Exercises

# Chapter 6: Assessing the evidence base

## Exercise 6.1: Study selection

Now revisit your inclusion and exclusion criteria from Chapter 4 and write them down in the grid provided. Keep this grid to hand whilst you go through your study selection process.

|  |
| --- |
| *Tool for focusing your question – PICO, PICOC, SPICE, etc* |
| **Inclusion criteria**Population = Intervention/exposure = Comparator = Outcome (s) = Context =  |

### Sample Response

For our earlier (fictitious) systematic review on mobile health interventions for maternal, newborn and child health in low- and middle-income countries, we have devised the

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **Inclusions** | **Exclusions** |
| **P**opulation  | Mothers, their newborns and young children [NB: Define Upper Age Range for children) | • Pregnant women• Women in antenatal, intranatal and postnatal periods• Newborns• Children aged 0–5 years |  |
| **I**ntervention OR Exposure  | Mobile health interventions [NB: Brainstorm list of intervention types] | Any intervention, delivered via mobile ICT, which is designed to support the maternal, newborn and child health at national, state, city, or community level in Low- and Middle-Income Countries |  |
| **C**omparison  | Usual (traditional) support to mothers and children (NB: Speak with experts to find out what this type is) | Usual care | Another recent technological intervention |
| **O**utcome(s)  | Improved maternal health, Improved child health, Body Mass Index, maternal and infant well-being | **Primary outcomes**: maternal mortality; maternal morbidity; newborn and child mortality; newborn and child morbidity**Secondary outcomes:** number of planned antenatal and postnatal visits; the number of unscheduled care visits and emergency care; quality of life; quality of care (delivery by skilled birth attendants, appropriate use of evidence-based medical and obstetric interventions where available); self-efficacy; cost-effectiveness; immunisation cover and child developmental milestones |  |
| **C**ontext | Low- and Middle-Income Countries (LMICs) | Any healthcare setting in an LMIC | • Studies from developed countries |
| **S**tudy types | Empirical comparative studies (randomised controlled trials; cohort studies) |  | • Expert opinion• Case studies, case series• Technical reports, reviews |

## Exercise 6.3: When is an article worth reading?

Before introducing quality assessment in more detail, when you come across a research study in a journal, consider the following:

1. How do you decide whether an article is worth reading?

2. What makes an article believable?

In addressing Exercise 6.3, you can probably divide your factors into three distinct groups:

● *Applicability –* the topic of the study, how important it is to you at the moment, the similarity of the setting to the one in which you work, the level at which it is written, the professional group or discipline for whom it is written, etc.

● *Extrinsic factors* – those external factors assumed to relate to, but not always associated with, the quality of the article: who wrote it, where they work, what their job or qualifications are, whether you have heard of them, who paid for the study, in which journal it is published, whether the authors have written on the subject before.

● *Intrinsic factors* – factors that relate to the study itself, i.e., the appropriateness of the study design to the question being asked, the suitability of the sample, the methods used to recruit the sample, the methods used to obtain the results.

## Exercise 6.4: Study scenario

**Can an ICT intervention improve primary schoolchildren’s performance in solving maths problems?**

A group of 20 primary school children are selected by their teacher to take part in this study as a reward for good behaviour. At the start of the study, each child completes a maths problems test consisting of 10 questions. Each child then completes a 1-hour online workbook on the class computer, in which they complete a series of maths problems in the form of fun games and exercises. The children then complete another maths problems test of 10 questions, and the number answered correctly is recorded. The teacher asks each child to recall how many they got correct in the maths test before the study started.

1. *What sources of bias can you see?*

Hint 1: Might the children selected to take part in the study differ in any way from those who did not take part?

Selection bias – well-behaved children Recall bias – can the children accurately recall their pre-study test score? Will the children be truthful about their pre-study test scores?! Investigator bias – might the teacher want the outcome of the study to go in a particular direction?

Hint 2: Do you think the children could accurately (and truthfully!) recall their score for the test they took at the beginning of the study?

1) Small sample size Confounding factors: We are told very little about the children that took part and there are likely to be confounding factors that might influence the results, for example, consider the following pre-study test score on a weekly maths test, parental help with maths homework.

2. *Aside from the sources of bias in this study, what other factors might limit the credibility of the study findings?*

Incentive: ICT intervention is a reward for good behaviour.

## Exercise 6.5: Quality assessment

|  |
| --- |
| ***CASP Randomised Controlled Trial Standard Checklist:******11 questions to help you make sense of a randomised controlled trial (RCT)*** |
| **Study and citation** | **Mogollon et al., 2014** |
| **Section A: Is the basic study design valid for a randomised controlled trial?** | **Yes/No/Can’t tell** |
| **1. Did the study address a clearly focused research question?***CONSIDER:**Was the study designed to assess the outcomes of an intervention?**Is the research question ‘focused’ in terms of:**• Population studied**• Intervention given**• Comparator chosen**• Outcomes measured?* | **Yes**“to investigate the impact of 12-week high-flavanol chocolate (HFC) consumption vs. LFC on skin sensitivity to UV radiation, measured by minimal erythema dose (MED)”... “enrolled nonsmoking healthy women aged 20 to 65 years who had normalskin types I or II, as described by Fitzpatrick” |
| **2. Was the assignment of participants to interventions randomised?***CONSIDER:**• How was randomisation carried out? Was**the method appropriate?**• Was randomisation sufficient to eliminate**systematic bias?**• Was the allocation sequence concealed**from investigators and participants?* | **Can’t tell**Randomisation method is robust -“A blocked randomisation (4) was computer-generated by a statistician who was not involved in the study. It was stratified according to skin type (I and II) and age (30–35 years; 36–49 years; 50–65years).” However, does not report how an allocation to groups was performed. |
| **3. Were all participants who entered the study accounted for at its conclusion?***CONSIDER:**• Were losses to follow-up and exclusions**after randomisation accounted for?**• Were participants analysed in the study**groups to which they were randomised**(intention-to-treat analysis)?**• Was the study stopped early? If so, what**was the reason?* | **Yes**All participants accounted for the flow diagram in Figure 1 and were analysed in groups randomised. The study was not stopped early. |
| **Section B: Was the study methodologically sound?** | **Yes/No/Can’t tell** |
| **4.** **• Were the participants ‘blind’ to the intervention they were given?****• Were the investigators ‘blind’ to the intervention they were giving to participants?****• Were the people assessing/analysing****outcome/s ‘blinded’?** | **Yes****Yes****Yes** |
| **5. Were the study groups similar at the start of the randomised controlled trial?***CONSIDER:**• Were the baseline characteristics of each**study group (e.g., age, sex, socio-economic**group) clearly, set out?**• Were there any differences between the**study groups that could affect the**outcome/s?* | **No**“At randomization, the 2 arms were well balanced with regard to social and baseline characteristics, excluding BMI which was 1.8 higher in the LFC group and season at week 12. Specifically, the percentage of participants in the HFC group who were evaluated before the spring season was lower than in the LFC group (Table 1).” |
| **6. Apart from the experimental intervention, did each study group receive the same level of care (that is, were they treated equally)?***CONSIDER:**• Was there a clearly defined study protocol?**• If any additional interventions were given**(e.g., tests or treatments), were they similar**between the study groups?**• Were the follow-up intervals the same for**each study group?* | **Yes**“During the study period, both groups tended towards decreased dietary polyphenols (not including interventional chocolate consumption). No differences were evident in changes of dietary polyphenol consumption between the HFC and the LFC groups (Additional file 1: Table S12).” |
| **Section C: What are the results?** | **Yes/No/Can’t tell** |
| **7. Were the effects of intervention reported comprehensively?***CONSIDER:**• What outcomes were measured, and were they clearly specified?**• How were the results expressed? For binary outcomes, where are relative and absolute effects reported?**• Were the results reported for each outcome in each study group at each follow-up interval?**• Was there any missing or incomplete data?**• Was there differential drop-out between the study groups that could affect the results?**• Were potential sources of bias identified?**• Which statistical tests were used?**• Were p values reported?* | **Yes**The sample size estimate and statistical methods are robust. |
| **8. Was the precision of the estimate of the intervention or treatment effect reported?***CONSIDER:**• Were confidence intervals (CIs) reported?* | **No** |
| **9. Do the benefits of the experimental intervention outweigh the harms and costs?***CONSIDER:**• What was the size of the intervention or treatment effect?**• Were harms or unintended effects reported for each study group?**• Was a cost-effectiveness analysis undertaken? (Cost-effectiveness analysis allows a comparison to be made between different interventions used in the care of the same condition or problem.)* | **Can’t tell**95% CIs reported for minimal erythema dose, but not skin elasticity and hydrationparameters |
| **Section D: Will the results help locally?** | **Yes/No/Can’t tell** |
| **10. Can the results be applied to your local population/in your context?***CONSIDER:**• Are the study participants similar to the**people in your care?**• Would any differences between your**population and the study participants alter**the outcomes reported in the study?**• Are the outcomes important to your**population?**• Are there any outcomes you would have**wanted information on that have not been**studied or reported?**• Are there any limitations of the study that**would affect your decision?* | This study was undertaken in Canada. If your setting was in a different country, or a different population (e.g., children, men), you would need to consider if the results of the study could be generalized to your context. Also, if the study has captured all of the outcomes relevant to your context. |
| **11. Would the experimental intervention provide greater value to the people in your care than any of the existing interventions?***CONSIDER:**• What resources are needed to introduce**this intervention takes into account time,**finances, and skills development or training**needs?**• Are you able to disinvest resources in one**or more existing interventions in order to**be able to re-invest in the new**intervention?* | **Can’t tell**The study evaluates the clinical effectiveness of chocolate flavanols in preventing cutaneous UV damage. There is no evaluation of the costs of treatment or resource usage reported. |
| **Appraisal summary** |  |
| *Record key points from your critical appraisal in this box. What is your conclusion about the paper? Would you use it to change your practice or to recommend changes to care/interventions used by your organisation? Could you judiciously implement this intervention without delay?* | The limitations of the study (also noted in the study discussion – non-optimal power of the study sample, the large CIs revealed by multivariate analysis, imbalance in numbers allocated to treatment groups, and differences between study groups for skin type and date of entry) indicate that the results of this study should be interpreted with caution. |
| ***CASP Cohort Study Checklist:******12 questions to help you make sense of a Cohort Study*** |
| **Study and citation** | **Bellis et al., 2012** |
| **Section A: Are the results of the study valid?** | **Yes/No/Can’t tell** |
| **1) Did the study address a clearly****focused issue?***HINT: A question can be ‘focused’**in terms of**• the population studied**• the risk factors studied**• is it clear whether the study tried to**detect a beneficial or harmful effect**• the outcomes considered* | **Yes**“Here, we examine the impact of fame on mortality in North American and European rock and pop stars. We update a previous epidemiological analysis to include more recent stars (reaching fame between 2000 and 2006) and incorporate larger numbers of older and experformers. We examine risk and protective factors for mortality in stars. For the first time, we also explore the relative contributions of ACEs (adverse childhood experiences) and other performer characteristics to cause of premature death among rock and pop stars.” |
| **2) Was the cohort recruited in an acceptable way?***HINT: Look for selection bias which might compromise the generalisability of the**findings:**• was the cohort representative of a**defined population**• was there something special about the**cohort**• was everybody included who should**have been* | **No**“With no internationally agreed definition of what constitutes a rock or pop star, we used large, established music polls to identify which individuals to include. An international poll of over 200 000 fans, experts and critics identified the all-time top 1000 albums up to the year 1999” |
| **3) Was the exposure accurately measured to minimise bias?***HINT: Look for measurement or classification bias:**• did they use subjective or objective**measurements**• do the measurements truly reflect what**you want them to (have they been**validated)**• were all the subjects classified**into exposure groups using the**same procedure* | **Yes**“Causes of death were dichotomised into ‘substance use or risk-related deaths’ (drug or alcohol-related chronic disorder, overdose or accident and other risk-related causes that may or may not have been related to substance use, i.e., suicide and violence) and ‘other’. For those who had died, ACEs were identified through the same online and published biographical sources. ACEs were taken from the WHO standardised ACE questionnaire and here included suffering as a child: (1) physical abuse; (2) sexual abuse; (3) substantive verbal abuse; living with: (4) a depressed, mentally ill, suicidal or chronically ill person; (5) a substance-abusing household member; (6) a family with an incarcerated household member; (7) a separated family or (8) domestic violence.” |
| **4) Was the outcome accurately measured to minimise bias?***HINT: Look for measurement or classification bias:**• did they use subjective or objective measurements**• do the measurements truly reflect what you want them to (have they been validated)**• has a reliable system been**established for detecting all the cases (for measuring disease occurrence)**• were the measurement methods similar in the different groups**• were the subjects and/or the outcome assessor blinded to exposure (does this matter)* | **Can’t tell**"For an objective measure of age and date of fame, we used the earliest date of first chart success (n = 1012) or date of release of earliest album included in the study (n = 477). Chart success was measured as the earliest of when an individual first appeared on an album in the Top 40 UK Official Chart (n = 636) or Top 40 US Billboard 200 (n = 239).”“Survival since becoming famous was calculated for comparison to expected survival based on general populations (matched to stars for sex, nationality, ethnicity, date of fame and age at fame).”Insufficient detail. |
| **5)** **a) Have the authors identified all-important confounding factors?***HINT:**• list the ones you think might be important, and ones the author missed* | Yes‘Further, including possible confounders (performer type, continent, ethnicity, gender, age of fame, year of birth and year reached fame) in logistic regression analysis...’ |
| **b) Have they taken account of the confounding factors in the design and/or analysis?***HINT:**• look for restrictions in design, and techniques, e.g., modelling, stratified-, regression-, or sensitivity analysis to correct,* *control or adjust for confounding factors* | **Yes**‘Further, including possible confounders (performer type, continent, ethnicity, gender, age of fame, year of birth and year reached fame) in logistic regression analysis impact of increasing ACEs on the cause of death (Wald = 8.95,*p* < 0.005; AOR 2.40; 95%CIs 1.35 to 4.25).” |
| **6)** **a) Was the follow up of subjects complete enough?**  | **Yes** |
| **b) Was the follow up of subjects long enough?***HINT: Consider**• the good or bad effects should have had long enough to reveal themselves**• the persons that are lost to follow-up may have different outcomes than those available for assessment**• in an open or dynamic cohort, was there anything special about the outcome of the people leaving, or the exposure of the people entering the* | **Yes**‘Along with the 1000 albums up to 1999, the top 30 albums each year from 2000 to 2006 were included in this study (total n = 1210), with a minimum of 5 years fame considered necessary to calculate survival.’ |
| **Section B: What are the results?** | **Yes/No/Can’t tell** |
| **7. What are the results of this study?***HINT: Consider**• what are the bottom-line results**• have they reported the rate or the proportion between the**exposed/unexposed, the ratio/rate difference**• how strong is the association between exposure and outcome (RR)**• what is the absolute risk reduction (ARR)* | European Band (Reference): N = 606, Died = 5.4%, Chi-squared 36.32, *p* < 0.001, Wald 32.21, *p* < 0.001European Solo: N = 51, Died = 9.8%, Adjusted Hazard ratio (HR) 1.12, 95%CI 0.44-2.88, Wald 0.06, *p* = 0.812North American Band: N = 718, Died = 10.2% HR 2.09, 95%CI 1.39-3.16, Wald 12.36, *p* < 0.001North American Solo: N = 114, Died = 22.8%, HR 4.24, 95%CI 2.53-7.09, Wald 30.26, *p* < 0.001 |
| **8. How precise are the results?***HINT:**• look for the range of the confidence intervals, if given* | In this retrospective cohort study, relationships between fame and mortality vary with performers’ characteristics were analyzed appropriately as adjusted hazard ratios with 95% confidence intervals |
| **9. Do you believe the results?***HINT: Consider**• big effect is hard to ignore**• can it be due to bias, chance or confounding**• are the design and methods of this study sufficiently flawed to make the results unreliable**• Bradford Hills criteria (e.g., time sequence, dose-response gradient, biological plausibility, consistency)* | **Can’t tell**In this retrospective cohort study, relationships between fame and mortality vary with performers’ characteristics were analyzed appropriately as adjusted hazard ratios with 95% confidence intervals |
| **Section C: Will the results help locally?** | **Yes/No/Can’t tell** |
| **10. Can the results be applied to the local population?***HINT: Consider whether**• a cohort study was the appropriate method to answer this question**• the subjects covered in this study could be sufficiently different from your population to cause concern**• your local setting is likely to differ much from that of the study**• you can quantify the local benefits and harms* | **No**There is a possibility that the populations included in the study, might not be representative of all rock and pop stars. |
| **11. Do the results of this study fit with other available evidence?** | The reviewer would need to consider the findings of this study, and all other studies included in the review. |
| **12. What are the implications of this study for practice?** | **Can’t tell**As a standalone study, this study would provide insufficient evidence to recommend any changes in policy or practice. Other studies included in the review would need considering. |
| ***CASP Qualitative Study Checklist:******10 questions to help you make sense of a Qualitative research*** |
| **Study and citation** | **Birks et al., 2016** |
| **Section A: Are the results valid?** | **Yes/No/Can’t tell** |
| **1. Was there a clear statement of the aims of the research?***HINT: Consider**• what was the goal of the research**• why it was thought important**• its relevance* | **Yes**“The research presented in this paper aims to contribute to the existing body of knowledge by exploring the experiences of therapists using Paro as a therapeutic tool with a more diverse group of residents in an aged care facility in regional Australia.” |
| **2. Is a qualitative methodology appropriate?***HINT: Consider**•If the research seeks to interpret or illuminate the actions and/or subjective experiences of research participants**•Is qualitative research the right methodology for addressing the research goal* | **Yes**The aim of the study was to identify, explore, and describe the impact of the use of Paro robotic seals in an aged care facility in a regional Australian city. A qualitative, descriptive, exploratory design was employed, which was appropriate. |
| **3. Was the research design appropriate to address the aims of the research?***HINT: Consider**• if the researcher has justified the research design (e.g., have they discussed how they decided which method to use)* | **Can’t tell**The rationale for using the research design employed or justification for using it is not reported in the paper. |
| **4. Was the recruitment strategy appropriate to the aims of the research?***HINT: Consider**• If the researcher has explained how the participants were selected**• If they explained why the participants they selected were the most appropriate to provide access to the type of knowledge sought by the study**• If there are any discussions around recruitment (e.g., why some people chose not to take part)* | **Yes**The setting was an aged care facility in a regional Australian city. The original intention was to recruit from a broad group of stakeholders (staff, residents, family and friends). The three recreational (diversional) therapists employed at the facility were those most involved with the use of the robotic seals, however, and it was these staff who participated in the study. |
| **5. Was the data collected in a way that addressed the research issue?***HINT: Consider**• If the setting for the data collection was justified**• If it is clear how data were collected (e.g., focus group, semi-structured interview etc.)**• If the researcher has justified the methods chosen**• If the researcher has made the methods explicit (e.g., for interview method, is there an indication of how interviews are conducted or did they use a topic guide)**• If methods were modified during the study. If so, has the researcher explained how and why**• If the form of data is clear (e.g., tape recordings, video material, notes etc.)**• If the researcher has discussed saturation of data* | **Can’t tell**“semistructured interviews were conducted with the participants for approximately one hour each. While a group interview was planned, one therapist was unavailable and so was interviewed separately from the other two. Two members of the research team undertook an inductive thematic analysis of the data obtained from the verbatim interview transcripts and therapists’ journals.”No justification for the methods chosen. No details of how interviews were conducted. |
| **6. Has the relationship between researcher and participants been adequately considered?***HINT: Consider**• If the researcher critically examined their own role, potential bias and influence during (a) formulation of the research questions (b) data collection, including sample recruitment and choice of location**• How the researcher responded to events during the study and whether they considered the implications of any changes in the research design* | **Can’t tell**No reporting on examination of potential bias or influence of the researcher during the research design or conduct. |
| **Section B: What are the results?** | **Yes/No/Can’t tell** |
| **7. Have ethical issues been taken into consideration?***HINT: Consider**• If there are sufficient details of how the research was explained to participants for the reader to assess whether ethical standards were maintained**• If the researcher has discussed issues raised by the study (e.g., issues around informed consent or confidentiality or how they have handled the effects of the study on the participants during and after the study)**• If approval has been sought from the ethics committee* | **Can’t tell**“Following approval by the university’s Human Research Ethics Committee, the therapists were asked to maintain a journal of their interactions with residents when using Paro.”However, no detail of how the research was explained to participants, consent, or confidentiality. |
| **8. Was the data analysis sufficiently rigorous?***HINT: Consider**• If there is an in-depth description of the analysis process**• If the thematic analysis is used. If so, is it clear how the categories/themes were derived from the data**• Whether the researcher explains how the data presented were selected from the original sample to demonstrate the analysis process**• If sufficient data are presented to support the findings**• To what extent contradictory data are taken into account**• Whether the researcher critically examined their own role, potential bias and influence during analysis and selection of data for presentation* | **Can’t tell**“Two members of the research team undertook an inductive thematic analysis of the data obtained from the verbatim interview transcripts and therapists’ journals. Thisanalysis was checked by a third researcher...”Insufficient detail reported |
| **9. Is there a clear statement of findings?***HINT: Consider**• If the findings are explicit**• If there is an adequate discussion of the evidence both for and against the researcher’s arguments**• If the researcher has discussed the credibility of their findings (e.g., triangulation, respondent validation, more than one analyst)**• If the findings are discussed in relation to the original research question* | **No**Little discussion of the evidence both for and against the researcher’s argument. Credibility is not discussed. |
| **Section C: Will the results help locally?** |  |
| **10. How valuable is the research?***HINT: Consider whether**• If the researcher discusses the contribution the study makes to existing knowledge or understanding (e.g., do they consider the findings in relation to current practice or policy or relevant research-based literature**• If they identify new areas where research is necessary**• If the researchers have discussed whether or how the findings can be transferred to other populations or considered other ways the research may be used* | The authors recommend “In respect of practice, it is clear that if robotic seals are to be introduced into the practice environment, it is necessary for those who will be involved in their use to be adequately prepared.” However, the limited reporting concerning Ethical conduct, the research and analysis methods, and the discussion of the findings, should be taken into consideration. |