Supplementary Files

Table S1. Major statements in manuals or articles about choosing study designs in systematic reviews

ID [Reference]	Major statements
Institutions	
AHRQ 2009[13]	Inclusion and exclusion criteria for studies to be used in the review are determined based on the specific questions to be addressed. Criteria may vary for each question in the review. Key Questions 1. Does any single NPWT system or any component of a NPWT system have a substantial therapeutic distinction in terms of wound healing outcomes compared to any other NPWT system or any other similar component of a NPWT system for the treatment of acute or chronic wounds? 2. Do patients being treated with one NPWT system have a substantial therapeutic distinction in terms of less pain, bleeding, infection, other complications, or mortality than other NPWT systems? 3. What are the reported occurrences of pain, bleeding, infection, other complications, and mortality for NPWT systems? Inclusion Criteria: Study Design: For Key Question 1 and 2, study must have been a controlled study comparing one NPWT system or components of a system to another NPWT system or components. Randomization to a NPWT system group was not required. For Key Question 3, no control group was required, because the focus of the question was simply to identify adverse events rather than compare rates across systems or components. However, because of the potential for bias in case series studies, no analyses were performed using adverse event data from these studies.
ASERNIP-S[14]	Interpretation of assessments should take into consideration the likely robustness of the evidence, as indicated by the type of study design
CADTH 2003[15]	If only a limited number of poor quality RCTs can be identified, other study designs can be considered. In some cases, even if RCTs are available, studies of other types should be reviewed by the authors, e.g. to identify long-term effectiveness and/or rare or long-term adverse effects.
CEBM 2014[5]	The type of study can generally be worked at by looking at three issues: 1. What was the aim of the study, descriptive or analytic? 2. If analytic, was the intervention randomly allocated? Yes: RCT; No: Observational study. 3. If observational, when were the outcomes determined? After the exposure: prospective (typically cohort study); at the same time as the exposure: cross sectional study; before the exposure: retrospective (typically case-control study).
Cochrane 2011[16]	Randomization is the only way to prevent systematic differences between baseline characteristics of participants in different intervention groups in terms of both known and unknown (or unmeasured) confounders. For some Cochrane reviews, the question of interest cannot be answered by randomized trials, and review authors may be justified in including non-randomized studies.

	Non-randomized studies (NRS) are defined here as any quantitative study estimating the effectiveness of an intervention (harm or benefit) that does not use randomization to allocate units to comparison groups. There are three main reasons for including NRS in a Cochrane review: 1. To examine the case for undertaking a randomized trial by providing an explicit evaluation of the weaknesses of available NRS 2. To provide evidence of the effects (benefit or harm) of interventions that cannot be randomized, or which are extremely unlikely to be studied in randomized trials. 3. To provide evidence of effects (benefit or harm) that cannot be adequately studied in randomized trials, such as long-term and rare outcomes Key issues about the inclusion of non-randomized studies in a Cochrane review • Susceptibility to selection bias (understood in this Handbook to mean differences in the baseline characteristics of individuals in different intervention groups. • Confounding occurs when selection bias gives rise to imbalances between intervention and control groups (or case and control groups in case-control studies) on prognostic factors, i.e. the distributions of the factors differ between groups AND the factors are associated with outcome. Determining which of non-randomized study designs to include (Chapter 13.2.1.3 of the Cochrane handbook): The diversity of NRS designs raises two related questions. First, should all NRS designs of a particular effectiveness question be included in a review? Second, if review authors do not include all NRS designs, what criteria should be used to decide which study designs to include and which to exclude?
CRD 2009[7]	Randomised controlled trials (RCTs) and controlled trials for the objective to assess the clinical effectiveness of treatments. If information from controlled trials is not available, cohort studies are eligible for inclusion provided that data from a comparison group are reported. In some cases a range of study designs may be needed to address different questions within the same review. For example, a review seeking to include information on adverse events will often include case-control and/or case-series whilst a review incorporating participants' experiences of an intervention is likely to include qualitative studies. In cases where it is unworkable or unethical to randomise participants (e.g. when evaluating the effects of smoking on health), researchers may instead have to use a quasi-experimental or an observational design.
HAS 2007[17]	Literature search and analysis of scientific data. A level of scientific evidence (HAS grading scheme) is allocated to each study. I- High powered randomised controlled trials, meta-analyses, decision analyses. II- Low powered randomised controlled trials, or non-randomised trials, cohort studies. III- Case-control studies. IV- Retrospective studies, case series, descriptive epidemiological studies, and controlled trials with bias.
IQWiG 2013[18]	Für die Nutzenbewertung von Interventionen ist an erster Stelle eine Kontrollgruppe zu fordern. Aus einem reinen Vorher-nachher-Vergleich in einem Design mit abhängigen Stichproben ohne Kontrollgruppe lässt sich in der Regel kein Beleg für einen Effekt einer Intervention ableiten. Gütekriterien, die die Aussagekraft kontrollierter Studien erhöhen, sind Randomisierung und Verblindung. Den ersten Informationsgewinn gibt es häufig aus Fallberichten oder Fallserien. Grundsätzlich sind prospektive Studien retrospektiven Designs vorzuziehen. Allerdings sind zum Beispiel Fall-Kontroll-Studien häufig die einzige praktikable Möglichkeit, Informationen über Zusammenhänge zwischen Expositionen und seltenen Erkrankungen zu gewinnen. Neuere Studiendesigns der modernen Epidemiologie enthalten Elemente sowohl von Fall-Kontroll-Studien als auch von Kohortenstudien

	und sind nicht mehr eindeutig als retrospektiv oder prospektiv zu klassifizieren. Zumindest im Rahmen von Therapiestudien wird der höchste Evidenzgrad RCTs und systematischen Übersichten von RCTs zugeordnet. Auf den nächsten Plätzen folgen nicht randomisierte Interventionsstudien, prospektive Beobachtungsstudien, retrospektive Beobachtungsstudien, nicht experimentelle Studien (Fallserien und Fallberichte) und – mit niedrigstem Evidenzgrad – Expertenmeinungen ohne wissenschaftliche Begründung.
MRC 2008[19]	There are many study designs to choose from, and different designs suit different questions and different circumstances.45 Awareness of the whole range of experimental and non-experimental approaches should lead to more appropriate methodological choices. You should always consider randomization when assessing effectiveness. An experimental approach may not be feasible or randomization may be unnecessary and other designs preferable. Examples where non-randomised designs have been used successfully: to ban smoking in public places, for studying rare adverse events, to evaluate the impact on the incidence of sudden infant deaths using case-control methods.
MSAC	No handbook found
NICE 2013[6]	RCTs directly comparing the technology under appraisal with relevant comparators provide the most valid evidence of relative efficacy. However, such evidence may not always be available and may not be sufficient to quantify the effect of treatment over the course of the disease. Therefore, data from non-randomised studies may be required to supplement RCT data. Any potential bias arising from the design of the studies used in the assessment should be explored and documented.
OHTAC	No handbook found
Authors	
Black 1996[20]	The view is widely held that experimental methods (randomised controlled trials) are the "gold standard" for evaluation and that observational methods (cohort and case control studies) have little or no value. This ignores the limitations of randomized trials, which may prove unnecessary, inappropriate, impossible, or inadequate.
Deeks 2003[21]	There must be hundreds of examples of interventions for which RCTs would be possible but have not yet been carried out, leaving the medical and policy community to rely on nonrandomised evidence. It is therefore essential to have an understanding of the biases that may influence non-randomised studies.
Egger 2006[22]	Etiological hypothesis, however, cannot generally be tested in randomized experiments. Due to the limited size of such trials, less common adverse effects of drugs may only be detected in case-control studies, or in analyses of databases from postmarketing surveillance schemes. Also, because follow-up is generally limited, adverse effects occurring later will not be identified. Women, the elderly, and minority ethnic groups are often excluded from randomized trials. Similarly, the university hospital typically participating in clinical trials differ from the settings where most patients are treated. Finally, both patient and therapist preferences may preclude a randomized controlled experiment. Example: ingesting your own urine. Even if adjustments for confounding factors have been made in the analysis, residual confounding remains a potentially serious problem in observational research. It is well known that people from less favourable social circumstances are more likely to be non-responders in studies evaluating patient-reported outcomes. The thorough consideration of possible sources of heterogeneity between observational study results will provide more insights than the mechanistic calculation of an overall measure of effect, which may often be biased. That different recall of past exposures may introduce bias. Sensitivity analysis is important to test the stability of findings across different study designs, different approaches to exposure ascertainment and to selection of study participants.
Guyatt 2005[23]	Randomized controlled trials to assess treatment. Observational studies to assess harm.

	Ideally, we would also look to randomized trials to address issues of harm. However, for many potentially harmful exposures, randomly allocating patients is neither practical nor ethical. For instance, one could not suggest to potential study participants that an investigator will decide by the flip of a coin whether or not they smoke during the next 20 years or whether they will be exposed to potentially harmful ionizing radiation. For exposures like smoking and radiation, the best one can do is identify studies in which personal choice, or happenstance, determines whether people are exposed or not exposed.
Khan 2011[24]	Some reviewers consider certain study designs to be superior because they feel that the design has an inherent value in itself. For example, they may focus exclusively on randomized studies when conducting reviews. Such a view ignores the fact that addressing different types of questions may require the use of different study designs. Assessment of long-term or rare outcomes, particularly when examining the safety of interventions, would be more suited to an observational design, not an experimental study. RCTs may be unethical. There may be the case that no RCT has been published. Educational interventions are often studied using a range of designs. It is improbable that rare harmful outcomes are detected by RCTs.
Reeves 2005[25]	Sources of bias in different study designs should be accounted for. Systematic reviews of NRS may be helpful in three situations.

Abbreviation. ID: Identifier; further abbreviations see Table 3