV infect* OR human immune*
#2	HIV-exposed uninfected OR HIV exposed uninfected OR HEU OR HIV-EU OR HIV-unexposed OR HIV unexposed OR HUU OR HU OR HIV-UE OR HIV-uninfected OR HIV uninfected
#3	Children OR child OR infant OR infants OR pediatric OR pediatrics OR paediatric OR paediatrics
#4	Growth OR development OR stunting OR malnutrition OR wasting OR underweight

	OR weight OR length OR height OR BMI OR head circumference OR growth faltering
#5	Limits: from 1 January 1989 to 3 November 2017
#6	#1 AND #2 AND #3 AND #4 AND #5

Searching other sources

The list of references of eligible articles as well as relevant review articles will also be manually searched.

Selecting studies for inclusion in the review

All articles returned by the search will be saved to the Endnote version X4 software which will be used to remove duplicates. The titles and abstracts of the articles remaining after exclusion of duplicates will be reviewed to identify potentially relevant papers according to the inclusion and exclusion criteria. Two investigators will independently review retrieved full text of all identified papers for eligibility and then retain those that are eligible by consensus. Any disagreement between investigators will be settled by a third reviewer. A PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram will be used to detail the number of articles identified, screened, included and excluded [12] (online supplementary file 1).

Data extraction and management

A tool for data abstraction will be developed and the following data will be extracted:

1. Details about the author: Name, Journal
2. Characteristics of the study: design, country, period when study was conducted, year of publication, duration of follow-up, sample size, study endpoints,
3. Characteristics of participants: Age, sex, type of maternal ART, mode of infant feeding (no BF, mixed feeding, exclusive BF), duration of breastfeeding (BF), Cotrimoxazole prophylaxis, growth references used,
4. Outcome measures at different time points: weight-for-age, length-for-age, weight-for-length, BMI for age.

Two investigators will independently extract data and any disagreements will be reconciled through discussion or by a third investigator as necessary.

Assessing methodological quality of included studies and risk of bias

The quality assessment tool for Observational Cohort and Cross-sectional studies of the National Heart, Lung and Blood Institute (NHLBI) will be used to assess the quality of observational cohort studies [13] (online supplementary file 2) while for intervention studies, the Effective Public Health Practice (EPHPP) tool [14] (online supplementary file 3) will be used by two investigators independently. Agreement between the two investigators will be

measured by Cohen's kappa statistic. For the NHLBI tool, the score will be graded as Good, Fair and Poor while for the EPHPP tool, it will be graded as Strong, Moderate and Weak.

Data synthesis and analysis

Prevalences of underweight, stunting and wasting in the HEU and HUU groups will be recalculated based on the information provided by individual studies and odd ratios will be computed at each age point. The exposure variable will be HIV and ART exposure in utero and the outcome variable will be the different measures of growth: length-for-age, height-for-age, weight-for-age. We will use the statistic $I^2 = 100*(Q-df)/Q$ Q (with Q being the statistic of the Cochran heterogeneity Q-test and df the number of degrees of freedom corresponding to the number of studies minus one), which makes it possible to capture the proportion of the total variance observed due to a real difference in the measures of effects between the studies, in order to explore heterogeneity. Unlike the Q-test, this statistic is not influenced by the number of studies. Moreover, a non-significant Q-test does not necessarily imply that there is no heterogeneity between the studies because the power of the test depends on the number of studies. For statistical I^2, the values of 0%, 25%, 50% and 75% respectively represent the following heterogeneity levels: absence, weak, moderate and high. For each risk factor, the combined measure of relative risk will be estimated from a random-effect model if $I^2 \geq 25\%$ and a fixed-effect model if $I^2 < 25\%$. If substantial heterogeneity is observed, a meta-regression will be used to identify the characteristics of the studies (e.g study quality, study location, sample size, adjusting of confounders or not, antiretroviral therapy agents and regimens, duration and mode of breastfeeding, maternal socioeconomic status, etc.) that may explain the observed heterogeneity. All data will be analyzed using Stata V.14 (StataCorp, Texas, USA). A forest plot will be presented for each factor studied in the studies found. Qualitative synthesis will be used in cases where data extracted is insufficient to perform quantitative synthesis or when the included studies significantly differ in design, setting and outcome measures.

Assessment of reporting bias

To explore a potential publication bias, funnel plots and Egger's test will be done and a p value below 0.10 will be considered indicative of significant publication bias.

Presentation and reporting of the results

The results of this meta-analysis will be presented by HIV and ART exposure status, by pre-ART and post-ART era, by country income level and by infant feeding mode. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocol (PRISMA-P) will be used for reporting this present protocol (online supplementary file 4).[14] Results of the search method will be presented with a flow chart, detailing the selection process. Causes of exclusion for studies firstly identified as eligible will be documented.

POTENTIAL PROTOCOL AMENDMENTS

The current protocol as written will not be amended in the course of the study. This is to avoid any outcome reporting bias.

ETHICAL CONSIDERATIONS AND DISSEMINATION

In this study, only published data will be used. No ethical clearance will therefore be needed. The final results will be published in peer-reviewed journals and also presented at conferences.

CONCLUSION

A comprehensive summary of the evidence concerning how in utero/postnatal exposure to HIV and ART affect growth is critical to shed more light on our understanding of why HEU children have higher morbidity and mortality to certain infectious diseases compared to HUU. It is hoped that the results of this review will quantify the problem of growth impairment and draw attention to the necessity to possibly develop appropriate nutritional interventions for this growing population.

AUTHOR STATEMENT: EG, LV had the idea and EG designed the protocol. LV, JJ, EP critically reviewed intellectual content and revised the methodology of the study. EG and LV are guarantors of the study and all the authors approved the final version of the manuscript.

FUNDING: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

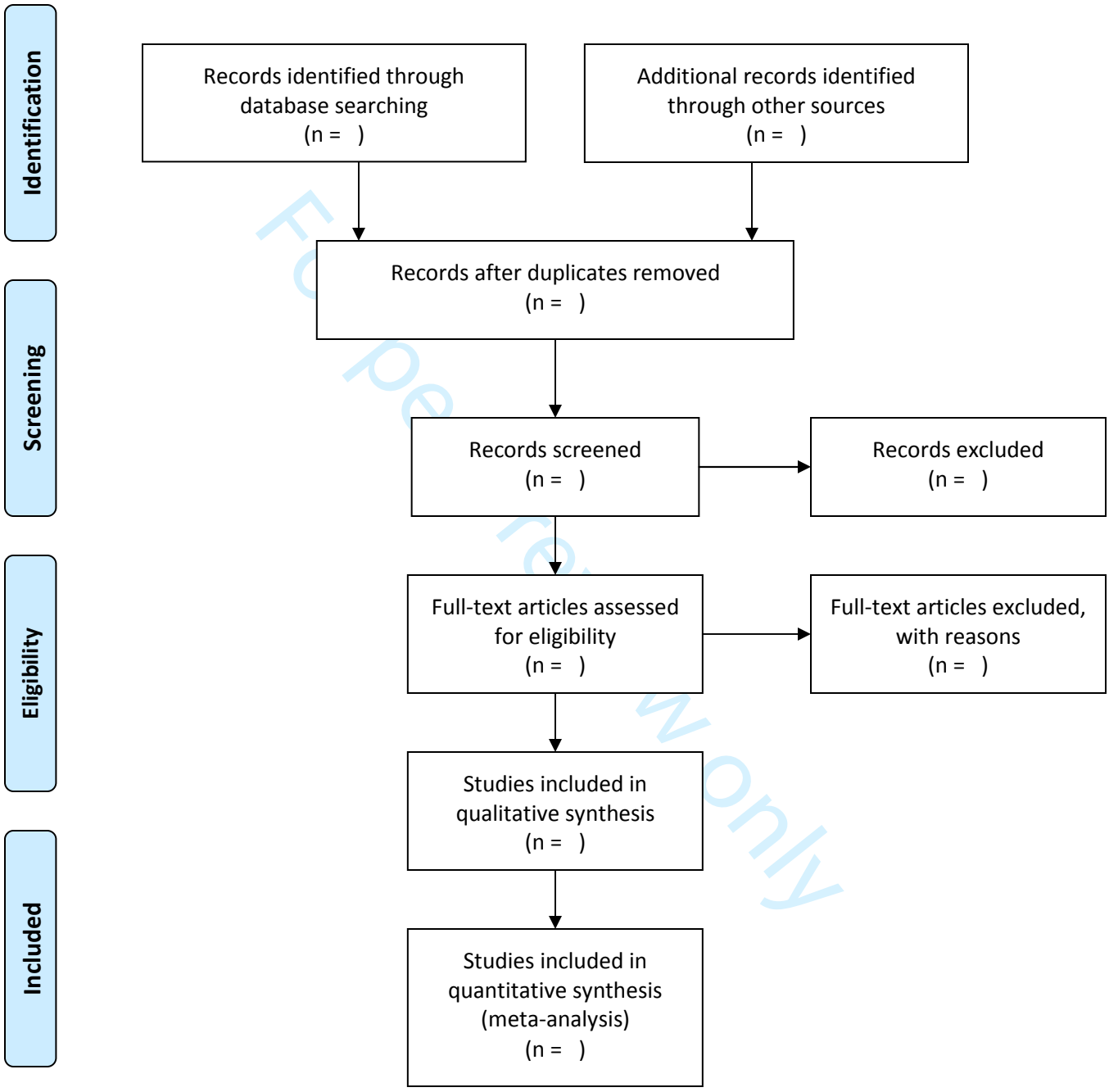
COMPETING INTEREST: None declared

REFERENCES

1. Mofenson LM. Mother-child HIV-1 transmission: Timing and determinants. *Obstet Gynecol Clin North Am* 1997; 24:759.
2. Lehman DA, Farquhar C. Biological mechanisms of vertical human immunodeficiency virus (HIV-1) transmission. *Rev Med Virol*. 2007;17:381–403
3. Brocklehurst P: Interventions for reducing the risk of mother-to-child transmission of HIV Infection. *Cochrane Database Syst Rev* 2002;1:CD000102
4. UNAIDS data 2017. Available at http://www.unaids.org/sites/default/files/media_asset/20170720_Data_book_2017_en.pdf. Accessed on January 29th 2018 at 00h24minutes.
5. Slogrove AL, Goetghebuer T, Cotton MF et al. Pattern of Infectious Morbidity in HIV-Exposed Uninfected Infants and Children. *Front Immunol* 2016 May 6;7:164
6. Brennan AT, Bonawitz R, Gill CJ, et al. A meta-analysis assessing all-cause mortality in HIV-exposed uninfected compared with HIV-unexposed uninfected infants and children. *AIDS* 2016 Sep 24;30(15):2351-60
7. Desmonde S, Goetghebuer T, Thorne C, et al. Health and survival of HIV perinatally exposed but uninfected children born to HIV-infected mothers. *Curr Opin HIV AIDS* 2016 Sep;11(5):465-476

8. Abu-Raya B, Kollmann TR, Marchant A, et al. The Immune System of HIV-Exposed Uninfected Infants. *Front Immunol* 2016 Sep 28;7:383
9. Ruck C, Reikie BA, Marchant A, et al. Linking Susceptibility to Infectious Diseases to Immune System Abnormalities among HIV-Exposed Uninfected Infants. *Front Immunol* 2016 Aug 19;7:310
10. WHO Media Center. Children : reducing child mortality Fact sheet. Available from http://www.who.int/mediacentre/factsheets/fs178/en/. Accessed on February 2nd 2018 at 23h56.
11. Centers for Reviews and Dissemination. CRD's guidance for undertaking reviews in healthcare. Centers for Reviews and Dissemination; 2009. Available from: http://www.york.ac.uk/crd/SysRev/!SSL!/WebHelp/SysRev3.htm. Accessed February 2nd 2018 at 01h04
12. National Heart, Lung and Blood Institute. Quality assessment tool for Observational Cohort and Cross-sectional studies. Available from https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools and accessed February 2nd 2018 at 01h06.
13. Effective Public Health Practice Project. Quality assessment tool for quantitative studies. Available at http://www.ephpp.ca/PDF/Quality%20Assessment%20Tool_2010_2.pdf . Accessed on February 2nd 2018 at 01h07.
14. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.doi:10.1186/2046-4053-4-1

PRISMA 2009 Flow Diagram



From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit www.prisma-statement.org.

Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies

Criteria	Yes	No	Other (CD, NR, NA)*
1. Was the research question or objective in this paper clearly stated?			
2. Was the study population clearly specified and defined?			
3. Was the participation rate of eligible persons at least 50%?			
4. Were all the subjects selected or recruited from the same or similar populations (including the same time period)? Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?			
5. Was a sample size justification, power description, or variance and effect estimates provided?			
6. For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?			
7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?			
8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?			
9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?			
10. Was the exposure(s) assessed more than once over time?			
11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?			
12. Were the outcome assessors blinded to the exposure status of participants?			
13. Was loss to follow-up after baseline 20% or less?			
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?			

Quality Rating (Good, Fair, or Poor) (see guidance)
Rater #1 initials:
Rater #2 initials:
Additional Comments (If POOR, please state why):

*CD, cannot determine; NA, not applicable; NR, not reported

Guidance for Assessing the Quality of Observational Cohort and Cross-Sectional Studies

The guidance document below is organized by question number from the tool for quality assessment of observational cohort and cross-sectional studies.

Question 1. Research question

Did the authors describe their goal in conducting this research? Is it easy to understand what they were looking to find? This issue is important for any scientific paper of any type. Higher quality scientific research explicitly defines a research question.

Questions 2 and 3. Study population

Did the authors describe the group of people from which the study participants were selected or recruited, using demographics, location, and time period? If you were to conduct this study again, would you know who to recruit, from where, and from what time period? Is the cohort population free of the outcomes of interest at the time they were recruited?

An example would be men over 40 years old with type 2 diabetes who began seeking medical care at Phoenix Good Samaritan Hospital between January 1, 1990 and December 31, 1994. In this example, the population is clearly described as: (1) who (men over 40 years old with type 2 diabetes); (2) where (Phoenix Good Samaritan Hospital); and (3) when (between January 1, 1990 and December 31, 1994). Another example is women ages 34 to 59 years of age in 1980 who were in the nursing profession and had no known coronary disease, stroke, cancer, hypercholesterolemia, or diabetes, and were recruited from the 11 most populous States, with contact information obtained from State nursing boards.

In cohort studies, it is crucial that the population at baseline is free of the outcome of interest. For example, the nurses' population above would be an appropriate group in which to study incident coronary disease. This information is usually found either in descriptions of population recruitment, definitions of variables, or inclusion/exclusion criteria.

You may need to look at prior papers on methods in order to make the assessment for this question. Those papers are usually in the reference list.

If fewer than 50% of eligible persons participated in the study, then there is concern that the study population does not adequately represent the target population. This increases the risk of bias.

Question 4. Groups recruited from the same population and uniform eligibility criteria

Were the inclusion and exclusion criteria developed prior to recruitment or selection of the study population? Were the same underlying criteria used for all of the subjects involved? This issue is related to the description of the study population, above, and you may find the information for both of these questions in the same section of the paper.

Most cohort studies begin with the selection of the cohort; participants in this cohort are then measured or evaluated to determine their exposure status. However, some cohort studies may recruit or select exposed participants in a different time or place than unexposed participants, especially retrospective cohort studies–which is when data are obtained from the past (retrospectively), but the analysis examines exposures prior to outcomes. For example, one research question could be whether diabetic men with clinical depression are at higher risk for cardiovascular disease than those without clinical depression. So, diabetic men with depression might be selected from a mental health clinic, while diabetic men without depression might be selected from an internal medicine or endocrinology clinic. This study recruits groups from different clinic populations, so this example would get a "no."

However, the women nurses described in the question above were selected based on the same inclusion/exclusion criteria, so that example would get a "yes."

Question 5. Sample size justification

Did the authors present their reasons for selecting or recruiting the number of people included or analyzed? Do they note or discuss the statistical power of the study? This question is about whether or not the study had enough participants to detect an association if one truly existed.

A paragraph in the methods section of the article may explain the sample size needed to detect a hypothesized difference in outcomes. You may also find a discussion of power in the discussion section (such as the study had 85 percent power to detect a 20 percent increase in the rate of an outcome of interest, with a 2-sided alpha of 0.05). Sometimes estimates of variance and/or estimates of effect size are given, instead of sample size calculations. In any of these cases, the answer would be "yes."

However, observational cohort studies often do not report anything about power or sample sizes because the analyses are exploratory in nature. In this case, the answer would be "no." This is not a "fatal flaw." It just may indicate that attention was not paid to whether the study was sufficiently sized to answer a prespecified question–i.e., it may have been an exploratory, hypothesis-generating study.

Question 6. Exposure assessed prior to outcome measurement

This question is important because, in order to determine whether an exposure causes an outcome, the exposure must come before the outcome.

For some prospective cohort studies, the investigator enrolls the cohort and then determines the exposure status of various members of the cohort (large epidemiological studies like Framingham used this approach). However, for other cohort studies, the cohort is selected based on its exposure status, as in the example above of depressed diabetic men (the exposure being depression). Other examples include a cohort identified by its exposure to fluoridated drinking water and then compared to a cohort living in an area without fluoridated water, or a cohort of military personnel exposed to combat in the Gulf War compared to a cohort of military personnel not deployed in a combat zone.

With either of these types of cohort studies, the cohort is followed forward in time (i.e., prospectively) to assess the outcomes that occurred in the exposed members compared to nonexposed members of the cohort. Therefore, you begin the study in the present by looking at groups that were exposed (or not) to some biological or behavioral factor, intervention, etc., and then you follow them forward in time to examine outcomes. If a cohort study is conducted properly, the answer to this question should be "yes," since the exposure status of members of the cohort was determined at the beginning of the study before the outcomes occurred.

For retrospective cohort studies, the same principal applies. The difference is that, rather than identifying a cohort in the present and following them forward in time, the investigators go back in time (i.e., retrospectively) and select a cohort based on their exposure status in the past and then follow them forward to assess the outcomes that occurred in the exposed and nonexposed cohort members. Because in retrospective cohort studies the exposure and outcomes may have already occurred (it depends on how long they follow the cohort), it is important to make sure that the exposure preceded the outcome.

Sometimes cross-sectional studies are conducted (or cross-sectional analyses of cohort-study data), where the exposures and outcomes are measured during the same timeframe. As a result, cross-sectional analyses provide weaker evidence than regular cohort studies regarding a potential causal relationship between exposures and outcomes. For cross-sectional analyses, the answer to Question 6 should be "no."

Question 7. Sufficient timeframe to see an effect

Did the study allow enough time for a sufficient number of outcomes to occur or be observed, or enough time for an exposure to have a biological effect on an outcome? In the examples given above, if clinical depression has a biological effect on increasing risk for CVD, such an effect may take years. In the other example, if higher dietary sodium increases BP, a short timeframe may be sufficient to assess its association with BP, but a longer timeframe would be needed to examine its association with heart attacks.

The issue of timeframe is important to enable meaningful analysis of the relationships between exposures and outcomes to be conducted. This often requires at least several years, especially when looking at health outcomes, but it depends on the research question and outcomes being examined.

Cross-sectional analyses allow no time to see an effect, since the exposures and outcomes are assessed at the same time, so those would get a "no" response.

Question 8. Different levels of the exposure of interest

If the exposure can be defined as a range (examples: drug dosage, amount of physical activity, amount of sodium consumed), were multiple categories of that exposure assessed? (for example, for drugs: not on the medication, on a low dose, medium dose, high dose; for dietary sodium, higher than average U.S. consumption, lower than recommended consumption, between the two). Sometimes discrete categories of exposure are not used, but instead exposures are measured as continuous variables (for example, mg/day of dietary sodium or BP values).

In any case, studying different levels of exposure (where possible) enables investigators to assess trends or dose-response relationships between exposures and outcomes–e.g., the higher the exposure, the greater the rate of the health outcome. The presence of trends or dose-response relationships lends credibility to the hypothesis of causality between exposure and outcome.

For some exposures, however, this question may not be applicable (e.g., the exposure may be a dichotomous variable like living in a rural setting versus an urban setting, or vaccinated/not vaccinated with a one-time vaccine). If there are only two possible exposures (yes/no), then this question should be given an "NA," and it should not count negatively towards the quality rating.

Question 9. Exposure measures and assessment

Were the exposure measures defined in detail? Were the tools or methods used to measure exposure accurate and reliable–for example, have they been validated or are they objective? This issue is important as it influences confidence in the reported exposures. When exposures are measured with less accuracy or validity, it is

harder to see an association between exposure and outcome even if one exists. Also as important is whether the exposures were assessed in the same manner within groups and between groups; if not, bias may result.

For example, retrospective self-report of dietary salt intake is not as valid and reliable as prospectively using a standardized dietary log plus testing participants' urine for sodium content. Another example is measurement of BP, where there may be quite a difference between usual care, where clinicians measure BP however it is done in their practice setting (which can vary considerably), and use of trained BP assessors using standardized equipment (e.g., the same BP device which has been tested and calibrated) and a standardized protocol (e.g., patient is seated for 5 minutes with feet flat on the floor, BP is taken twice in each arm, and all four measurements are averaged). In each of these cases, the former would get a "no" and the latter a "yes."

Here is a final example that illustrates the point about why it is important to assess exposures consistently across all groups: If people with higher BP (exposed cohort) are seen by their providers more frequently than those without elevated BP (nonexposed group), it also increases the chances of detecting and documenting changes in health outcomes, including CVD-related events. Therefore, it may lead to the conclusion that higher BP leads to more CVD events. This may be true, but it could also be due to the fact that the subjects with higher BP were seen more often; thus, more CVD-related events were detected and documented simply because they had more encounters with the health care system. Thus, it could bias the results and lead to an erroneous conclusion.

Question 10. Repeated exposure assessment

Was the exposure for each person measured more than once during the course of the study period? Multiple measurements with the same result increase our confidence that the exposure status was correctly classified. Also, multiple measurements enable investigators to look at changes in exposure over time, for example, people who ate high dietary sodium throughout the followup period, compared to those who started out high then reduced their intake, compared to those who ate low sodium throughout. Once again, this may not be applicable in all cases. In many older studies, exposure was measured only at baseline. However, multiple exposure measurements do result in a stronger study design.

Question 11. Outcome measures

Were the outcomes defined in detail? Were the tools or methods for measuring outcomes accurate and reliable–for example, have they been validated or are they objective? This issue is important because it influences confidence in the validity of study results. Also important is whether the outcomes were assessed in the same manner within groups and between groups.

An example of an outcome measure that is objective, accurate, and reliable is death–the outcome measured with more accuracy than any other. But even with a measure as objective as death, there can be differences in the accuracy and reliability of how death was assessed by the investigators. Did they base it on an autopsy report, death certificate, death registry, or report from a family member? Another example is a study of whether dietary fat intake is related to blood cholesterol level (cholesterol level being the outcome), and the cholesterol level is measured from fasting blood samples that are all sent to the same laboratory. These examples would get a "yes." An example of a "no" would be self-report by subjects that they had a heart attack, or self-report of how much they weigh (if body weight is the outcome of interest).

Similar to the example in Question 9, results may be biased if one group (e.g., people with high BP) is seen more frequently than another group (people with normal BP) because more frequent encounters with the health care system increases the chances of outcomes being detected and documented.

Question 12. Blinding of outcome assessors

Blinding means that outcome assessors did not know whether the participant was exposed or unexposed. It is also sometimes called "masking." The objective is to look for evidence in the article that the person(s) assessing the outcome(s) for the study (for example, examining medical records to determine the outcomes that occurred in the exposed and comparison groups) is masked to the exposure status of the participant. Sometimes the person measuring the exposure is the same person conducting the outcome assessment. In this case, the outcome assessor would most likely not be blinded to exposure status because they also took measurements of exposures. If so, make a note of that in the comments section.

As you assess this criterion, think about whether it is likely that the person(s) doing the outcome assessment would know (or be able to figure out) the exposure status of the study participants. If the answer is no, then blinding is adequate. An example of adequate blinding of the outcome assessors is to create a separate committee, whose members were not involved in the care of the patient and had no information about the study participants' exposure status. The committee would then be provided with copies of participants' medical records, which had been stripped of any potential exposure information or personally identifiable information. The committee would then review the records for prespecified outcomes according to the study protocol. If blinding was not possible, which is sometimes the case, mark "NA" and explain the potential for bias.

Question 13. Followup rate

Higher overall followup rates are always better than lower followup rates, even though higher rates are expected in shorter studies, whereas lower overall followup rates are often seen in studies of longer duration. Usually, an acceptable overall followup rate is considered 80 percent or more of participants whose exposures were measured at baseline. However, this is just a general guideline. For example, a 6-month cohort study examining the relationship between dietary sodium intake and BP level may have over 90 percent followup, but a 20-year cohort study examining effects of sodium intake on stroke may have only a 65 percent followup rate.

Question 14. Statistical analyses

Were key potential confounding variables measured and adjusted for, such as by statistical adjustment for baseline differences? Logistic regression or other regression methods are often used to account for the influence of variables not of interest.

This is a key issue in cohort studies, because statistical analyses need to control for potential confounders, in contrast to an RCT, where the randomization process controls for potential confounders. All key factors that may be associated both with the exposure of interest and the outcome–that are not of interest to the research question–should be controlled for in the analyses.

For example, in a study of the relationship between cardiorespiratory fitness and CVD events (heart attacks and strokes), the study should control for age, BP, blood cholesterol, and body weight, because all of these factors are associated both with low fitness and with CVD events. Well-done cohort studies control for multiple potential confounders.

Some general guidance for determining the overall quality rating of observational cohort and cross-sectional studies

The questions on the form are designed to help you focus on the key concepts for evaluating the internal validity of a study. They are not intended to create a list that you simply tally up to arrive at a summary judgment of quality.

Internal validity for cohort studies is the extent to which the results reported in the study can truly be attributed to the exposure being evaluated and not to flaws in the design or conduct of the study–in other words, the ability of the study to draw associative conclusions about the effects of the exposures being studied on outcomes. Any such flaws can increase the risk of bias.

Critical appraisal involves considering the risk of potential for selection bias, information bias, measurement bias, or confounding (the mixture of exposures that one cannot tease out from each other). Examples of confounding include co-interventions, differences at baseline in patient characteristics, and other issues throughout the questions above. High risk of bias translates to a rating of poor quality. Low risk of bias translates to a rating of good quality. (Thus, the greater the risk of bias, the lower the quality rating of the study.)

In addition, the more attention in the study design to issues that can help determine whether there is a causal relationship between the exposure and outcome, the higher quality the study. These include exposures occurring prior to outcomes, evaluation of a dose-response gradient, accuracy of measurement of both exposure and outcome, sufficient timeframe to see an effect, and appropriate control for confounding–all concepts reflected in the tool.

Generally, when you evaluate a study, you will not see a "fatal flaw," but you will find some risk of bias. By focusing on the concepts underlying the questions in the quality assessment tool, you should ask yourself about the potential for bias in the study you are critically appraising. For any box where you check "no" you should ask, "What is the potential risk of bias resulting from this flaw in study design or execution?" That is, does this factor cause you to doubt the results that are reported in the study or doubt the ability of the study to accurately assess an association between exposure and outcome?

The best approach is to think about the questions in the tool and how each one tells you something about the potential for bias in a study. The more you familiarize yourself with the key concepts, the more comfortable you will be with critical appraisal. Examples of studies rated good, fair, and poor are useful, but each study must be assessed on its own based on the details that are reported and consideration of the concepts for minimizing bias.

Last Updated March 2014

QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES

COMPONENT RATINGS

A) SELECTION BIAS

(Q1) Are the individuals selected to participate in the study likely to be representative of the target population?

1 Very likely
2 Somewhat likely
3 Not likely
4 Can't tell

(Q2) What percentage of selected individuals agreed to participate?

1 80 - 100% agreement
2 60 – 79% agreement
3 less than 60% agreement
4 Not applicable
5 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

B) STUDY DESIGN

Indicate the study design

1 Randomized controlled trial
2 Controlled clinical trial
3 Cohort analytic (two group pre + post)
4 Case-control
5 Cohort (one group pre + post (before and after))
6 Interrupted time series
7 Other specify _____
8 Can't tell

Was the study described as randomized? If NO, go to Component C.

No Yes

If Yes, was the method of randomization described? (See dictionary)

No Yes

If Yes, was the method appropriate? (See dictionary)

No Yes

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

C) CONFOUNDERS

(Q1) Were there important differences between groups prior to the intervention?

1 Yes
2 No
3 Can't tell

The following are examples of confounders:

1 Race
2 Sex
3 Marital status/family
4 Age
5 SES (income or class)
6 Education
7 Health status
8 Pre-intervention score on outcome measure

(Q2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design (e.g. stratification, matching) or analysis)?

1 80 – 100% (most)
2 60 – 79% (some)
3 Less than 60% (few or none)
4 Can't Tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

D) BLINDING

(Q1) Was (were) the outcome assessor(s) aware of the intervention or exposure status of participants?

1 Yes
2 No
3 Can't tell

(Q2) Were the study participants aware of the research question?

1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

E) DATA COLLECTION METHODS

(Q1) Were data collection tools shown to be valid?

1 Yes
2 No
3 Can't tell

(Q2) Were data collection tools shown to be reliable?

1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

F) WITHDRAWALS AND DROP-OUTS

(Q1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?

1 Yes
2 No
3 Can't tell
4 Not Applicable (i.e. one time surveys or interviews)

(Q2) Indicate the percentage of participants completing the study. (If the percentage differs by groups, record the lowest).

1 80 -100%
2 60 - 79%
3 less than 60%
4 Can't tell
5 Not Applicable (i.e. Retrospective case-control)

RATE THIS SECTION	STRONG	MODERATE	WEAK	
See dictionary	1	2	3	Not Applicable

G) INTERVENTION INTEGRITY

(Q1) What percentage of participants received the allocated intervention or exposure of interest?

1 80 -100%
2 60 - 79%
3 less than 60%
4 Can't tell

(Q2) Was the consistency of the intervention measured?

1 Yes
2 No
3 Can't tell

(Q3) Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results?

4 Yes
5 No
6 Can't tell

H) ANALYSES

(Q1) Indicate the unit of allocation (circle one)

community organization/institution practice/office individual

(Q2) Indicate the unit of analysis (circle one)

community organization/institution practice/office individual

(Q3) Are the statistical methods appropriate for the study design?

1 Yes
2 No
3 Can't tell

(Q4) Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?

1 Yes
2 No
3 Can't tell

GLOBAL RATING

COMPONENT RATINGS

Please transcribe the information from the gray boxes on pages 1-4 onto this page. See dictionary on how to rate this section.

A	**SELECTION BIAS**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
B	**STUDY DESIGN**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
C	**CONFOUNDERS**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
D	**BLINDING**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
E	**DATA COLLECTION METHOD**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
F	**WITHDRAWALS AND DROPOUTS**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	Not Applicable

GLOBAL RATING FOR THIS PAPER (circle one):

1	STRONG	(no WEAK ratings)
2	MODERATE	(one WEAK rating)
3	WEAK	(two or more WEAK ratings)

With both reviewers discussing the ratings:

Is there a discrepancy between the two reviewers with respect to the component (A-F) ratings?

No Yes

If yes, indicate the reason for the discrepancy

1	Oversight
2	Differences in interpretation of criteria
3	Differences in interpretation of study

Final decision of both reviewers (circle one):

1	**STRONG**
2	**MODERATE**
3	**WEAK**

Reporting checklist for protocol of a systematic review.

Based on the PRISMA-P guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the PRISMA-P reporting guidelines, and cite them as:

Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA. Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 statement. Syst Rev. 2015;4(1):1.

		Reporting Item	Page Number
Identification	#1a	Identify the report as a protocol of a systematic review	1
Update	#1b	If the protocol is for an update of a previous systematic review, identify as such	n/a
	#2	If registered, provide the name of the registry (such as PROSPERO) and registration number	2
Contact	#3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author	1
Contribution	#3b	Describe contributions of protocol authors and identify the guarantor of the review	8
	#4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important	7

		protocol amendments	
Sources	#5a	Indicate sources of financial or other support for the review	8
Sponsor	#5b	Provide name for the review funder and / or sponsor	n/a
Role of sponsor or funder	#5c	Describe roles of funder(s), sponsor(s), and / or institution(s), if any, in developing the protocol	n/a
Rationale	#6	Describe the rationale for the review in the context of what is already known	3
Objectives	#7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	4
Eligibility criteria	#8	Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review	4,5
Information sources	#9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage	5
Search strategy	#10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	5
Study records - data management	#11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	6
Study records - selection process	#11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis)	6
Study records - data collection process	#11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	6
Data items	#12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications	6

Outcomes and prioritization	#13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	5
Risk of bias in individual studies	#14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	6
Data synthesis	#15a	Describe criteria under which study data will be quantitatively synthesised	7
	#15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I2, Kendall's τ)	7
	#15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)	7
	#15d	If quantitative synthesis is not appropriate, describe the type of summary planned	7
Meta-bias(es)	#16	Specify any planned assessment of meta-bias(es) (such as publication bias across studies, selective reporting within studies)	7
Confidence in cumulative evidence	#17	Describe how the strength of the body of evidence will be assessed (such as GRADE)	6

BMJ Open

The effect of in-utero exposure to HIV and antiretroviral drugs on growth in HIV-exposed uninfected children: a systematic review and meta-analysis protocol

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-023937.R1
Article Type:	Protocol
Date Submitted by the Author:	11-Nov-2018
Complete List of Authors:	Ekali, Gabriel; Universite de Yaounde I, Biotechnology Center; National AIDS Control Committee, Jesson, Julie; Universite Toulouse III Paul Sabatier, Inserm UMR 1027 Enok, Pascal; University of Montreal, Department of Preventive Medicine Leroy, Valériane; Universite Toulouse III Paul Sabatier
Primary Subject Heading:	HIV/AIDS
Secondary Subject Heading:	Paediatrics
Keywords:	HIV & AIDS < INFECTIOUS DISEASES, Nutrition < TROPICAL MEDICINE, PAEDIATRICS, Growth, Wasting, Stunting

The effect of *in-utero* exposure to HIV and antiretroviral drugs on growth in HIV-exposed uninfected children: a systematic review and meta-analysis protocol

Gabriel Loni Ekali[1,2,3,*], Julie Jesson[2], Enok Pascal[4], Valériane Leroy[2]

1.Biotechnology Center, University of Yaounde 1, Cameroon.

2. Inserm UMR 1027, Université Toulouse 3, Toulouse, France.

3. National AIDS Control Committee, Ministry of Public Health, Cameroon.

4.Department of Preventive Medicine, University of Montreal.

***Corresponding author**: Dr Gabriel Loni Ekali

Biotechnology Center, University of Yaounde 1, Cameroon; National AIDS Control Committee, Ministry of Public Health, Cameroon.

Email: eloni2000@yahoo.com

Telephone: 237 - 677333023

PO Box: 3077, Yaounde, Cameroon.

Word count: 2,396

ABSTRACT

Introduction

HIV-exposed uninfected children (HEU) have higher morbidity and mortality compared to HIV unexposed uninfected children. Despite the fact that malnutrition contributes to about half of all infant deaths below 5 years of age in low and middle income countries and that growth impairment has been reported in the HEU population, the spectrum of growth disorders associated to HIV and ART exposure during the *in utero* and perinatal periods is yet to comprehensively summarized amongst the global HEU population. This protocol for a systematic review and meta-analysis aims to critically synthesize data concerning the prevalence of underweight, stunting and wasting at different ages in the global HEU population.

Methods and analysis

Medline, EMBASE, Cochrane Library, TOXLINE, WHO Global Index Medicus and the Web of Science will be searched for relevant articles published between January 1st 1989 and December 1st 2017 without language restriction. In addition, conference abstracts and reference lists of eligible papers and relevant review articles will be screened. Authors will screen and select studies, extract data, assess risk of bias as well as studies individually for heterogeneity. Study-specific estimates will be pooled through a random-effects meta-analysis model for studies that are clinically homogenous while funnel plots and Egger's test will be used to detect publication bias. Results will be presented by ART availability period, country income levels, mode of breastfeeding.

Ethics and Dissemination

Ethical approval will not be required for this study because it will be based on published data. The final report of this study will be published in a peer-reviewed journal and presented at scientific conferences. This review will summarize the evidence and quantify the problem of growth impairment in HEU infants and so shed more light on our understanding of the higher morbidity and mortality in this growing population.

PROSPERO registration number: CRD42018091762

Key words: Growth, wasting, stunting, underweight, HIV-exposed uninfected children

STRENGTHS AND LIMITATIONS OF THE STUDY

- One of the limitations of the current review could be insufficient data to quantify the growth outcomes of interest.
- To our knowledge, this will be the first study to summarize the effect of HIV and ART exposure on growth of HIV-exposed uninfected infants compared to an appropriate control group of HIV-unexposed uninfected children.
- This review will cover both the era before widespread combination ART and the era after widespread access to combination ART so as to investigate whether any differences in growth outcomes exists between these two eras.

BACKGROUND: One of the major strategies of reducing new pediatric Human Immunodeficiency Virus (HIV) infections is by preventing mother-to-child transmission (MTCT) of the virus. This transmission can occur in utero, during delivery or breastfeeding.[1] In the absence of any intervention MTCT of HIV can be as high as 45%.[2] However, with antiretroviral therapy (ART), the rate of MTCT can be reduced to less than 2%.[3] With widespread access to ART, the UNAIDS report that new pediatric HIV infections have declined by 47% between 2010 and 2016, while coverage of antiretroviral medicines provided to pregnant women living with HIV to prevent transmission to their children rose from 47% to 76% during the same period.[4]

As MTCT rates drop, the proportion of children who are exposed in utero to HIV and ART and are born uninfected with the virus (HEU) is increasing worldwide, especially in sub-Saharan Africa where the greatest burden of the HIV pandemic is found. Indeed, HEU children constitute about 30% of the Southern African pediatric population.[5] Even though HEU children are not infected with HIV, a growing body of literature shows evidence of higher morbidity and mortality compared to HIV/ART unexposed uninfected children (HUU). Higher frequency of hospitalisations, more severe respiratory tract infections especially from encapsulated bacteria as well as diarrheal diseases have been described.[5-6] Children exposed to maternal HIV and ART have modest but significant impairment of development and a higher risk of growth impairment.[7] Growth failure, reflected by restricted linear growth and weight gain, have been reported amongst HEU children.A number of recent reviews have dwelt on summarizing the available evidence of higher morbidity, mortality, immune dysfunction and its possible role in the higher susceptibility of HEU to diseases.[5-7,8-9] However, the spectrum of growth disorders found in this growing population and its possible association and role to the higher susceptibility to infectious diseases they experience is yet to be documented. This is important as it is known that impaired growth clinically represents an altered nutritional status, which is significantly associated to mortality in infancy. Indeed, nutrition-related factors contribute to about 45% of deaths in children under 5 years of age in low and middle income countries, translating into the loss of approximately 3 million young lives annually.[10] Furthermore, restricted linear growth in

early life also affects the quality of life an individual as it is associated to cognitive impairment [11] and lower lifetime earnings [12]. In addition, appropriate comparisons of HEU and HUU infants were rare before the antiretroviral era. With increasing ART access, it is methodologically challenging to disentangle effects of HIV exposure from those associated with ART exposure, particularly in regions where malnutrition and co-morbidities are common and could increase the potential effects of these exposures.

In view of the recent reports of altered linear growth and weight gain among HEU infants, the main goal of this paper is to review and summarize evidence on the effects of perinatal/postnatal exposure to HIV and ART on growth outcomes in HEU children until 60 months of age and in later life.

Review question

What is the effect of in utero/postnatal exposure to HIV and ART on growth outcomes (stunting, wasting, underweight) until 60 months of age in HEU compared to HUU children as documented in studies reported between January 1st, 1989 to December 1st, 2017.

Specific objectives

To answer this review question, a systematic review and meta-analysis of the existing evidence will be conducted, the objectives of which will be to determine:

1. The effect of HIV and ART exposure on underweight (weight-for-age < 2 SD) until 60 months of age in the global HEU population compared to HUU children.
2. The effect of HIV and ART exposure on stunting (height/length-for-age < 2 SD) until 60 months of age in the global HEU population compared to HUU children.
3. The effect of HIV and ART exposure on wasting (weight-for-height/length < 2 SD) until 60 months of age in the global HEU population compared to HUU children.

METHODS AND ANALYSIS

Design

This will be a systematic review and meta-analysis of the published literature. The Centre for Reviews and Dissemination guidelines [13] will be used for the methodology of this review and the review will be registered in the PROSPERO International Prospective Register of systematic reviews.

Inclusion criteria:

1. Type of studies: all observational cohort studies and randomized controlled trials that have data on any growth parameter of interest, either as main or secondary outcomes.

Only comparative studies with HEU and HUU children will be included in the meta-analysis.

2. Type of participants: adequately defined HEU (children exposed in utero to HIV and combination ART of any class, who are declared HIV negative by either DNA PCR between 6 weeks and 18 months OR appropriate HIV serological tests above 18 months) and HUU children (children born of mothers with documented HIV negative serological test), born at term (≥37 weeks) aged 0 to 5 years. All ART exposure will be considered.
3. Outcome measures: prevalence of, stunting, wasting, underweight; growth velocity or studies that have enough data to compute the outcomes of interest at yearly time intervals, available from 12 to 60 months
4. Type and period of publications: published studies and conference abstracts from 1 January 1989 to 1 December 2017.

For studies with several publications of their findings over time, we will retain in the review the one that has the best quality and the most appropriate regarding our objective. For the meta-analysis, studies with a documented measure of effect of in utero/postnatal HIV and ART exposure on growth comparing HEU and HUU groups will be included. For studies without reported measures of effect, these will be computed if the provided data is adequate.

Exclusion criteria:

We will exclude:

1. Type of studies: unpublished manuscripts, case reviews, policy reports, commentaries and editorials.
2. Studies whose data will not be sufficient to calculate the appropriate measures of effect.

Outcome measures: The following measures will be considered at different age points (6, 9,12,18,24, 36,48,60 months)

- height-for-age
- weight-for-height/length
- weight-for-age
- growth velocity
- prevalence of stunting, wasting and underweight at each age point.

Search strategy to identify relevant studies

Databases: An exhaustive literature search will be carried out in Medline, EMBASE, Cochrane Library, TOXLINE, WHO Global Index Medicus and the Web of Science to identify relevant articles published on growth in HIV-exposed uninfected infants between January 1st 1989 and

December 1st 2017. The proceedings of the following conferences will also be searched: International AIDS Society Conference on HIV Pathogenesis, Treatment and Prevention; International AIDS Conference; Conference of Retroviruses and Opportunistic Infections. The search strategy that will be used for online databases is presented in table 1.

Table 1: Search strategy for the review of impact of in utero exposure to HIV and ART on growth in HIV-exposed uninfected infants

Search	Search Terms
#1	HIV OR HIV-1 OR HIV-2 OR HIV1 OR HIV2 OR HIV infect* OR human immune*
#2	HIV-exposed uninfected OR HIV exposed uninfected OR HEU OR HIV-EU OR HIV-unexposed OR HIV unexposed OR HUU OR HU OR HIV-UE OR HIV-uninfected OR HIV uninfected
#3	Children OR child OR infant OR infants OR pediatric OR pediatrics OR paediatric OR paediatrics
#4	Growth OR development OR stunting OR malnutrition OR wasting OR underweight OR weight OR length OR height OR BMI OR head circumference OR growth faltering OR Growth velocity
#5	Limits: from 1 January 1989 to 1 December 2017
#6	#1 AND #2 AND #3 AND #4 AND #5

Searching other sources

The list of references of eligible articles as well as relevant review articles will also be manually searched.

Selecting studies for inclusion in the review

All articles returned by the search will be saved to the Endnote version X4 software which will be used to remove duplicates. The titles and abstracts of the articles remaining after exclusion of duplicates will be reviewed to identify potentially relevant papers according to the inclusion and exclusion criteria. Two investigators will independently review retrieved full text of all identified papers for eligibility and then retain those that are eligible by consensus. Any disagreement between investigators will be settled by a third reviewer. A PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram will be used to detail the number of articles identified, screened, included and excluded (online supplementary file 1).

Data extraction and management

A tool for data abstraction will be developed and the following data will be extracted:

1. Details about the author: Name, Journal
2. Characteristics of the study: design, country, period when study was conducted, year of publication, duration of follow-up, sample size, study endpoints,
3. Characteristics of participants: Age, sex, class of maternal ART, term of delivery, low birth-weight, mode of infant feeding (no BF, mixed feeding, exclusive BF), duration of breastfeeding (BF), Cotrimoxazole prophylaxis, growth references used,
4. Outcome measures at different time points: weight-for-age, length-for-age, weight-for-length, BMI for age and growth velocity.

Two investigators will independently extract data and any disagreements will be reconciled through discussion or by a third investigator as necessary.

Assessing methodological quality of included studies and risk of bias

The quality assessment tool for Observational Cohort and Cross-sectional studies of the National Heart, Lung and Blood Institute (NHLBI) will be used to assess the quality of observational cohort studies [14] (online supplementary file 2) while for intervention studies, the Effective Public Health Practice (EPHPP) tool [15] (online supplementary file 3) will be used by two investigators independently. Agreement between the two investigators will be measured by Cohen's kappa statistic. For the NHLBI tool, the score will be graded as Good, Fair and Poor while for the EPHPP tool, it will be graded as Strong, Moderate and Weak.

Data synthesis and analysis

Prevalences of underweight, stunting and wasting in the HEU and HUU groups will be recalculated based on the information provided by individual studies and odd ratios will be computed at each age point. The exposure variable will be HIV and combination ART exposure in utero and the outcome variable will be the different measures of growth: length/height-for-age, weight-for-height, then weight-for-age. We will use the statistic $I^2 = 100*(Q-df)/Q$ Q (with Q being the statistic of the Cochran heterogeneity Q-test and df the number of degrees of freedom corresponding to the number of studies minus one), which makes it possible to capture the proportion of the total variance observed due to a real difference in the measures of effects between the studies, in order to explore heterogeneity. Unlike the Q-test, this statistic is not influenced by the number of studies. Moreover, a non-significant Q-test does not necessarily imply that there is no heterogeneity between the studies because the power of the test depends on the number of studies. For statistical I^2, the values of 0%, 25%, 50% and 75% respectively represent the following heterogeneity levels: absence, weak, moderate and high. For each risk factor, the combined measure of relative risk will be estimated from a random-effect model if

$I^2 \geq 25\%$ and a fixed-effect model if $I^2 < 25\%$. If substantial heterogeneity is observed, a meta-regression will be used to identify the characteristics of the studies (e.g study quality, study location, sample size, adjusting of confounders or not, antiretroviral therapy agents and regimens, duration and mode of breastfeeding, maternal socioeconomic status, etc.) that may explain the observed heterogeneity. All data will be analyzed using Stata V.14 (StataCorp, Texas, USA). A forest plot will be presented for each factor studied in the studies found. Qualitative synthesis will be used in cases where data extracted is insufficient to perform quantitative synthesis or when the included studies significantly differ in design, setting and outcome measures.

Assessment of reporting bias

To explore a potential publication bias, funnel plots and Egger's test will be done and a p value below 0.10 will be considered indicative of significant publication bias.

Presentation and reporting of the results

The results of this meta-analysis will be presented by HIV and ART exposure status, by pre-ART and post-ART era according to the different classes of ART, by infant feeding mode, and if feasible by country income level. We will prioritize length/height-for-age, weight-for-height outcomes report first, weight-for-age being less discriminant. [16]

As NCHS/CDC/WHO [17] growth standards were used before 2005, and WHO child growth standards were used after 2005 [18], this may affect the longitudinal comparisons of the outcomes. Therefore, we will stratify our results according to the different standards. We will carry out sensitivity analysis to determine whether low birth weight has an impact on our results. Adherence to ART will be documented if reported by authors of the selected papers and discussed in the qualitative synthesis section of our manuscript.

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocol (PRISMA-P) will be used for reporting this present protocol (online supplementary file 4).[19] Results of the search method will be presented with a flow chart, detailing the selection process. Causes of exclusion for studies firstly identified as eligible will be documented.

Patient and Public Involvement

Patients were not involved in the design of this study protocol.

POTENTIAL PROTOCOL AMENDMENTS

The current protocol as written will not be amended in the course of the study. This is to avoid any outcome reporting bias.

ETHICAL CONSIDERATIONS AND DISSEMINATION

In this study, only published data will be used. No ethical clearance will therefore be needed. The final results will be published in peer-reviewed journals and also presented at conferences.

CONCLUSION

A comprehensive summary of the evidence concerning how in utero/postnatal exposure to HIV and ART affect growth is critical to shed more light on our understanding of why HEU children have higher morbidity and mortality to certain infectious diseases compared to HUU. It is hoped that the results of this review will quantify the problem of growth impairment and draw attention to the necessity to possibly develop appropriate nutritional interventions for this growing population.

AUTHOR STATEMENT: EG, LV had the idea and EG designed the protocol. LV, JJ, EP critically reviewed intellectual content and revised the methodology of the study. EG and LV are guarantors of the study and all the authors approved the final version of the manuscript.

FUNDING: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

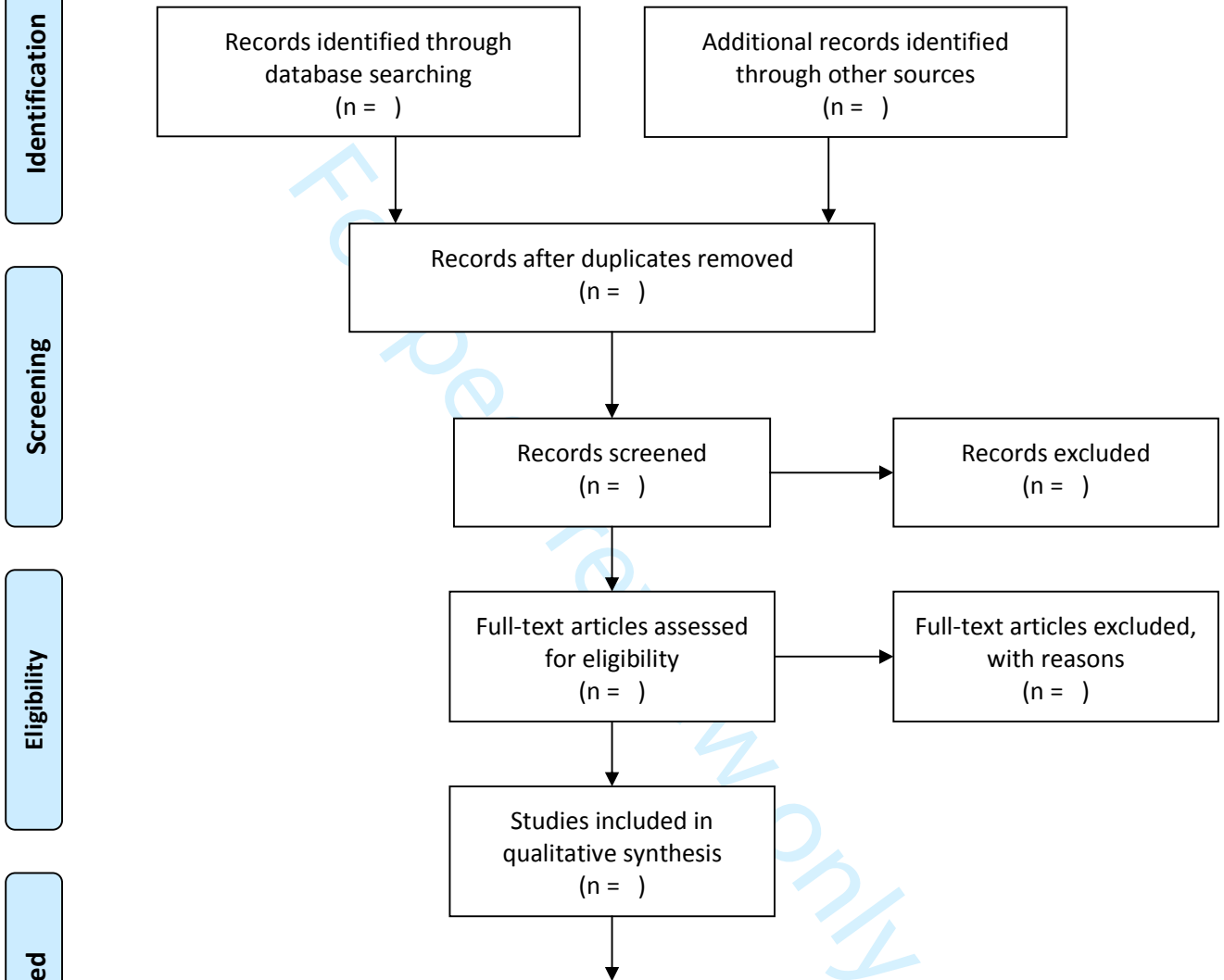
COMPETING INTEREST: None declared

REFERENCES

1. Mofenson LM. Mother-child HIV-1 transmission: Timing and determinants. *Obstet Gynecol Clin North Am* 1997; 24:759.
2. Lehman DA, Farquhar C. Biological mechanisms of vertical human immunodeficiency virus (HIV-1) transmission. *Rev Med Virol*. 2007;17:381–403
3. Brocklehurst P: Interventions for reducing the risk of mother-to-child transmission of HIV Infection. *Cochrane Database Syst Rev* 2002;1:CD000102
4. UNAIDS data 2017. Available at http://www.unaids.org/sites/default/files/media_asset/20170720_Data_book_2017_en.pdf . Accessed on January 29th 2018 at 00h24minutes.
5. Slogrove AL, Goetghebuer T, Cotton MF et al. Pattern of Infectious Morbidity in HIV-Exposed Uninfected Infants and Children. *Front Immunol* 2016 May 6;7:164
6. Brennan AT, Bonawitz R, Gill CJ, et al. A meta-analysis assessing all-cause mortality in HIV-exposed uninfected compared with HIV-unexposed uninfected infants and children. *AIDS* 2016 Sep 24;30(15):2351-60
7. Desmonde S, Goetghebuer T, Thorne C, et al. Health and survival of HIV perinatally exposed but uninfected children born to HIV-infected mothers. *Curr Opin HIV AIDS* 2016 Sep;11(5):465-476
8. Abu-Raya B, Kollmann TR, Marchant A, et al. The Immune System of HIV-Exposed Uninfected Infants. *Front Immunol* 2016 Sep 28;7:383
9. Ruck C, Reikie BA, Marchant A, et al. Linking Susceptibility to Infectious Diseases to Immune System Abnormalities among HIV-Exposed Uninfected Infants. *Front Immunol* 2016 Aug 19;7:310

10. WHO Media Center. Children : reducing child mortality Fact sheet. Available from http://www.who.int/mediacentre/factsheets/fs178/en/. Accessed on February 2nd 2018 at 23h56.
11. Sudfeld CR, McCoy DC, Danaie G, et al. Linear growth and child development and low- and middle-income countries: a meta-analysis. Pediatrics. 2015;135:e1266–e1275.
12. Adair LS, Fall CH, Osmond C, et al. Associations of linear growth and relative weight gain during early life with adult health and human capital in countries of low and middle income: findings five birth cohort studies. Lancet. 2013;382:525–534
13. Centers for Reviews and Dissemination. CRD's guidance for undertaking reviews in healthcare. Centers for Reviews and Dissemination; 2009. Available from: http://www.york.ac.uk/crd/SysRev/!SSL!/WebHelp/SysRev3.htm. Accessed February 2nd 2018 at 01h04
14. National Heart, Lung and Blood Institute. Quality assessment tool for Observational Cohort and Cross-sectional studies. Available from https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools and accessed February 2nd 2018 at 01h06.
15. Effective Public Health Practice Project. Quality assessment tool for quantitative studies. Available at http://www.ephpp.ca/PDF/Quality%20Assessment%20Tool_2010_2.pdf . Accessed on February 2nd 2018 at 01h07.
16. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.doi:10.1186/2046-4053-4-1
17. Geneva: WHO; 1997. Global database on child growth and malnutrition. Available at http://apps.who.int/iris/bitstream/handle/10665/63750/WHO_NUT_97.4.pdf;jsessionid=3467A6BA2F61719518EEF6309CE1ACD7?sequence=1. Accessed on September 24th 2019 at 3:43am WAT.
18. Waterlow JC, Buzina R, Keller W, Lane JM, Nichaman MZ, Tanner JM. The presentation and use of height and weight data for comparing the nutritional status of groups of children under the age of 10 years. Bull World Health Organ. 1977;55(4):489-98
19. Geneva: WHO; 2007. WHO Anthro for personal computers, version Software for assessing growth and development of the world's children. http://www.who.int/childgrowth/software/en/ [Ref list]

PRISMA 2009 Flow Diagram



From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit www.prisma-statement.org.

National Heart, Lung, and Blood Institute

Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies

Criteria	Yes	No	Other (CD, NR, NA)*
1. Was the research question or objective in this paper clearly stated?			
2. Was the study population clearly specified and defined?			
3. Was the participation rate of eligible persons at least 50%?			
4. Were all the subjects selected or recruited from the same or similar populations (including the same time period)? Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?			
5. Was a sample size justification, power description, or variance and effect estimates provided?			
6. For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?			
7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?			
8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?			
9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?			
10. Was the exposure(s) assessed more than once over time?			
11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?			
12. Were the outcome assessors blinded to the exposure status of participants?			
13. Was loss to follow-up after baseline 20% or less?			
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?			

Quality Rating (Good, Fair, or Poor) (see guidance)
Rater #1 initials:
Rater #2 initials:
Additional Comments (If POOR, please state why):

*CD, cannot determine; NA, not applicable; NR, not reported

Guidance for Assessing the Quality of Observational Cohort and Cross-Sectional Studies

The guidance document below is organized by question number from the tool for quality assessment of observational cohort and cross-sectional studies.

Question 1. Research question

Did the authors describe their goal in conducting this research? Is it easy to understand what they were looking to find? This issue is important for any scientific paper of any type. Higher quality scientific research explicitly defines a research question.

Questions 2 and 3. Study population

Did the authors describe the group of people from which the study participants were selected or recruited, using demographics, location, and time period? If you were to conduct this study again, would you know who to recruit, from where, and from what time period? Is the cohort population free of the outcomes of interest at the time they were recruited?

An example would be men over 40 years old with type 2 diabetes who began seeking medical care at Phoenix Good Samaritan Hospital between January 1, 1990 and December 31, 1994. In this example, the population is clearly described as: (1) who (men over 40 years old with type 2 diabetes); (2) where (Phoenix Good Samaritan Hospital); and (3) when (between January 1, 1990 and December 31, 1994). Another example is women ages 34 to 59 years of age in 1980 who were in the nursing profession and had no known coronary disease, stroke, cancer, hypercholesterolemia, or diabetes, and were recruited from the 11 most populous States, with contact information obtained from State nursing boards.

In cohort studies, it is crucial that the population at baseline is free of the outcome of interest. For example, the nurses' population above would be an appropriate group in which to study incident coronary disease. This information is usually found either in descriptions of population recruitment, definitions of variables, or inclusion/exclusion criteria.

You may need to look at prior papers on methods in order to make the assessment for this question. Those papers are usually in the reference list.

If fewer than 50% of eligible persons participated in the study, then there is concern that the study population does not adequately represent the target population. This increases the risk of bias.

Question 4. Groups recruited from the same population and uniform eligibility criteria

Were the inclusion and exclusion criteria developed prior to recruitment or selection of the study population? Were the same underlying criteria used for all of the subjects involved? This issue is related to the description of the study population, above, and you may find the information for both of these questions in the same section of the paper.

Most cohort studies begin with the selection of the cohort; participants in this cohort are then measured or evaluated to determine their exposure status. However, some cohort studies may recruit or select exposed participants in a different time or place than unexposed participants, especially retrospective cohort studies–which is when data are obtained from the past (retrospectively), but the analysis examines exposures prior to outcomes. For example, one research question could be whether diabetic men with clinical depression are at higher risk for cardiovascular disease than those without clinical depression. So, diabetic men with depression might be selected from a mental health clinic, while diabetic men without depression might be selected from an internal medicine or endocrinology clinic. This study recruits groups from different clinic populations, so this example would get a "no."

However, the women nurses described in the question above were selected based on the same inclusion/exclusion criteria, so that example would get a "yes."

Question 5. Sample size justification

Did the authors present their reasons for selecting or recruiting the number of people included or analyzed? Do they note or discuss the statistical power of the study? This question is about whether or not the study had enough participants to detect an association if one truly existed.

A paragraph in the methods section of the article may explain the sample size needed to detect a hypothesized difference in outcomes. You may also find a discussion of power in the discussion section (such as the study had 85 percent power to detect a 20 percent increase in the rate of an outcome of interest, with a 2-sided alpha of 0.05). Sometimes estimates of variance and/or estimates of effect size are given, instead of sample size calculations. In any of these cases, the answer would be "yes."

However, observational cohort studies often do not report anything about power or sample sizes because the analyses are exploratory in nature. In this case, the answer would be "no." This is not a "fatal flaw." It just may indicate that attention was not paid to whether the study was sufficiently sized to answer a prespecified question–i.e., it may have been an exploratory, hypothesis-generating study.

Question 6. Exposure assessed prior to outcome measurement

This question is important because, in order to determine whether an exposure causes an outcome, the exposure must come before the outcome.

For some prospective cohort studies, the investigator enrolls the cohort and then determines the exposure status of various members of the cohort (large epidemiological studies like Framingham used this approach). However, for other cohort studies, the cohort is selected based on its exposure status, as in the example above of depressed diabetic men (the exposure being depression). Other examples include a cohort identified by its exposure to fluoridated drinking water and then compared to a cohort living in an area without fluoridated water, or a cohort of military personnel exposed to combat in the Gulf War compared to a cohort of military personnel not deployed in a combat zone.

With either of these types of cohort studies, the cohort is followed forward in time (i.e., prospectively) to assess the outcomes that occurred in the exposed members compared to nonexposed members of the cohort. Therefore, you begin the study in the present by looking at groups that were exposed (or not) to some biological or behavioral factor, intervention, etc., and then you follow them forward in time to examine outcomes. If a cohort study is conducted properly, the answer to this question should be "yes," since the exposure status of members of the cohort was determined at the beginning of the study before the outcomes occurred.

For retrospective cohort studies, the same principal applies. The difference is that, rather than identifying a cohort in the present and following them forward in time, the investigators go back in time (i.e., retrospectively) and select a cohort based on their exposure status in the past and then follow them forward to assess the outcomes that occurred in the exposed and nonexposed cohort members. Because in retrospective cohort studies the exposure and outcomes may have already occurred (it depends on how long they follow the cohort), it is important to make sure that the exposure preceded the outcome.

Sometimes cross-sectional studies are conducted (or cross-sectional analyses of cohort-study data), where the exposures and outcomes are measured during the same timeframe. As a result, cross-sectional analyses provide weaker evidence than regular cohort studies regarding a potential causal relationship between exposures and outcomes. For cross-sectional analyses, the answer to Question 6 should be "no."

Question 7. Sufficient timeframe to see an effect

Did the study allow enough time for a sufficient number of outcomes to occur or be observed, or enough time for an exposure to have a biological effect on an outcome? In the examples given above, if clinical depression has a biological effect on increasing risk for CVD, such an effect may take years. In the other example, if higher dietary sodium increases BP, a short timeframe may be sufficient to assess its association with BP, but a longer timeframe would be needed to examine its association with heart attacks.

The issue of timeframe is important to enable meaningful analysis of the relationships between exposures and outcomes to be conducted. This often requires at least several years, especially when looking at health outcomes, but it depends on the research question and outcomes being examined.

Cross-sectional analyses allow no time to see an effect, since the exposures and outcomes are assessed at the same time, so those would get a "no" response.

Question 8. Different levels of the exposure of interest

If the exposure can be defined as a range (examples: drug dosage, amount of physical activity, amount of sodium consumed), were multiple categories of that exposure assessed? (for example, for drugs: not on the medication, on a low dose, medium dose, high dose; for dietary sodium, higher than average U.S. consumption, lower than recommended consumption, between the two). Sometimes discrete categories of exposure are not used, but instead exposures are measured as continuous variables (for example, mg/day of dietary sodium or BP values).

In any case, studying different levels of exposure (where possible) enables investigators to assess trends or dose-response relationships between exposures and outcomes–e.g., the higher the exposure, the greater the rate of the health outcome. The presence of trends or dose-response relationships lends credibility to the hypothesis of causality between exposure and outcome.

For some exposures, however, this question may not be applicable (e.g., the exposure may be a dichotomous variable like living in a rural setting versus an urban setting, or vaccinated/not vaccinated with a one-time vaccine). If there are only two possible exposures (yes/no), then this question should be given an "NA," and it should not count negatively towards the quality rating.

Question 9. Exposure measures and assessment

Were the exposure measures defined in detail? Were the tools or methods used to measure exposure accurate and reliable–for example, have they been validated or are they objective? This issue is important as it influences confidence in the reported exposures. When exposures are measured with less accuracy or validity, it is

harder to see an association between exposure and outcome even if one exists. Also as important is whether the exposures were assessed in the same manner within groups and between groups; if not, bias may result.

For example, retrospective self-report of dietary salt intake is not as valid and reliable as prospectively using a standardized dietary log plus testing participants' urine for sodium content. Another example is measurement of BP, where there may be quite a difference between usual care, where clinicians measure BP however it is done in their practice setting (which can vary considerably), and use of trained BP assessors using standardized equipment (e.g., the same BP device which has been tested and calibrated) and a standardized protocol (e.g., patient is seated for 5 minutes with feet flat on the floor, BP is taken twice in each arm, and all four measurements are averaged). In each of these cases, the former would get a "no" and the latter a "yes."

Here is a final example that illustrates the point about why it is important to assess exposures consistently across all groups: If people with higher BP (exposed cohort) are seen by their providers more frequently than those without elevated BP (nonexposed group), it also increases the chances of detecting and documenting changes in health outcomes, including CVD-related events. Therefore, it may lead to the conclusion that higher BP leads to more CVD events. This may be true, but it could also be due to the fact that the subjects with higher BP were seen more often; thus, more CVD-related events were detected and documented simply because they had more encounters with the health care system. Thus, it could bias the results and lead to an erroneous conclusion.

Question 10. Repeated exposure assessment

Was the exposure for each person measured more than once during the course of the study period? Multiple measurements with the same result increase our confidence that the exposure status was correctly classified. Also, multiple measurements enable investigators to look at changes in exposure over time, for example, people who ate high dietary sodium throughout the followup period, compared to those who started out high then reduced their intake, compared to those who ate low sodium throughout. Once again, this may not be applicable in all cases. In many older studies, exposure was measured only at baseline. However, multiple exposure measurements do result in a stronger study design.

Question 11. Outcome measures

Were the outcomes defined in detail? Were the tools or methods for measuring outcomes accurate and reliable–for example, have they been validated or are they objective? This issue is important because it influences confidence in the validity of study results. Also important is whether the outcomes were assessed in the same manner within groups and between groups.

An example of an outcome measure that is objective, accurate, and reliable is death–the outcome measured with more accuracy than any other. But even with a measure as objective as death, there can be differences in the accuracy and reliability of how death was assessed by the investigators. Did they base it on an autopsy report, death certificate, death registry, or report from a family member? Another example is a study of whether dietary fat intake is related to blood cholesterol level (cholesterol level being the outcome), and the cholesterol level is measured from fasting blood samples that are all sent to the same laboratory. These examples would get a "yes." An example of a "no" would be self-report by subjects that they had a heart attack, or self-report of how much they weigh (if body weight is the outcome of interest).

Similar to the example in Question 9, results may be biased if one group (e.g., people with high BP) is seen more frequently than another group (people with normal BP) because more frequent encounters with the health care system increases the chances of outcomes being detected and documented.

Question 12. Blinding of outcome assessors

Blinding means that outcome assessors did not know whether the participant was exposed or unexposed. It is also sometimes called "masking." The objective is to look for evidence in the article that the person(s) assessing the outcome(s) for the study (for example, examining medical records to determine the outcomes that occurred in the exposed and comparison groups) is masked to the exposure status of the participant. Sometimes the person measuring the exposure is the same person conducting the outcome assessment. In this case, the outcome assessor would most likely not be blinded to exposure status because they also took measurements of exposures. If so, make a note of that in the comments section.

As you assess this criterion, think about whether it is likely that the person(s) doing the outcome assessment would know (or be able to figure out) the exposure status of the study participants. If the answer is no, then blinding is adequate. An example of adequate blinding of the outcome assessors is to create a separate committee, whose members were not involved in the care of the patient and had no information about the study participants' exposure status. The committee would then be provided with copies of participants' medical records, which had been stripped of any potential exposure information or personally identifiable information. The committee would then review the records for prespecified outcomes according to the study protocol. If blinding was not possible, which is sometimes the case, mark "NA" and explain the potential for bias.

Question 13. Followup rate

Higher overall followup rates are always better than lower followup rates, even though higher rates are expected in shorter studies, whereas lower overall followup rates are often seen in studies of longer duration. Usually, an acceptable overall followup rate is considered 80 percent or more of participants whose exposures were measured at baseline. However, this is just a general guideline. For example, a 6-month cohort study examining the relationship between dietary sodium intake and BP level may have over 90 percent followup, but a 20-year cohort study examining effects of sodium intake on stroke may have only a 65 percent followup rate.

Question 14. Statistical analyses

Were key potential confounding variables measured and adjusted for, such as by statistical adjustment for baseline differences? Logistic regression or other regression methods are often used to account for the influence of variables not of interest.

This is a key issue in cohort studies, because statistical analyses need to control for potential confounders, in contrast to an RCT, where the randomization process controls for potential confounders. All key factors that may be associated both with the exposure of interest and the outcome–that are not of interest to the research question–should be controlled for in the analyses.

For example, in a study of the relationship between cardiorespiratory fitness and CVD events (heart attacks and strokes), the study should control for age, BP, blood cholesterol, and body weight, because all of these factors are associated both with low fitness and with CVD events. Well-done cohort studies control for multiple potential confounders.

Some general guidance for determining the overall quality rating of observational cohort and cross-sectional studies

The questions on the form are designed to help you focus on the key concepts for evaluating the internal validity of a study. They are not intended to create a list that you simply tally up to arrive at a summary judgment of quality.

Internal validity for cohort studies is the extent to which the results reported in the study can truly be attributed to the exposure being evaluated and not to flaws in the design or conduct of the study–in other words, the ability of the study to draw associative conclusions about the effects of the exposures being studied on outcomes. Any such flaws can increase the risk of bias.

Critical appraisal involves considering the risk of potential for selection bias, information bias, measurement bias, or confounding (the mixture of exposures that one cannot tease out from each other). Examples of confounding include co-interventions, differences at baseline in patient characteristics, and other issues throughout the questions above. High risk of bias translates to a rating of poor quality. Low risk of bias translates to a rating of good quality. (Thus, the greater the risk of bias, the lower the quality rating of the study.)

In addition, the more attention in the study design to issues that can help determine whether there is a causal relationship between the exposure and outcome, the higher quality the study. These include exposures occurring prior to outcomes, evaluation of a dose-response gradient, accuracy of measurement of both exposure and outcome, sufficient timeframe to see an effect, and appropriate control for confounding–all concepts reflected in the tool.

Generally, when you evaluate a study, you will not see a "fatal flaw," but you will find some risk of bias. By focusing on the concepts underlying the questions in the quality assessment tool, you should ask yourself about the potential for bias in the study you are critically appraising. For any box where you check "no" you should ask, "What is the potential risk of bias resulting from this flaw in study design or execution?" That is, does this factor cause you to doubt the results that are reported in the study or doubt the ability of the study to accurately assess an association between exposure and outcome?

The best approach is to think about the questions in the tool and how each one tells you something about the potential for bias in a study. The more you familiarize yourself with the key concepts, the more comfortable you will be with critical appraisal. Examples of studies rated good, fair, and poor are useful, but each study must be assessed on its own based on the details that are reported and consideration of the concepts for minimizing bias.

Last Updated March 2014

QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES

COMPONENT RATINGS

A) SELECTION BIAS

(Q1) Are the individuals selected to participate in the study likely to be representative of the target population?

1 Very likely
2 Somewhat likely
3 Not likely
4 Can't tell

(Q2) What percentage of selected individuals agreed to participate?

1 80 - 100% agreement
2 60 – 79% agreement
3 less than 60% agreement
4 Not applicable
5 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

B) STUDY DESIGN

Indicate the study design

1 Randomized controlled trial
2 Controlled clinical trial
3 Cohort analytic (two group pre + post)
4 Case-control
5 Cohort (one group pre + post (before and after))
6 Interrupted time series
7 Other specify _____
8 Can't tell

Was the study described as randomized? If NO, go to Component C.

No Yes

If Yes, was the method of randomization described? (See dictionary)

No Yes

If Yes, was the method appropriate? (See dictionary)

No Yes

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

C) CONFOUNDERS

(Q1) Were there important differences between groups prior to the intervention?

1 Yes
2 No
3 Can't tell

The following are examples of confounders:

1 Race
2 Sex
3 Marital status/family
4 Age
5 SES (income or class)
6 Education
7 Health status
8 Pre-intervention score on outcome measure

(Q2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design (e.g. stratification, matching) or analysis)?

1 80 – 100% (most)
2 60 – 79% (some)
3 Less than 60% (few or none)
4 Can't Tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

D) BLINDING

(Q1) Was (were) the outcome assessor(s) aware of the intervention or exposure status of participants?

1 Yes
2 No
3 Can't tell

(Q2) Were the study participants aware of the research question?

1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

E) DATA COLLECTION METHODS

(Q1) Were data collection tools shown to be valid?

1 Yes
2 No
3 Can't tell

(Q2) Were data collection tools shown to be reliable?

1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

F) WITHDRAWALS AND DROP-OUTS

(Q1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?

1 Yes
2 No
3 Can't tell
4 Not Applicable (i.e. one time surveys or interviews)

(Q2) Indicate the percentage of participants completing the study. (If the percentage differs by groups, record the lowest).

1 80 -100%
2 60 - 79%
3 less than 60%
4 Can't tell
5 Not Applicable (i.e. Retrospective case-control)

RATE THIS SECTION	STRONG	MODERATE	WEAK	
See dictionary	1	2	3	Not Applicable

G) INTERVENTION INTEGRITY

(Q1) What percentage of participants received the allocated intervention or exposure of interest?

1 80 -100%
2 60 - 79%
3 less than 60%
4 Can't tell

(Q2) Was the consistency of the intervention measured?

1 Yes
2 No
3 Can't tell

(Q3) Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results?

4 Yes
5 No
6 Can't tell

H) ANALYSES

(Q1) Indicate the unit of allocation (circle one)

community organization/institution practice/office individual

(Q2) Indicate the unit of analysis (circle one)

community organization/institution practice/office individual

(Q3) Are the statistical methods appropriate for the study design?

1 Yes
2 No
3 Can't tell

(Q4) Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?

1 Yes
2 No
3 Can't tell

GLOBAL RATING

COMPONENT RATINGS

Please transcribe the information from the gray boxes on pages 1-4 onto this page. See dictionary on how to rate this section.

A	**SELECTION BIAS**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
B	**STUDY DESIGN**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
C	**CONFOUNDERS**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
D	**BLINDING**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
E	**DATA COLLECTION METHOD**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
F	**WITHDRAWALS AND DROPOUTS**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	Not Applicable

GLOBAL RATING FOR THIS PAPER (circle one):

1	STRONG	(no WEAK ratings)
2	MODERATE	(one WEAK rating)
3	WEAK	(two or more WEAK ratings)

With both reviewers discussing the ratings:

Is there a discrepancy between the two reviewers with respect to the component (A-F) ratings?

No Yes

If yes, indicate the reason for the discrepancy

1	Oversight
2	Differences in interpretation of criteria
3	Differences in interpretation of study

Final decision of both reviewers (circle one):

1	**STRONG**
2	**MODERATE**
3	**WEAK**

Reporting checklist for protocol of a systematic review.

Based on the PRISMA-P guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the PRISMA-P reporting guidelines, and cite them as:

Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA. Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 statement. Syst Rev. 2015;4(1):1.

		Reporting Item	Page Number
Identification	#1a	Identify the report as a protocol of a systematic review	1
Update	#1b	If the protocol is for an update of a previous systematic review, identify as such	n/a
	#2	If registered, provide the name of the registry (such as PROSPERO) and registration number	2
Contact	#3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author	1
Contribution	#3b	Describe contributions of protocol authors and identify the guarantor of the review	8
	#4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important	7

		protocol amendments	
Sources	#5a	Indicate sources of financial or other support for the review	8
Sponsor	#5b	Provide name for the review funder and / or sponsor	n/a
Role of sponsor or funder	#5c	Describe roles of funder(s), sponsor(s), and / or institution(s), if any, in developing the protocol	n/a
Rationale	#6	Describe the rationale for the review in the context of what is already known	3
Objectives	#7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	4
Eligibility criteria	#8	Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review	4,5
Information sources	#9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage	5
Search strategy	#10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	5
Study records - data management	#11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	6
Study records - selection process	#11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis)	6
Study records - data collection process	#11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	6
Data items	#12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications	6

Outcomes and prioritization	#13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	5
Risk of bias in individual studies	#14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	6
Data synthesis	#15a	Describe criteria under which study data will be quantitatively synthesised	7
	#15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I2, Kendall's τ)	7
	#15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)	7
	#15d	If quantitative synthesis is not appropriate, describe the type of summary planned	7
Meta-bias(es)	#16	Specify any planned assessment of meta-bias(es) (such as publication bias across studies, selective reporting within studies)	7
Confidence in cumulative evidence	#17	Describe how the strength of the body of evidence will be assessed (such as GRADE)	6

BMJ Open

The effect of in-utero exposure to HIV and antiretroviral drugs on growth in HIV-exposed uninfected children: a systematic review and meta-analysis protocol

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-023937.R2
Article Type:	Protocol
Date Submitted by the Author:	14-Feb-2019
Complete List of Authors:	Ekali, Gabriel; Universite de Yaounde I, Biotechnology Center; National AIDS Control Committee, Jesson, Julie; Universite Toulouse III Paul Sabatier, Inserm UMR 1027 Enok, Pascal; University of Montreal, Department of Preventive Medicine Leroy, Valériane; Universite Toulouse III Paul Sabatier
Primary Subject Heading:	HIV/AIDS
Secondary Subject Heading:	Paediatrics
Keywords:	HIV & AIDS < INFECTIOUS DISEASES, Nutrition < TROPICAL MEDICINE, PAEDIATRICS, Growth, Wasting, Stunting

The effect of *in-utero* exposure to HIV and antiretroviral drugs on growth in HIV-exposed uninfected children: a systematic review and meta-analysis protocol

Gabriel Loni Ekali[1,2,3,*], Julie Jesson[2], Enok Pascal[4], Valériane Leroy[2]

1.Biotechnology Center, University of Yaounde 1, Cameroon.

2. Inserm UMR 1027, Université Toulouse 3, Toulouse, France.

3. National AIDS Control Committee, Ministry of Public Health, Cameroon.

4.Department of Preventive Medicine, University of Montreal.

***Corresponding author**: Dr Gabriel Loni Ekali

Biotechnology Center, University of Yaounde 1, Cameroon; National AIDS Control Committee, Ministry of Public Health, Cameroon.

Email: eloni2000@yahoo.com

Telephone: 237 - 677333023

PO Box: 3077, Yaounde, Cameroon.

Word count: 2,396

ABSTRACT

Introduction

HIV-exposed uninfected children (HEU) have higher morbidity and mortality compared to HIV unexposed uninfected children. Despite the fact that malnutrition contributes to about half of all infant deaths below 5 years of age in low and middle income countries and that growth impairment has been reported in the HEU population, the spectrum of growth disorders associated to HIV and ART exposure during the *in utero* and perinatal periods is yet to comprehensively summarized amongst the global HEU population. This protocol for a systematic review and meta-analysis aims to critically synthesize data concerning the prevalence of underweight, stunting and wasting at different ages in the global HEU population.

Methods and analysis

Medline, EMBASE, Cochrane Library, TOXLINE, WHO Global Index Medicus and the Web of Science will be searched for relevant articles published between January 1st 1989 and December 1st 2017 without language restriction. In addition, conference abstracts and reference lists of eligible papers and relevant review articles will be screened. Authors will screen and select studies, extract data, assess risk of bias as well as studies individually for heterogeneity. Study-specific estimates will be pooled through a random-effects meta-analysis model for studies that are clinically homogenous while funnel plots and Egger's test will be used to detect publication bias. Results will be presented by ART availability period, country income levels, mode of breastfeeding.

Ethics and Dissemination

Ethical approval will not be required for this study because it will be based on published data. The final report of this study will be published in a peer-reviewed journal and presented at scientific conferences. This review will summarize the evidence and quantify the problem of growth impairment in HEU infants and so shed more light on our understanding of the higher morbidity and mortality in this growing population.

PROSPERO registration number: CRD42018091762

Key words: Growth, wasting, stunting, underweight, HIV-exposed uninfected children

STRENGTHS AND LIMITATIONS OF THE STUDY

- One of the limitations of the current review could be insufficient data to quantify the growth outcomes of interest.
- To our knowledge, this will be the first study to summarize the effect of HIV and ART exposure on growth of HIV-exposed uninfected infants compared to an appropriate control group of HIV-unexposed uninfected children.

- This review will cover both the era before widespread combination ART and the era after widespread access to combination ART so as to investigate whether any differences in growth outcomes exists between these two eras.

BACKGROUND: One of the major strategies of reducing new pediatric Human Immunodeficiency Virus (HIV) infections is by preventing mother-to-child transmission (MTCT) of the virus. This transmission can occur in utero, during delivery or breastfeeding.[1] In the absence of any intervention MTCT of HIV can be as high as 45%.[2] However, with antiretroviral therapy (ART), the rate of MTCT can be reduced to less than 2%.[3] With widespread access to ART, the UNAIDS report that new pediatric HIV infections have declined by 47% between 2010 and 2016, while coverage of antiretroviral medicines provided to pregnant women living with HIV to prevent transmission to their children rose from 47% to 76% during the same period.[4]

As MTCT rates drop, the proportion of children who are exposed in utero to HIV and ART and are born uninfected with the virus (HEU) is increasing worldwide, especially in sub-Saharan Africa where the greatest burden of the HIV pandemic is found. Indeed, HEU children constitute about 30% of the Southern African pediatric population.[5] Even though HEU children are not infected with HIV, a growing body of literature shows evidence of higher morbidity and mortality compared to HIV/ART unexposed uninfected children (HUU). Higher frequency of hospitalisations, more severe respiratory tract infections especially from encapsulated bacteria as well as diarrheal diseases have been described.[5-6] Children exposed to maternal HIV and ART have modest but significant impairment of development and a higher risk of growth impairment.[7] Growth failure, reflected by restricted linear growth and weight gain, have been reported amongst HEU children.A number of recent reviews have dwelt on summarizing the available evidence of higher morbidity, mortality, immune dysfunction and its possible role in the higher susceptibility of HEU to diseases.[5-7,8-9] However, the spectrum of growth disorders found in this growing population and its possible association and role to the higher susceptibility to infectious diseases they experience is yet to be documented. This is important as it is known that impaired growth clinically represents an altered nutritional status, which is significantly associated to mortality in infancy. Indeed, nutrition-related factors contribute to about 45% of deaths in children under 5 years of age in low and middle income countries, translating into the loss of approximately 3 million young lives annually.[10] Furthermore, restricted linear growth in early life also affects the quality of life an individual as it is associated to cognitive impairment [11] and lower lifetime earnings [12]. In addition, appropriate comparisons of HEU and HUU infants were rare before the antiretroviral era. With increasing ART access, it is methodologically challenging to disentangle effects of HIV exposure from those associated with ART exposure, particularly in regions where malnutrition and co-morbidities are common and could increase the potential effects of these exposures.

In view of the recent reports of altered linear growth and weight gain among HEU infants, the main goal of this paper is to review and summarize evidence on the effects of perinatal/postnatal

exposure to HIV and ART on growth outcomes in HEU children until 60 months of age and in later life.

Review question

What is the effect of in utero/postnatal exposure to HIV and ART on growth outcomes (stunting, wasting, underweight) until 60 months of age in HEU compared to HUU children as documented in studies reported between January 1st, 1989 to December 1st, 2017.

Specific objectives

To answer this review question, a systematic review and meta-analysis of the existing evidence will be conducted, the objectives of which will be to determine:

1. The effect of HIV and ART exposure on underweight (weight-for-age < 2 SD) until 60 months of age in the global HEU population compared to HUU children.
2. The effect of HIV and ART exposure on stunting (height/length-for-age < 2 SD) until 60 months of age in the global HEU population compared to HUU children.
3. The effect of HIV and ART exposure on wasting (weight-for-height/length < 2 SD) until 60 months of age in the global HEU population compared to HUU children.

METHODS AND ANALYSIS

Design

This will be a systematic review and meta-analysis of the published literature. The Centre for Reviews and Dissemination guidelines [13] will be used for the methodology of this review and the review will be registered in the PROSPERO International Prospective Register of systematic reviews.

Inclusion criteria:

1. Type of studies: all observational cohort studies and randomized controlled trials that have data on any growth parameter of interest, either as main or secondary outcomes. Only comparative studies with HEU and HUU children will be included in the meta-analysis.
2. Type of participants: adequately defined HEU (children exposed in utero to HIV and combination ART of any class, who are declared HIV negative by either DNA PCR between 6 weeks and 18 months OR appropriate HIV serological tests above 18 months) and HUU children (children born of mothers with documented HIV negative serological test), aged 0 to 5 years. All ART exposure will be considered.
3. Outcome measures: prevalence of, stunting, wasting, underweight; growth velocity or studies that have enough data to compute the outcomes of interest at yearly time intervals, available from 12 to 60 months
4. Type and period of publications: published studies and conference abstracts from 1 January 1989 to 1 December 2017.

For studies with several publications of their findings over time, we will retain in the review the one that has the best quality and the most appropriate regarding our objective. For the meta-analysis, studies with a documented measure of effect of in utero/postnatal HIV and ART exposure on growth comparing HEU and HUU groups will be included. For studies without reported measures of effect, these will be computed if the provided data is adequate.

Exclusion criteria:

We will exclude:

1. Type of studies: unpublished manuscripts, case reviews, policy reports, commentaries and editorials.
2. Studies whose data will not be sufficient to calculate the appropriate measures of effect.

Outcome measures: The following measures will be considered at different age points (6, 9,12,18,24, 36,48,60 months)

- height-for-age
- weight-for-height/length
- weight-for-age
- growth velocity
- prevalence of stunting, wasting and underweight at each age point.

Search strategy to identify relevant studies

Databases: An exhaustive literature search will be carried out in Medline, EMBASE, Cochrane Library, TOXLINE, WHO Global Index Medicus and the Web of Science to identify relevant articles published on growth in HIV-exposed uninfected infants between January 1st 1989 and December 1st 2017. The proceedings of the following conferences will also be searched: International AIDS Society Conference on HIV Pathogenesis, Treatment and Prevention; International AIDS Conference; Conference of Retroviruses and Opportunistic Infections. The search strategy that will be used for online databases is presented in table 1.

Table 1: Search strategy for the review of impact of in utero exposure to HIV and ART on growth in HIV-exposed uninfected infants

Search	Search Terms
#1	HIV OR HIV-1 OR HIV-2 OR HIV1 OR HIV2 OR HIV infect* OR human immune*
#2	HIV-exposed uninfected OR HIV exposed uninfected OR HEU OR HIV-EU OR HIV-unexposed OR HIV unexposed OR HUU OR HU OR HIV-UE OR HIV-uninfected OR HIV uninfected

#3	Children OR child OR infant OR infants OR pediatric OR pediatrics OR paediatric OR paediatrics
#4	Growth OR development OR stunting OR malnutrition OR wasting OR underweight OR weight OR length OR height OR BMI OR head circumference OR growth faltering OR Growth velocity
#5	Limits: from 1 January 1989 to 1 December 2017
#6	#1 AND #2 AND #3 AND #4 AND #5

Searching other sources

The list of references of eligible articles as well as relevant review articles will also be manually searched.

Selecting studies for inclusion in the review

All articles returned by the search will be saved to the Endnote version X4 software which will be used to remove duplicates. The titles and abstracts of the articles remaining after exclusion of duplicates will be reviewed to identify potentially relevant papers according to the inclusion and exclusion criteria. Two investigators will independently review retrieved full text of all identified papers for eligibility and then retain those that are eligible by consensus. Any disagreement between investigators will be settled by a third reviewer. A PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram will be used to detail the number of articles identified, screened, included and excluded (online supplementary file 1).

Data extraction and management

A tool for data abstraction will be developed and the following data will be extracted:

1. Details about the author: Name, Journal
2. Characteristics of the study: design, country, period when study was conducted, year of publication, duration of follow-up, sample size, study endpoints,
3. Characteristics of participants: Age, sex, class of maternal ART, term of delivery, low birth-weight, mode of infant feeding (no BF, mixed feeding, exclusive BF), duration of breastfeeding (BF), Cotrimoxazole prophylaxis, growth references used,
4. Outcome measures at different time points: weight-for-age, length-for-age, weight-for-length, BMI for age and growth velocity.

Two investigators will independently extract data and any disagreements will be reconciled through discussion or by a third investigator as necessary.

Assessing methodological quality of included studies and risk of bias

The quality assessment tool for Observational Cohort and Cross-sectional studies of the National Heart, Lung and Blood Institute (NHLBI) will be used to assess the quality of observational cohort studies [14] (online supplementary file 2) while for intervention studies, the Effective Public Health Practice (EPHPP) tool [15] (online supplementary file 3) will be used by two investigators independently. Agreement between the two investigators will be measured by Cohen's kappa statistic. For the NHLBI tool, the score will be graded as Good, Fair and Poor while for the EPHPP tool, it will be graded as Strong, Moderate and Weak.

Data synthesis and analysis

Prevalences of underweight, stunting and wasting in the HEU and HUU groups will be recalculated based on the information provided by individual studies and odd ratios will be computed at each age point. The exposure variable will be HIV and combination ART exposure in utero and the outcome variable will be the different measures of growth: length/height-for-age, weight-for-height, then weight-for-age. We will use the statistic $I^2 = 100*(Q-df)/Q$ Q (with Q being the statistic of the Cochran heterogeneity Q-test and df the number of degrees of freedom corresponding to the number of studies minus one), which makes it possible to capture the proportion of the total variance observed due to a real difference in the measures of effects between the studies, in order to explore heterogeneity. Unlike the Q-test, this statistic is not influenced by the number of studies. Moreover, a non-significant Q-test does not necessarily imply that there is no heterogeneity between the studies because the power of the test depends on the number of studies. For statistical I^2, the values of 0%, 25%, 50% and 75% respectively represent the following heterogeneity levels: absence, weak, moderate and high. For each risk factor, the combined measure of relative risk will be estimated from a random-effect model if $I^2 \geq 25\%$ and a fixed-effect model if $I^2 < 25\%$. If substantial heterogeneity is observed, a meta-regression will be used to identify the characteristics of the studies (e.g study quality, study location, sample size, adjusting of confounders or not, antiretroviral therapy agents and regimens, duration and mode of breastfeeding, maternal socioeconomic status, etc.) that may explain the observed heterogeneity. All data will be analyzed using Stata V.14 (StataCorp, Texas, USA). A forest plot will be presented for each factor studied in the studies found. Qualitative synthesis will be used in cases where data extracted is insufficient to perform quantitative synthesis or when the included studies significantly differ in design, setting and outcome measures.

Assessment of reporting bias

To explore a potential publication bias, funnel plots and Egger's test will be done and a p value below 0.10 will be considered indicative of significant publication bias.

Presentation and reporting of the results

The results of this meta-analysis will be presented by HIV and ART exposure status, by pre-ART and post-ART era according to the different classes of ART, by infant feeding mode, and if feasible by country income level. We will prioritize length/height-for-age, weight-for-height outcomes report first, weight-for-age being less discriminant. [16]

As NCHS/CDC/WHO [17] growth standards were used before 2005, and WHO child growth standards were used after 2005 [18], this may affect the longitudinal comparisons of the outcomes. Therefore, we will stratify our results according to the different standards. We will carry out sensitivity analysis to determine whether low birth weight has an impact on our results. Adherence to ART will be documented if reported by authors of the selected papers and discussed in the qualitative synthesis section of our manuscript.
The Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocol (PRISMA-P) will be used for reporting this present protocol (online supplementary file 4).[19] Results of the search method will be presented with a flow chart, detailing the selection process. Causes of exclusion for studies firstly identified as eligible will be documented.

Patient and Public Involvement
Patients were not involved in the design of this study protocol.

POTENTIAL PROTOCOL AMENDMENTS
The current protocol as written will not be amended in the course of the study. This is to avoid any outcome reporting bias.

ETHICAL CONSIDERATIONS AND DISSEMINATION
In this study, only published data will be used. No ethical clearance will therefore be needed. The final results will be published in peer-reviewed journals and also presented at conferences.

CONCLUSION
A comprehensive summary of the evidence concerning how in utero/postnatal exposure to HIV and ART affect growth is critical to shed more light on our understanding of why HEU children have higher morbidity and mortality to certain infectious diseases compared to HUU. It is hoped that the results of this review will quantify the problem of growth impairment and draw attention to the necessity to possibly develop appropriate nutritional interventions for this growing population.

AUTHOR STATEMENT: EG, LV had the idea and EG designed the protocol. LV, JJ, EP critically reviewed intellectual content and revised the methodology of the study. EG and LV are guarantors of the study and all the authors approved the final version of the manuscript.

FUNDING: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

COMPETING INTEREST: None declared

REFERENCES

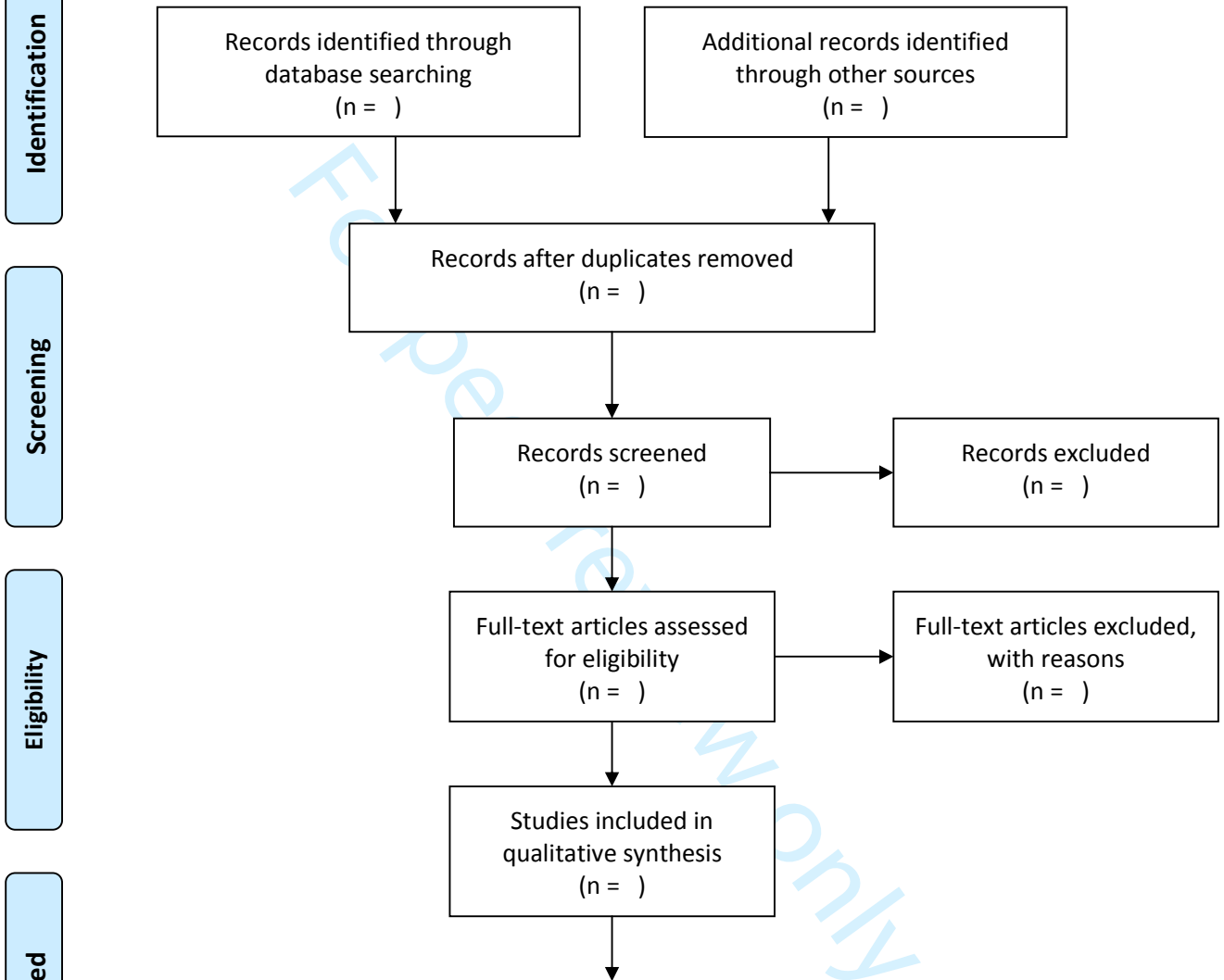
1. Mofenson LM. Mother-child HIV-1 transmission: Timing and determinants. *Obstet Gynecol Clin North Am* 1997; 24:759.
2. Lehman DA, Farquhar C. Biological mechanisms of vertical human immunodeficiency virus (HIV-1) transmission. *Rev Med Virol*. 2007;17:381–403

3. Brocklehurst P: Interventions for reducing the risk of mother-to-child transmission of HIV Infection. *Cochrane Database Syst Rev* 2002;1:CD000102
4. UNAIDS data 2017. Available at http://www.unaids.org/sites/default/files/media_asset/20170720_Data_book_2017_en.pdf. Accessed on January 29th 2018 at 00h24minutes.
5. Slogrove AL, Goetghebuer T, Cotton MF et al. Pattern of Infectious Morbidity in HIV-Exposed Uninfected Infants and Children. *Front Immunol* 2016 May 6;7:164
6. Brennan AT, Bonawitz R, Gill CJ, et al. A meta-analysis assessing all-cause mortality in HIV-exposed uninfected compared with HIV-unexposed uninfected infants and children. *AIDS* 2016 Sep 24;30(15):2351-60
7. Desmonde S, Goetghebuer T, Thorne C, et al. Health and survival of HIV perinatally exposed but uninfected children born to HIV-infected mothers. *Curr Opin HIV AIDS* 2016 Sep;11(5):465-476
8. Abu-Raya B, Kollmann TR, Marchant A, et al. The Immune System of HIV-Exposed Uninfected Infants. *Front Immunol* 2016 Sep 28;7:383
9. Ruck C, Reikie BA, Marchant A, et al. Linking Susceptibility to Infectious Diseases to Immune System Abnormalities among HIV-Exposed Uninfected Infants. *Front Immunol* 2016 Aug 19;7:310
10. WHO Media Center. Children : reducing child mortality Fact sheet. Available from http://www.who.int/mediacentre/factsheets/fs178/en/. Accessed on February 2nd 2018 at 23h56.
11. Sudfeld CR, McCoy DC, Danaie G, et al. Linear growth and child development and low- and middle-income countries: a meta-analysis. Pediatrics. 2015;135:e1266–e1275.
12. Adair LS, Fall CH, Osmond C, et al. Associations of linear growth and relative weight gain during early life with adult health and human capital in countries of low and middle income: findings five birth cohort studies. Lancet. 2013;382:525–534
13. Centers for Reviews and Dissemination. CRD's guidance for undertaking reviews in healthcare. Centers for Reviews and Dissemination; 2009. Available from: http://www.york.ac.uk/crd/SysRev/!SSL!/WebHelp/SysRev3.htm. Accessed February 2nd 2018 at 01h04
14. National Heart, Lung and Blood Institute. Quality assessment tool for Observational Cohort and Cross-sectional studies. Available from https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools and accessed February 2nd 2018 at 01h06.
15. Effective Public Health Practice Project. Quality assessment tool for quantitative studies. Available at http://www.ephpp.ca/PDF/Quality%20Assessment%20Tool_2010_2.pdf . Accessed on February 2nd 2018 at 01h07.
16. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.doi:10.1186/2046-4053-4-1
17. Geneva: WHO; 1997. Global database on child growth and malnutrition. Available at http://apps.who.int/iris/bitstream/handle/10665/63750/WHO_NUT_97.4.pdf;jsessioni

d=3467A6BA2F61719518EEF6309CE1ACD7?sequence=1. Accessed on September 24th 2019 at 3:43am WAT.

18. Waterlow JC, Buzina R, Keller W, Lane JM, Nichaman MZ, Tanner JM. The presentation and use of height and weight data for comparing the nutritional status of groups of children under the age of 10 years. Bull World Health Organ. 1977;55(4):489-98
19. Geneva: WHO; 2007. WHO Anthro for personal computers, version Software for assessing growth and development of the world's children. http://www.who.int/childgrowth/software/en/ [Ref list]

PRISMA 2009 Flow Diagram



From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit www.prisma-statement.org.

National Heart, Lung, and Blood Institute

Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies

Criteria	Yes	No	Other (CD, NR, NA)*
1. Was the research question or objective in this paper clearly stated?			
2. Was the study population clearly specified and defined?			
3. Was the participation rate of eligible persons at least 50%?			
4. Were all the subjects selected or recruited from the same or similar populations (including the same time period)? Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?			
5. Was a sample size justification, power description, or variance and effect estimates provided?			
6. For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?			
7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?			
8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?			
9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?			
10. Was the exposure(s) assessed more than once over time?			
11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?			
12. Were the outcome assessors blinded to the exposure status of participants?			
13. Was loss to follow-up after baseline 20% or less?			
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?			

Quality Rating (Good, Fair, or Poor) (see guidance)
Rater #1 initials:
Rater #2 initials:
Additional Comments (If POOR, please state why):

*CD, cannot determine; NA, not applicable; NR, not reported

Guidance for Assessing the Quality of Observational Cohort and Cross-Sectional Studies

The guidance document below is organized by question number from the tool for quality assessment of observational cohort and cross-sectional studies.

Question 1. Research question

Did the authors describe their goal in conducting this research? Is it easy to understand what they were looking to find? This issue is important for any scientific paper of any type. Higher quality scientific research explicitly defines a research question.

Questions 2 and 3. Study population

Did the authors describe the group of people from which the study participants were selected or recruited, using demographics, location, and time period? If you were to conduct this study again, would you know who to recruit, from where, and from what time period? Is the cohort population free of the outcomes of interest at the time they were recruited?

An example would be men over 40 years old with type 2 diabetes who began seeking medical care at Phoenix Good Samaritan Hospital between January 1, 1990 and December 31, 1994. In this example, the population is clearly described as: (1) who (men over 40 years old with type 2 diabetes); (2) where (Phoenix Good Samaritan Hospital); and (3) when (between January 1, 1990 and December 31, 1994). Another example is women ages 34 to 59 years of age in 1980 who were in the nursing profession and had no known coronary disease, stroke, cancer, hypercholesterolemia, or diabetes, and were recruited from the 11 most populous States, with contact information obtained from State nursing boards.

In cohort studies, it is crucial that the population at baseline is free of the outcome of interest. For example, the nurses' population above would be an appropriate group in which to study incident coronary disease. This information is usually found either in descriptions of population recruitment, definitions of variables, or inclusion/exclusion criteria.

You may need to look at prior papers on methods in order to make the assessment for this question. Those papers are usually in the reference list.

If fewer than 50% of eligible persons participated in the study, then there is concern that the study population does not adequately represent the target population. This increases the risk of bias.

Question 4. Groups recruited from the same population and uniform eligibility criteria

Were the inclusion and exclusion criteria developed prior to recruitment or selection of the study population? Were the same underlying criteria used for all of the subjects involved? This issue is related to the description of the study population, above, and you may find the information for both of these questions in the same section of the paper.

Most cohort studies begin with the selection of the cohort; participants in this cohort are then measured or evaluated to determine their exposure status. However, some cohort studies may recruit or select exposed participants in a different time or place than unexposed participants, especially retrospective cohort studies–which is when data are obtained from the past (retrospectively), but the analysis examines exposures prior to outcomes. For example, one research question could be whether diabetic men with clinical depression are at higher risk for cardiovascular disease than those without clinical depression. So, diabetic men with depression might be selected from a mental health clinic, while diabetic men without depression might be selected from an internal medicine or endocrinology clinic. This study recruits groups from different clinic populations, so this example would get a "no."

However, the women nurses described in the question above were selected based on the same inclusion/exclusion criteria, so that example would get a "yes."

Question 5. Sample size justification

Did the authors present their reasons for selecting or recruiting the number of people included or analyzed? Do they note or discuss the statistical power of the study? This question is about whether or not the study had enough participants to detect an association if one truly existed.

A paragraph in the methods section of the article may explain the sample size needed to detect a hypothesized difference in outcomes. You may also find a discussion of power in the discussion section (such as the study had 85 percent power to detect a 20 percent increase in the rate of an outcome of interest, with a 2-sided alpha of 0.05). Sometimes estimates of variance and/or estimates of effect size are given, instead of sample size calculations. In any of these cases, the answer would be "yes."

However, observational cohort studies often do not report anything about power or sample sizes because the analyses are exploratory in nature. In this case, the answer would be "no." This is not a "fatal flaw." It just may indicate that attention was not paid to whether the study was sufficiently sized to answer a prespecified question–i.e., it may have been an exploratory, hypothesis-generating study.

Question 6. Exposure assessed prior to outcome measurement

This question is important because, in order to determine whether an exposure causes an outcome, the exposure must come before the outcome.

For some prospective cohort studies, the investigator enrolls the cohort and then determines the exposure status of various members of the cohort (large epidemiological studies like Framingham used this approach). However, for other cohort studies, the cohort is selected based on its exposure status, as in the example above of depressed diabetic men (the exposure being depression). Other examples include a cohort identified by its exposure to fluoridated drinking water and then compared to a cohort living in an area without fluoridated water, or a cohort of military personnel exposed to combat in the Gulf War compared to a cohort of military personnel not deployed in a combat zone.

With either of these types of cohort studies, the cohort is followed forward in time (i.e., prospectively) to assess the outcomes that occurred in the exposed members compared to nonexposed members of the cohort. Therefore, you begin the study in the present by looking at groups that were exposed (or not) to some biological or behavioral factor, intervention, etc., and then you follow them forward in time to examine outcomes. If a cohort study is conducted properly, the answer to this question should be "yes," since the exposure status of members of the cohort was determined at the beginning of the study before the outcomes occurred.

For retrospective cohort studies, the same principal applies. The difference is that, rather than identifying a cohort in the present and following them forward in time, the investigators go back in time (i.e., retrospectively) and select a cohort based on their exposure status in the past and then follow them forward to assess the outcomes that occurred in the exposed and nonexposed cohort members. Because in retrospective cohort studies the exposure and outcomes may have already occurred (it depends on how long they follow the cohort), it is important to make sure that the exposure preceded the outcome.

Sometimes cross-sectional studies are conducted (or cross-sectional analyses of cohort-study data), where the exposures and outcomes are measured during the same timeframe. As a result, cross-sectional analyses provide weaker evidence than regular cohort studies regarding a potential causal relationship between exposures and outcomes. For cross-sectional analyses, the answer to Question 6 should be "no."

Question 7. Sufficient timeframe to see an effect

Did the study allow enough time for a sufficient number of outcomes to occur or be observed, or enough time for an exposure to have a biological effect on an outcome? In the examples given above, if clinical depression has a biological effect on increasing risk for CVD, such an effect may take years. In the other example, if higher dietary sodium increases BP, a short timeframe may be sufficient to assess its association with BP, but a longer timeframe would be needed to examine its association with heart attacks.

The issue of timeframe is important to enable meaningful analysis of the relationships between exposures and outcomes to be conducted. This often requires at least several years, especially when looking at health outcomes, but it depends on the research question and outcomes being examined.

Cross-sectional analyses allow no time to see an effect, since the exposures and outcomes are assessed at the same time, so those would get a "no" response.

Question 8. Different levels of the exposure of interest

If the exposure can be defined as a range (examples: drug dosage, amount of physical activity, amount of sodium consumed), were multiple categories of that exposure assessed? (for example, for drugs: not on the medication, on a low dose, medium dose, high dose; for dietary sodium, higher than average U.S. consumption, lower than recommended consumption, between the two). Sometimes discrete categories of exposure are not used, but instead exposures are measured as continuous variables (for example, mg/day of dietary sodium or BP values).

In any case, studying different levels of exposure (where possible) enables investigators to assess trends or dose-response relationships between exposures and outcomes–e.g., the higher the exposure, the greater the rate of the health outcome. The presence of trends or dose-response relationships lends credibility to the hypothesis of causality between exposure and outcome.

For some exposures, however, this question may not be applicable (e.g., the exposure may be a dichotomous variable like living in a rural setting versus an urban setting, or vaccinated/not vaccinated with a one-time vaccine). If there are only two possible exposures (yes/no), then this question should be given an "NA," and it should not count negatively towards the quality rating.

Question 9. Exposure measures and assessment

Were the exposure measures defined in detail? Were the tools or methods used to measure exposure accurate and reliable–for example, have they been validated or are they objective? This issue is important as it influences confidence in the reported exposures. When exposures are measured with less accuracy or validity, it is

harder to see an association between exposure and outcome even if one exists. Also as important is whether the exposures were assessed in the same manner within groups and between groups; if not, bias may result.

For example, retrospective self-report of dietary salt intake is not as valid and reliable as prospectively using a standardized dietary log plus testing participants' urine for sodium content. Another example is measurement of BP, where there may be quite a difference between usual care, where clinicians measure BP however it is done in their practice setting (which can vary considerably), and use of trained BP assessors using standardized equipment (e.g., the same BP device which has been tested and calibrated) and a standardized protocol (e.g., patient is seated for 5 minutes with feet flat on the floor, BP is taken twice in each arm, and all four measurements are averaged). In each of these cases, the former would get a "no" and the latter a "yes."

Here is a final example that illustrates the point about why it is important to assess exposures consistently across all groups: If people with higher BP (exposed cohort) are seen by their providers more frequently than those without elevated BP (nonexposed group), it also increases the chances of detecting and documenting changes in health outcomes, including CVD-related events. Therefore, it may lead to the conclusion that higher BP leads to more CVD events. This may be true, but it could also be due to the fact that the subjects with higher BP were seen more often; thus, more CVD-related events were detected and documented simply because they had more encounters with the health care system. Thus, it could bias the results and lead to an erroneous conclusion.

Question 10. Repeated exposure assessment

Was the exposure for each person measured more than once during the course of the study period? Multiple measurements with the same result increase our confidence that the exposure status was correctly classified. Also, multiple measurements enable investigators to look at changes in exposure over time, for example, people who ate high dietary sodium throughout the followup period, compared to those who started out high then reduced their intake, compared to those who ate low sodium throughout. Once again, this may not be applicable in all cases. In many older studies, exposure was measured only at baseline. However, multiple exposure measurements do result in a stronger study design.

Question 11. Outcome measures

Were the outcomes defined in detail? Were the tools or methods for measuring outcomes accurate and reliable–for example, have they been validated or are they objective? This issue is important because it influences confidence in the validity of study results. Also important is whether the outcomes were assessed in the same manner within groups and between groups.

An example of an outcome measure that is objective, accurate, and reliable is death–the outcome measured with more accuracy than any other. But even with a measure as objective as death, there can be differences in the accuracy and reliability of how death was assessed by the investigators. Did they base it on an autopsy report, death certificate, death registry, or report from a family member? Another example is a study of whether dietary fat intake is related to blood cholesterol level (cholesterol level being the outcome), and the cholesterol level is measured from fasting blood samples that are all sent to the same laboratory. These examples would get a "yes." An example of a "no" would be self-report by subjects that they had a heart attack, or self-report of how much they weigh (if body weight is the outcome of interest).

Similar to the example in Question 9, results may be biased if one group (e.g., people with high BP) is seen more frequently than another group (people with normal BP) because more frequent encounters with the health care system increases the chances of outcomes being detected and documented.

Question 12. Blinding of outcome assessors

Blinding means that outcome assessors did not know whether the participant was exposed or unexposed. It is also sometimes called "masking." The objective is to look for evidence in the article that the person(s) assessing the outcome(s) for the study (for example, examining medical records to determine the outcomes that occurred in the exposed and comparison groups) is masked to the exposure status of the participant. Sometimes the person measuring the exposure is the same person conducting the outcome assessment. In this case, the outcome assessor would most likely not be blinded to exposure status because they also took measurements of exposures. If so, make a note of that in the comments section.

As you assess this criterion, think about whether it is likely that the person(s) doing the outcome assessment would know (or be able to figure out) the exposure status of the study participants. If the answer is no, then blinding is adequate. An example of adequate blinding of the outcome assessors is to create a separate committee, whose members were not involved in the care of the patient and had no information about the study participants' exposure status. The committee would then be provided with copies of participants' medical records, which had been stripped of any potential exposure information or personally identifiable information. The committee would then review the records for prespecified outcomes according to the study protocol. If blinding was not possible, which is sometimes the case, mark "NA" and explain the potential for bias.

Question 13. Followup rate

Higher overall followup rates are always better than lower followup rates, even though higher rates are expected in shorter studies, whereas lower overall followup rates are often seen in studies of longer duration. Usually, an acceptable overall followup rate is considered 80 percent or more of participants whose exposures were measured at baseline. However, this is just a general guideline. For example, a 6-month cohort study examining the relationship between dietary sodium intake and BP level may have over 90 percent followup, but a 20-year cohort study examining effects of sodium intake on stroke may have only a 65 percent followup rate.

Question 14. Statistical analyses

Were key potential confounding variables measured and adjusted for, such as by statistical adjustment for baseline differences? Logistic regression or other regression methods are often used to account for the influence of variables not of interest.

This is a key issue in cohort studies, because statistical analyses need to control for potential confounders, in contrast to an RCT, where the randomization process controls for potential confounders. All key factors that may be associated both with the exposure of interest and the outcome–that are not of interest to the research question–should be controlled for in the analyses.

For example, in a study of the relationship between cardiorespiratory fitness and CVD events (heart attacks and strokes), the study should control for age, BP, blood cholesterol, and body weight, because all of these factors are associated both with low fitness and with CVD events. Well-done cohort studies control for multiple potential confounders.

Some general guidance for determining the overall quality rating of observational cohort and cross-sectional studies

The questions on the form are designed to help you focus on the key concepts for evaluating the internal validity of a study. They are not intended to create a list that you simply tally up to arrive at a summary judgment of quality.

Internal validity for cohort studies is the extent to which the results reported in the study can truly be attributed to the exposure being evaluated and not to flaws in the design or conduct of the study–in other words, the ability of the study to draw associative conclusions about the effects of the exposures being studied on outcomes. Any such flaws can increase the risk of bias.

Critical appraisal involves considering the risk of potential for selection bias, information bias, measurement bias, or confounding (the mixture of exposures that one cannot tease out from each other). Examples of confounding include co-interventions, differences at baseline in patient characteristics, and other issues throughout the questions above. High risk of bias translates to a rating of poor quality. Low risk of bias translates to a rating of good quality. (Thus, the greater the risk of bias, the lower the quality rating of the study.)

In addition, the more attention in the study design to issues that can help determine whether there is a causal relationship between the exposure and outcome, the higher quality the study. These include exposures occurring prior to outcomes, evaluation of a dose-response gradient, accuracy of measurement of both exposure and outcome, sufficient timeframe to see an effect, and appropriate control for confounding–all concepts reflected in the tool.

Generally, when you evaluate a study, you will not see a "fatal flaw," but you will find some risk of bias. By focusing on the concepts underlying the questions in the quality assessment tool, you should ask yourself about the potential for bias in the study you are critically appraising. For any box where you check "no" you should ask, "What is the potential risk of bias resulting from this flaw in study design or execution?" That is, does this factor cause you to doubt the results that are reported in the study or doubt the ability of the study to accurately assess an association between exposure and outcome?

The best approach is to think about the questions in the tool and how each one tells you something about the potential for bias in a study. The more you familiarize yourself with the key concepts, the more comfortable you will be with critical appraisal. Examples of studies rated good, fair, and poor are useful, but each study must be assessed on its own based on the details that are reported and consideration of the concepts for minimizing bias.

Last Updated March 2014

QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES

COMPONENT RATINGS

A) SELECTION BIAS

(Q1) Are the individuals selected to participate in the study likely to be representative of the target population?

1 Very likely
2 Somewhat likely
3 Not likely
4 Can't tell

(Q2) What percentage of selected individuals agreed to participate?

1 80 - 100% agreement
2 60 – 79% agreement
3 less than 60% agreement
4 Not applicable
5 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

B) STUDY DESIGN

Indicate the study design

1 Randomized controlled trial
2 Controlled clinical trial
3 Cohort analytic (two group pre + post)
4 Case-control
5 Cohort (one group pre + post (before and after))
6 Interrupted time series
7 Other specify _____
8 Can't tell

Was the study described as randomized? If NO, go to Component C.

No Yes

If Yes, was the method of randomization described? (See dictionary)

No Yes

If Yes, was the method appropriate? (See dictionary)

No Yes

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

C) CONFOUNDERS

(Q1) Were there important differences between groups prior to the intervention?

1 Yes
2 No
3 Can't tell

The following are examples of confounders:

1 Race
2 Sex
3 Marital status/family
4 Age
5 SES (income or class)
6 Education
7 Health status
8 Pre-intervention score on outcome measure

(Q2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design (e.g. stratification, matching) or analysis)?

1 80 – 100% (most)
2 60 – 79% (some)
3 Less than 60% (few or none)
4 Can't Tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

D) BLINDING

(Q1) Was (were) the outcome assessor(s) aware of the intervention or exposure status of participants?

1 Yes
2 No
3 Can't tell

(Q2) Were the study participants aware of the research question?

1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

E) DATA COLLECTION METHODS

(Q1) Were data collection tools shown to be valid?

1 Yes
2 No
3 Can't tell

(Q2) Were data collection tools shown to be reliable?

1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

F) WITHDRAWALS AND DROP-OUTS

(Q1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?

1 Yes
2 No
3 Can't tell
4 Not Applicable (i.e. one time surveys or interviews)

(Q2) Indicate the percentage of participants completing the study. (If the percentage differs by groups, record the lowest).

1 80 -100%
2 60 - 79%
3 less than 60%
4 Can't tell
5 Not Applicable (i.e. Retrospective case-control)

RATE THIS SECTION	STRONG	MODERATE	WEAK	
See dictionary	1	2	3	Not Applicable

G) INTERVENTION INTEGRITY

(Q1) What percentage of participants received the allocated intervention or exposure of interest?

1 80 -100%
2 60 - 79%
3 less than 60%
4 Can't tell

(Q2) Was the consistency of the intervention measured?

1 Yes
2 No
3 Can't tell

(Q3) Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results?

4 Yes
5 No
6 Can't tell

H) ANALYSES

(Q1) Indicate the unit of allocation (circle one)

community organization/institution practice/office individual

(Q2) Indicate the unit of analysis (circle one)

community organization/institution practice/office individual

(Q3) Are the statistical methods appropriate for the study design?

1 Yes
2 No
3 Can't tell

(Q4) Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?

1 Yes
2 No
3 Can't tell

GLOBAL RATING

COMPONENT RATINGS

Please transcribe the information from the gray boxes on pages 1-4 onto this page. See dictionary on how to rate this section.

A	**SELECTION BIAS**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
B	**STUDY DESIGN**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
C	**CONFOUNDERS**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
D	**BLINDING**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
E	**DATA COLLECTION METHOD**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
F	**WITHDRAWALS AND DROPOUTS**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	Not Applicable

GLOBAL RATING FOR THIS PAPER (circle one):

1	STRONG	(no WEAK ratings)
2	MODERATE	(one WEAK rating)
3	WEAK	(two or more WEAK ratings)

With both reviewers discussing the ratings:

Is there a discrepancy between the two reviewers with respect to the component (A-F) ratings?

No Yes

If yes, indicate the reason for the discrepancy

1	Oversight
2	Differences in interpretation of criteria
3	Differences in interpretation of study

Final decision of both reviewers (circle one):

1	**STRONG**
2	**MODERATE**
3	**WEAK**

Reporting checklist for protocol of a systematic review.

Based on the PRISMA-P guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the PRISMA-P reporting guidelines, and cite them as:

Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA. Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 statement. Syst Rev. 2015;4(1):1.

		Reporting Item	Page Number
Identification	#1a	Identify the report as a protocol of a systematic review	1
Update	#1b	If the protocol is for an update of a previous systematic review, identify as such	n/a
	#2	If registered, provide the name of the registry (such as PROSPERO) and registration number	2
Contact	#3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author	1
Contribution	#3b	Describe contributions of protocol authors and identify the guarantor of the review	8
	#4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important	7

		protocol amendments	
Sources	#5a	Indicate sources of financial or other support for the review	8
Sponsor	#5b	Provide name for the review funder and / or sponsor	n/a
Role of sponsor or funder	#5c	Describe roles of funder(s), sponsor(s), and / or institution(s), if any, in developing the protocol	n/a
Rationale	#6	Describe the rationale for the review in the context of what is already known	3
Objectives	#7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	4
Eligibility criteria	#8	Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review	4,5
Information sources	#9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage	5
Search strategy	#10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	5
Study records - data management	#11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	6
Study records - selection process	#11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis)	6
Study records - data collection process	#11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	6
Data items	#12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications	6

Outcomes and prioritization	#13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	5
Risk of bias in individual studies	#14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	6
Data synthesis	#15a	Describe criteria under which study data will be quantitatively synthesised	7
	#15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I2, Kendall's τ)	7
	#15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)	7
	#15d	If quantitative synthesis is not appropriate, describe the type of summary planned	7
Meta-bias(es)	#16	Specify any planned assessment of meta-bias(es) (such as publication bias across studies, selective reporting within studies)	7
Confidence in cumulative evidence	#17	Describe how the strength of the body of evidence will be assessed (such as GRADE)	6