Published in final edited form as:

Semin Perinatol. 2016 October ; 40(6): 341–347. doi:10.1053/j.semperi.2016.05.004.

Recent Controversies on Comparative Effectiveness Research Investigations: Challenges, Opportunities and Pitfalls

Haresh Kirpalani, BM MSc¹, William E. Truog², Carl T. D'Angio, MD³, and M. Cotten, MD⁴

¹ Professor Pediatrics Division Neonatology, The Children's Hospital of Philadelphia at University Pennsylvania Philadelphia PA USA ; and Emeritus Professor Clinical Epidemiology McMaster University Ontario

² Professor of Pediatrics at University of Missouri Kansas City

³ Professor of Pediatrics and Medical Humanities & Bioethics, Division of Neonatology, University of Rochester School of Medicine and Dentistry, Rochester, NY, USA.

⁴ Professor Duke University

Abstract

The purpose of comparative effectiveness research (CER) is to improve health outcomes by developing and disseminating evidence-based information about which currently available interventions and practices are most effective for patients. While Randomized Controlled Trials (RCT) are the hallmark of scientific proof, when they have been used to compare interventions used in variable ways by different clinicians (Comparative Effectiveness RCTs, CER RCTs) they have at times, generated controversy. Usually the background for the CER RCT is a range of 'standard therapy' or 'standard of care'. This may have been adopted on observational data alone, pilot data. At times, such prior data may derive from populations that differ from the population in which the widely variable standard approach is being applied. We believe controversies related to these CER-RCTs result from confusing 'accepted' therapies and 'rigorously evaluated therapies'. We first define evidence based medicine, and consider how well neonatology conforms to that definition. We then contrast the approach of testing new therapies and those already existing and widely adopted, as in CER-RCTs. We next examine a central challenge in incorporating the control arm within CER, and aspects of the 'titrated' trial. We finally briefly consider some ethical issues that have arisen, and briefly discuss the wide range of neonatology practices that could be subject to CER-RCTs or alternative CER-based strategies that might inform practice in the absence of RCTs. Throughout, we emphasize the lack of awareness of the lay community, and indeed many researchers or commentators, in appreciating the wide variation of standard of care. There is a corresponding need to identify the best uses of available resources that will lead to the best outcomes for our patients. We conclude that CER is an essential methodology in modern neonatology to address many unanswered questions and test unproven therapies in newborn care.

Corresponding author H.Kirpalani.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Evidence and Standard Clinical Practice

In recent decades, clinical experts have embraced the goal to provide ‘evidence’ to guide medical practice and research. Evidence based Medicine (EBM) has become the guiding force in clinical practice and has become mandatory in the opinion of professional societies and clinical leaders. One pioneer of clinical epidemiology and EBM - David Sackett - explicitly linked the practicing clinician, the individual patient and research:

“Evidence-based medicine, whose philosophical origins extend back to mid-19th century Paris and earlier, is the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients. The practice of evidence-based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research” (1).

Sackett's broad vision of EBM is an inspiring definition of health care standard practices. This approach explicitly embraces the individual patient and her/his preferences. Admittedly, how standards of care are defined, and the levels of evidence supporting commonly use practices, vary. Although the terms “standard practice,” “standard therapy” and “standard of care” have subtle and meaningful distinctions, we will use them interchangeably in our discussions.

Commonly used standard of care practices can be based on very little evidence. They may also be affected by non-scientific considerations, such as potential legal concerns regarding harm and responsibility (2). However, recent approaches to adoption of new therapies into standard practice, nowadays often emphasize the degree of evidential-scientific rigor (3). Such an approach bases itself explicitly on the evidence-based pyramid, and prioritizes randomized evidence. In fact some go as far as to argue that in the absence of RCTs there is no standard of care. (4)

However as neonatologists, we are forced to grapple with a somber reality. This is the reality of a low number of RCTs in the neonatal population (5).

For example one specific area, has seen an excellent process to develop guidelines for care, of recommendations by the International Liaison Committee on Resuscitation (ILCOR) for neonatal care. The recommendations by ILCOR for the resuscitation of newborn infants, rest on currently available, poor level evidence. Less than 15 % of recommendations were based on RCT data and full 75% of recommendations either had no controls or were extrapolated from other populations, animals or mechanical models (6). Neonatal resuscitation is an area where many existing interventions could be compared in CER studies to advance the field. The same need occurs in almost every area of neonatal practice where there are many existing interventions that need to be compared.

So, currently, many standards of care are established without rigorous prior evidence. Even without new evidence from traditional RCT's becoming available, clinicians may adopt new approaches to existing ‘standards of care’. For example, in the case of blood transfusions, the advent of ‘new’ blood-related infections such as HIV (7); or the entity labelled as Transfusion Associated NEC) (8) have influenced practice – towards less transfusions. In

addition, randomized trials in adult and pediatric critical care suggested added risk with liberal transfusion guidelines that encourage maintenance of lower hemoglobin and hematocrit (9,10). These findings may have influenced at least some neonatal centers to adopt stricter transfusion guidelines on the basis of inferred risk.

In another example, oxygen radical disease was increasingly implicated in the etiology of many diseases of the preterm beyond the already recognized retinopathy of prematurity, pushing some centers to adopt strategies to limit supplemental oxygen exposure even in the absence of RCT data (11). Other stimuli for changing care in the absence of significant new evidence include rising recognition of the risk:benefit ratio of some therapies and of the need to consider health care costs (12). This is so especially with new expensive technologies (13). In effect, these patterns confirm that standards of care, once subject to testing, are not necessarily eternal and may warrant re-evaluation.

Ideally, changes in standards should be driven by new high-quality evidence. There is little debate that new or innovative therapies must fulfill certain imperatives of rigorous testing to conform to US Food and Drug Administration (FDA) and other agency norms before acceptance into standard of care (3). However, what is the situation for practices already embraced by clinicians, but based on low levels of evidence? Therapies prescribed in clinical practice are adopted over time from myriad sources, including historically time-enshrined ones, even if based on well-intended rationale. In neonatology, for example, many feeding regimens have passed down over the years, as ‘established’. Rarely have these therapies been subject to randomized controlled trial evaluation. While not everything prescribed can be rigorously tested, the number of such evaluations can, and should be, extended. This sentiment has gained support. Concomitantly interest has grown in evaluating not only ground-breaking ‘new therapies’, but also therapies of long-standing in the armamentarium. This awareness led to a widely accepted need for “comparative effectiveness research” (CER).

What is Comparative Effectiveness Research?

CER studies were defined by the **Institute of Medicine** in 2009 as:

“...the generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat and monitor a clinical condition, or to improve the delivery of care. The purpose of CER is to assist consumers, clinicians, purchasers and policy makers to make informed decision that will improve health care at both the individual and population levels”

(14-16)

Others have termed these studies “research on medical practice” (17) or “standard of care research” (18). We will use the more common term of CER.

By definition, CER comparisons are made between active agents or therapies already accepted in standard practice (19). It is important to note how wide the resonance has been to CER in medicine as a whole. As VanLare emphasizes, the extent to which a national agenda has been set to carry out CER is unique. The American Recovery and Reinvestment

Act appropriated \$1.1 billion to fund comparative-effectiveness research (CER) (20). Following the same impetus, an explicit mandate from the US Congress established the Patient-Centered Outcomes Research Institute (PCORI) in 2010. The mission of PCORI was to fund CER in order “to assist patients, clinicians, purchasers, and policy-makers in making informed health decisions.” (21)

What Study Design is best in CER Studies?

CER can be performed in the full range of study designs, ranging from observational to randomized studies (22). Without re-inventing the wheel, it can be most simply stated that the grades of evidence in CER-research parallel those of non-CER research (23-24). Even conclusions from large scale data-base collections are prone to an unpredictable discordance from results of RCTs (23-25).

Hence, there should be an appropriate emphasis on RCT's, if these are both feasible and possible, and when genuine uncertainty exists. It is recognized that not all possible studies can be performed and priorities are needed (25,26). At the outset the IOM laid out a top 100 goals (14), and similarly PECORI also sets out priorities (21). These reflect an increasing desire to reflect the real patient's needs. In turn an early recognition has grown and now become more widespread, that new generation of RCTs be performed in a ‘pragmatic’ manner, and will need to be large studies (26,27). In this context pragmatic refers to simple flexible designs that do not create a lot of exclusions from the trial, and enable a greater degree of generalizability of findings. While “personalized medicine” holds the promise of being able to tailor highly individualized plans of therapy, the first step in targeting many therapies is applying them across a broad range of participants to be able to evaluate variations in effectiveness at the population level. If possible, simultaneous study cohort analysis for characteristics (including ‘biomarkers’) could suggest which individuals or sub-groups may be more likely to benefit from the study intervention.

What should be the control arm in CER randomized studies?

One key principle of CER-RCTs is that there is no placebo or “no therapy” group. On the contrary, CER studies all involve comparisons of therapies considered as within accepted norms of standard practice. Considerable controversy has arisen historically over the appropriate control arm therapy to use in RCTs. In the 7th, and latest, iteration of the Declaration of Helsinki (28), the use of placebos in RCTs is restricted to exceptions, being for example where “no proven intervention exists” or where “compelling and scientifically sound methodological reasons” exist and with patients receiving placebo not being subject to any “additional risks of serious or irreversible harm” (29). This has been criticized, both from those who consider this as too lax (29) and those who consider it too restrictive (30). However, by definition, this particular aspect of trial design does not pertain to CER RCTs.

In contrast, CER RCTs test a minimum of two ‘accepted’ therapies against each other. At times, this may involve radically differing interventional pathways (e.g., a surgical therapy counterposed to a medical intervention) aimed at the same goal. More often, one particular therapy is widely believed to be needed, but differences of opinion exist on how such

therapies are to be optimally dosed or targeted. For dosing of medications, the test comparisons are sometimes straightforward – high and low dose ranges, long and short durations. The same applies for therapeutic target goals, or physiological parameters. For example ventilator tidal volumes, blood pressure, oxygen saturations, hemoglobins, heart rates etc – are expressed in a range of values along a continuum. Here trials may then devolve onto testing or target range. Such trials are termed ‘titrated trials’ (31-32).

Titration trials raise a different set of potential problems that have been termed a ‘misalignment’ between clinical goals and randomized controlled trial goals (31-33). The argument made is that if there is broad range of target values for a clinical parameter, then selecting a target range for therapy introduces sub-groups receiving differing levels of therapy. This debate erupted when the Acute Respiratory Distress Network (ARDS-NET) trial (34) was performed. The intent of ARDS-NET was to find ways to reduce mortality in ARDS, which had not been possible to achieve previously. Adult patients were randomized to either high (12 mL/kg) or low (6 mL/kg) Tidal Volume (TV) targets. Mean (SD) TV was 11.8 (0.8) mL/kg in the high target group and 6.2 (0.8) mL/kg in the low target group. A salient point to consider here is that the prevailing standard of care, appeared to be a lower TV than was prescribed in the high arm of the trial. This can be inferred from the fact that the pre-randomization mean TV was 10.3 mL/kg in the ARDS NET study and that an international survey found that the mean was 10.5 mL/kg (35-36). Thus, the high arm of the ARDS-NET trial may have been slightly outside of the standard practice continuum range, making the standard care group an inadequate comparator.

The solution proffered by some is that, in addition to two experimental arms, a true “usual practice” arm should be included, to represent the entire range of the possible interventions (31,37). However, even some of the proponents of this recognize the burdens of unfeasibility this places on the trialist. These derive in part from the large increases that would be needed in enrollment, and also therefore funding (37). In addition, since “usual practice” would likely cover a wide variety of approaches, it would be very difficult to draw conclusions from the comparison to experimental groups. Moreover, it should be acknowledged that in all science, progress is stepwise and a work in progress. Not all parts of a research question can be answered at one fell swoop. As Tennyson has it: “Science moves, but slowly, slowly, creeping on from point to point” (Tennyson, Lord A. Locksley Hall’; 1842).

As a minimum, in our view, it is reasonable to ensure that the testable targets are **within** the spectrum of ‘standard practice’. This evaluation could be performed by confirming prior to a trial what the range of practices are or by using the recommendations of authoritative bodies in the field. For instance, the PINT trial, one of two transfusion threshold trials in neonatology, tested two alternative poles of transfusion practices (38,39). However, these stayed within the recommended range as suggested by the Canadian Fetus and Newborn Committee.

Since PINT and a second neonatal randomized trial of higher versus lower transfusion thresholds found contradictory results for long-term outcomes (38-41) in secondary analyses, a larger trial was needed. This trial (The Transfusion of Prematurity (TOP) trial [ClinicalTrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT01702805) NCT01702805) adopted practices of transfusing across a spectrum of

hemoglobin values, but well within the clinically set target ranges found in a large international survey (42). Interestingly, before starting the trial, participating sites in the current Eunice Kennedy Shriver NICHD Neonatal Research Network (NRN) were surveyed to establish the low and high threshold transfusion guidelines. The lead investigators noted variation in practice wide enough among the 18 participating centers to define the two transfusion guideline arms that were within the range of current practice in this consortium of academic tertiary centers

Similarly for the oxygen targeting trial SUPPORT (43). It had long been recognized that there were very differing practices adopted by the community of neonatal physicians (11). SUPPORT, and the other international oxygen targeting trials, tested two ends of that well-accepted spectrum (43). This spectrum had been previously said by the American Academy of Pediatrics, to lie within the “pragmatically determined”, clinically acceptable range (44). When the low oxygen saturation arm of the trial unexpectedly showed increased mortality, some questioned whether the outcome could have been predicted from the information available at the outset of the trial. However others noted that prior observational data had actually suggested superiority of lower saturation targets (11). The ensuing controversy has been well ventilated in the medical and lay press, and all elements will not be reprised here (45). We echo the concluding words of the judge, who, in ruling against plaintiffs, explicitly expressed one rationale for performing prospective, randomized trials in the first place, when she said ““As the old axiom goes, correlation does not equal causation.” (46).

Another aspect of staying within an accepted continuum of practices, is that this of itself helps to ensure the feasibility of the trial. Clinicians are more likely to accept a randomization of therapy and explain it to parents-patients. This was pointed out by a leading doyen of trials – and encapsulated by the phrase “obtaining buy-in from the clinician” (47). Finally, we emphasize both internal and external validity of the trials for the interpreter of trials results, as stressed as key considerations for the user of studies (48). This was reiterated by Parshuram in the post-ARDS Net controversy, who urged examining some tests of validity when assessing the control arm in critical care studies (36).

Ethics

As previously mentioned we cannot here recapitulate the entire discussion that took place over the SUPPORT trial (43). The interested reader is pointed to cited references (45,46), and in addition - can be assured that several teams of commentators both supported (49) and, in contrast criticized (50) the SUPPORT team.

The ethical issues of CER RCT are similar to those for other RCTs. We accept as a basis the premises behind the Helsinki Declaration (28), notwithstanding some critiques (29-30). All would agree that protection of the enrolled subject, especially the vulnerable – including the neonate - is the first step. This was translated into research practices by some neonatal trialists in the following manner: “A skilled research coordinator is able to support the family while helping them to understand the risks and benefits of trial participation, and to make a decision that is best for their child. This is possible when the consent process is grounded in three principles: honesty, trust, and respect “ (51,52).

Moving beyond these affirmations, the first statement under this subheading should be that no trial should be performed if there is certainty about the risk to benefit ratio of a therapy. This applies both on a population basis and an individual basis. If either of the approaches to be evaluated in a trial, has already been shown conclusively to be superior, then it is reasonable to perform a trial comparing the two. We note, parenthetically, that many trials will establish the superiority of one therapy over another, but that this differs substantially both ethically and practically from foreknowledge.

Regarding individual risk-to-benefit calculations, Peto stated:

“A patient should not be entered if the responsible clinician or the patient are for any medical or non-medical reasons reasonably certain that one of the treatments that might be allocated would be inappropriate for this particular individual (in comparison with either no treatment or some other treatment that could be offered to the patient in or outside the trial)”(53).

Therefore, it is crucial when proposing to perform a research study to start with the certainty or uncertainty of understanding potential therapeutic efficacy (1,26,51,53)). As we have noted, if there is an undoubted certainty of benefit in one arm, especially if this includes a superior risk:benefit ratio, no trial should be conducted to verify the already proven. For example, the use of positive pressure ventilation in respiratory failure even in the early days, became rapidly ‘proven’ because of the high risk of death otherwise. A large confirmatory RCT was not performed. However such situations – where a valid decision is made to pick up a new therapy on the basis of no RCT data - are very rare in medicine. Interestingly, a Cochrane meta-analysis of large trials show that continuous positive pressure ventilation results in less death or bronchopulmonary dysplasia than the commonly used positive pressure ventilation which had been considered the standard of care (54). A neonatal-specific framework shows further concrete examples to help consider when to do a trial (51). This can be summarized in the dictum - “that if you know from evidence what needs to be done do it – do not study it!”

Moving beyond these two fundamental principles, we recognize that there are a range of views. For example, some ethicists highlight a stark difference between RCTs and clinical medicine as follows: “The RCT is not a form of personal therapy selected for particular patients” (37). However, this is often an artificial construct if the clinician does not know how or when to apply a therapy – for example, transfusions for preterms - in the first place (42). Further, there is little evidence that physicians providing standard care produce clinical results superior to those of clinical trials. Foglia examined the outcomes of 5389 infants eligible for one or more randomized trials in the NRN. Of these, 3795 were enrolled in at least one trial, and 1594 were not enrolled in any trials. There was no difference in a composite outcome of death or significant morbidity between groups (68% enrolled vs 69% eligible but not enrolled, adjusted P 0.29), nor was there a difference in component outcomes or composite outcome for each trial evaluated (55).

Many parents whose children are involved in research understand the core issues of the balance between participant safety and need for research (56). Zupancic et al found that most parents “strongly agreed that ‘It is important for children to take part in research

because what doctors learn may help other children” (57). Nonetheless, parents of subjects approached to participate in CER studies, should understand that infants enrolled to one arm of a CER-RCT could have a worse (or better) outcome than those enrolled in the other. They should also be informed that the study arm a baby is enrolled in, may not be the local practitioners’ usual choice. Since this of itself may reduce physician-driven variation, it may actually be appreciated by some parents. This has thus far, been our experience in the TOP trial.

Potential Path Forward

Nonetheless, ethicists, trialists and public lobbyists have strong views, and disagreements regarding the conduct and ethics of CER RCT, which no doubt will persist. Can we ever resolve some of the remaining disputes? They are so entangled in a tight knot, that possibly only an Olympian solution can untangle them. Alexander the Great is reputed to have cut the Gordian knot with his sword thereby resolving who was to be king of Asia (58). Do we have such a weapon to disentangle our modern Gordian knot?

We propose the relevant sword in this never-ending debate, is that of soliciting societal views, in a deliberative democratic discussion transcending select pressure groups (59-61). The ‘public democratic deliberation’ has been used in health policy discussions with broad ranges of both lay people, and informed consumers (61-63), and robust scientific methodologies have been developed (60). It has been used by workers in an array of fields of health policy, including pathology testing capabilities (64), setting large-scale epidemiology priorities for communities (65), cancer screening (66), prostate cancer screening (67), and engaging minorities in the health care agenda (68).

The sharp debate prompted by recent trials, has pitted many protagonists against each other. Each might be said to have a narrow interest. Only a broader societal debate can enable a rational synthesis. For example, societal discussion of the limited evidence for many neonatal practices might reveal similar insights similar those uncovered by Carman:

“Although participants perceived evidence as being essential to high-quality care, they also believed that personal choice or clinical judgment could trump evidence. They viewed doctors as central figures in discussing evidence with patients and key arbiters of whether to follow evidence in individual cases. They found evidence of harm to individuals or the community to be more compelling than evidence of effectiveness.” (69)

In a small way, many Institutional Review Boards and the NRN have already acknowledged the need and power of having parental-consumer representation in their discussions. Many Institutional review boards already include lay members. However, the process of policy-making in the sense of using, for example a citizens jury (61), is much more wide-ranging and inclusive. These strategies need to be incorporated into helping the parental consumer understand the limits of knowledge that the neonatologist faces in making any decision for the newborn infant. “Society” – and the parents neonatologists serve – need to understand that landscape. We believe that it is only in that way that the enterprise of prospective

comparative research can become a true collaboration between parents and clinicians, trialists and ethicists.

Conclusion

To summarize some main points we will revert to the “**PICO**” question, a staple in our teaching to novices in clinical trials. So the **P** (Population) must reflect a particular population, but the inclusion and exclusion criteria must allow generalization to the group of infants to whom the **I** (Intervention) will be applied to. To further ensure generalizability, the intervention must be relevant to that population. This is especially a consideration for titrated CER trials testing different target ranges. In these cases, generalizability is enhanced when pre-testing has assured that the **C** (Comparison) uses target ranges that are encompassed in the continuum of routinely used targets. Finally, the **O** (Outcome) must be a clinically relevant outcome, which may be ascertained both from a wide range of clinicians and from parent-consumers or a broader lay population. Ethical dilemmas are minimized by attention to standard ethical principles, but might potentially be further resolved through innovative strategies of deliberative democratic discussions with a broad array of societal representatives.

References

1. Sackett DL. Evidence-based medicine. *Semin Perinatol*. Feb; 1997 21(1):3–5. [PubMed: 9190027]
2. Strauss DC, Meirion TJ. What Does the Medical Profession Mean By “Standard of Care?”. *Journal Of Clinical Oncology*. 2009; 27:e192–3. [PubMed: 19770363]
3. Foglia EE, DeMauro SB, Dysart K, Kirpalani H. When has enough evidence accumulated to change neonatal practice? *Semin Fetal Neonatal Med*. Dec; 2015 20(6):424–30. [PubMed: 26441032]
4. Steinbrook R. Trial design and patient safety: the debate continues. *N Engl J Med*. 2003; 349:629–630. [PubMed: 12917298]
5. DeMauro SB, Giaccone A, Kirpalani H, Schmidt B. Quality of reporting of neonatal and infant trials in high-impact journals. *Pediatrics*. Sep; 2011 128(3):e639–44. [PubMed: 21859916]
6. Foglia EE, Owen LS, Kirpalani H. Delivery room research: when does poor quality evidence become an ethical issue? *Pediatrics*. May.2015 135(5):e1368.
7. Curran JW, Lawrence DN, Jaffe H, Kaplan JE, Zyla LD, Chamberland M, Weinstein R, Lui KJ, Schonberger LB, Spira TJ, et al. Acquired immunodeficiency syndrome (AIDS) associated with transfusions. *N Engl J Med*. Jan 12; 1984 310(2):69–75. [PubMed: 6606780]
8. Kirpalani H, Zupancic JA. Do transfusions cause necrotizing enterocolitis? The complementary role of randomized trials and observational studies. *Semin Perinatol*. Aug; 2012 36(4):269–76. [PubMed: 22818547]
9. Lacroix J, Hébert PC, Hutchison JS, Hume HA, Tucci M, Ducruet T, Gauvin F, Collet JP, Toledano BJ, Robillard P, Joffe A, Biarent D, Meert K, Peters MJ. TRIPICU Investigators; Canadian Critical Care Trials Group; Pediatric Acute Lung Injury and Sepsis Investigators Network. Transfusion strategies for patients in pediatric intensive care units. *N Engl J Med*. Apr 19; 2007 356(16):1609–19. [PubMed: 17442904]
10. Hébert PC1, Wells G, Blajchman MA, Marshall J, Martin C, Pagliarello G, Tweeddale M, Schweitzer I, Yetisir E. A multicenter, randomized, controlled clinical trial of transfusion requirements in critical care. Transfusion Requirements in Critical Care Investigators, Canadian Critical Care Trials Group. *N Engl J Med*. Feb 11; 1999 340(6):409–17. [PubMed: 9971864]
11. Tin, Win. Oxygen therapy: 50 years of uncertainty. *Pediatrics*. 2002; 110(3):615–6. [PubMed: 12205267]
12. Dukhovny D, Pursley DM, Kirpalani HM, Horbar JH, Zupancic JA. Evidence, Quality, and Waste: Solving the Value Equation in Neonatology. *Pediatrics*. Mar; 2016 137(3):1–10.

13. Makarov DV, Li H, Lepor H, Gross CP, Blustein J. Teaching Hospitals and the Disconnect Between Technology Adoption and Comparative Effectiveness Research: The Case of the Surgical Robot. *Med Care Res Rev.* Mar 30.2016 PMID: 27034439.
14. Institute of Medicine. Initial National Priorities for Comparative Effectiveness Research. Natl. Acad. Press; Washington, DC: 2009. p. 41<http://www.nap.edu/catalog/12648.html>
15. Sox HC, et al. CER. A report from the IOM. *Ann Intern Med.* 2009; 151:203–05. [PubMed: 19567618]
16. Conway PH, Clancy C. Comparative-Effectiveness Research — Implications of the Federal Coordinating Council's Report. *NEJM.* 2009; 361:328–330. [PubMed: 19567829]
17. Cho MK, Magnus D, Wilfond BS. Informed Consent for Research on Medical Practices. *Ann Intern Med.* Nov 3; 2015 163(9):725–6.
18. <http://www.hhs.gov/ohrp/regulations-and-policy/requests-for-comments/index.html>
19. Sox HC, Helfand M, Grimshaw J, Dickersin K, PLoS Medicine Editors. Tovey D, Knottnerus JA, Tugwell P. Comparative effectiveness research: challenges for medical journals. *Am J Manag Care.* May 1; 2010 16(5):e131–3. [PubMed: 20455639]
20. VanLare JM, Conway PH, Sox HC. Five next steps for a new national program for comparative-effectiveness research. *N Engl J Med.* Mar 18; 2010 362(11):970–3. [PubMed: 20164480]
21. Selby JV, Forsythe L, Sox HC. Stakeholder-Driven Comparative Effectiveness Research An Update From PCORI. *JAMA.* 2015; 314:2235–6. [PubMed: 26595059]
22. Sox HC, Goodman SN. The Methods of Comparative Effectiveness Research. *Annu. Rev. Public Health.* 2012; 33:425–45. [PubMed: 22224891]
23. Hemkens LG, Despina G, Contopoulos-Ioannidis, Ioannidis JPA. Routinely collected data and comparative effectiveness evidence: promises and limitations. *CMAJ.* Feb 16.2016 PMID: 26883316).
24. Kitsios GD, Dahabreh IJ, Callahan S, Paulus JK, Campagna AC, Dargin JM. Can We Trust Observational Studies Using Propensity Scores in the Critical Care Literature? A Systematic Comparison With Randomized Clinical Trials. *Crit Care Med.* Sep; 2015 43(9):1870–9. [PubMed: 26086943]
25. Hemkens LG, Despina G, Contopoulos-Ioannidis, Ioannidis JPA. Routinely collected data and comparative effectiveness evidence: promises and limitations. *CMAJ.* Feb 16.2016 PMID: 26883316.
26. Peto R, Baigent C. Trials: the next 50 years. Large scale randomised evidence of moderate benefits. *BMJ.* Oct 31; 1998 317(7167):1170–1. [PubMed: 9794846]
27. Chalkidou K, et al. The role for pragmatic RCTs (pRCTs) in CER. *Clinical Trials.* 2012; 9:436–446. [PubMed: 22752634]
28. World Medical Association. Declaration of Helsinki Ethical Principles for Medical Research Involving Human Subjects. *JAMA.* 2013; 310(20):2191–2194. [PubMed: 24141714]
29. Bertele' V, Banzi R, Gluud C, Garattini S. EMA's reflection on placebo does not reflect patients' interests. *Eur J Clin Pharmacol.* May; 2012 68(5):877–9. [PubMed: 22134525]
30. Millum J, Wendler D, Emanuel EJ. The 50th anniversary of the Declaration of Helsinki: progress but many remaining challenges. *JAMA.* Nov 27; 2013 310(20):2143–4. [PubMed: 24141885]
31. Deans KJ, Minneci PC, Eichacker PQ, Natanson C. Defining the standard of care in randomized controlled trials of titrated therapies. *Curr Opin Crit Care.* Dec; 2004 10(6):579–82. [PubMed: 15616405]
32. Deans KJ, Minneci PC, Danner RL, Eichacker PQ, Natanson C. Practice misalignments in randomized controlled trials: Identification, impact, and potential solutions. *Anesth Analg.* Aug; 2010 111(2):444–50. [PubMed: 19820238]
33. Miller FG, Silverman HJ. The ethical relevance of the standard of care in the design of clinical trials. *Am J Respir Crit Care Med.* 2004; 169:562–564. [PubMed: 14701713]
34. Acute Respiratory Distress Syndrome Network (ARDS NET). Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. *N Engl J Med.* 2000; 342:1301–1308. [PubMed: 10793162]

35. Esteban A, Anzueto A, Alia I, Gordo F, Apezteguia C, Palizas F, Cide D, Goldwaser R, Soto L, Buggedo G, et al. How is mechanical ventilation employed in the intensive care unit? An international utilization re-view. *Am J Respir Crit Care Med*. 2000; 161:1450–1458. [PubMed: 10806138]
36. Parshuram CS1, Kavanagh BP. *Am J Respir Crit Care Med*. 2004; 170(3):223–6. Positive clinical trials: understand the control group before implementing the result. [PubMed: 15151921]
37. Silverman HJ, Miller FG. Control group selection in critical care randomized controlled trials evaluating interventional strategies: An ethical assessment. *Crit Care Med*. 2004; 32(3):852–7. [PubMed: 15090973]
38. Kirpalani H, Whyte RK, Andersen C, Asztalos EV, Heddle N, Blajchman MA, Peliowski A, Rios A, LaCorte M, Connelly R, Barrington K, Roberts RS. The Premature Infants in Need of Transfusion (PINT) study: a randomized, controlled trial of a restrictive (low) versus liberal (high) transfusion threshold for extremely low birth weight infants. *J Pediatr*. Sep; 2006 149(3):301–307. [PubMed: 16939737]
39. Whyte RK, Kirpalani H, Asztalos EV, Andersen C, Blajchman M, Heddle N, LaCorte M, Robertson CM, Clarke MC, Vincer MJ, Doyle LW, Roberts RS. PINTOS Study Group. Neurodevelopmental outcome of extremely low birth weight infants randomly assigned to restrictive or liberal hemoglobin thresholds for blood transfusion. *Pediatrics*. Jan; 2009 123(1): 207–13. [PubMed: 19117884]
40. Bell EF, Strauss RG, Widness JA, Mahoney LT, Mock DM, Seward VJ, Cress GA, Johnson KJ, Kromer IJ, Zimmerman MB. Randomized trial of liberal versus restrictive guidelines for red blood cell transfusion in preterm infants. *Pediatrics*. Jun; 2005 115(6):1685–91. [PubMed: 15930233]
41. Nopoulos PC, Conrad AL, Bell EF, Strauss RG, Widness JA, Magnotta VA, Zimmerman MB, Georgieff MK, Lindgren SD, Richman LC. Long-term outcome of brain structure in premature infants: effects of liberal vs restricted red blood cell transfusions. *Arch Pediatr Adolesc Med*. May; 2011 165(5):443–5. [PubMed: 21199970]
42. Guillén U, Cummings JJ, Bell EF, Hosono S, Frantz AR, Maier RF, Whyte RK, Boyle E, Vento M, Widness JA, Kirpalani H. International survey of transfusion practices for extremely premature infants. *Semin Perinatol*. Aug; 2012 36(4):244–7. [PubMed: 22818544]
43. Carlo WA, Finer NN, Walsh MC, et al. SUPPORT Study Group of the Eunice Kennedy Shriver NICHD Neonatal Research Network. Target ranges of oxygen saturation in extremely preterm infants. *N Engl J Med*. 2010; 362(21):1959–1969. [PubMed: 20472937]
44. American Academy of Pediatrics and American College of Obstetricians and Gynecologists. Clinical considerations in the use of oxygen. 6th ed.. American Academy of Pediatrics; Elk Grove Village, IL: 2007. Guidelines for perinatal care; p. 259-62.
45. Lantos JD, Feudtner C. SUPPORT and the Ethics of Study Implementation: Lessons for Comparative Effectiveness Research from the Trial of Oxygen Therapy for Premature Babies. *Hastings Cent Rep*. Jan-Feb;2015 45(1):30–40. [PubMed: 25530316]
46. Lantos JD. Vindication for SUPPORT. *N Engl J Med*. Oct 8; 2015 373(15):1393–5. [PubMed: 26332442]
47. Sackett DL. Why randomized controlled trials fail but needn't: 1. Failure to gain “coal-face” commitment and to use the uncertainty principle. *CMAJ*. May 2; 2000 162(9):1311–4. [PubMed: 10813014]
48. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II How to sue an article about thepray or prevention. A. Are the results of the study valid? Evidence-based Medicine WOrkign Group. *JAMA*. 1993; 270:2598–2601. [PubMed: 8230645]
49. Wilfond BS, Magnus D, Antommaria AH, Appelbaum P, Aschner J, Barrington KJ, Beauchamp T, Boss RD, Burke W, Caplan AL, Capron AM, Cho M, Clayton EW, Cole FS, Darlow BA, Diekema D, Faden RR, Feudtner C, Fins JJ, Fost NC, Frader J, Hester DM, Janvier A, Joffe S, Kahn J, Kass NE, Kodish E, Lantos JD, McCullough L, McKinney R Jr, Meadow W, O'Rourke PP, Powderly KE, Pursley DM, Ross LF, Sayeed S, Sharp RR, Sugarman J, Tarnow-Mordi WO, Taylor H, Tomlinson T, Truog RD, Unguru YT, Weise KL, Woodrum D, Youngner S. The OHRP and SUPPORT. *N Engl J Med*. 2013 DOI: 10.1056/NEJMc1307008).
50. Macklin R, Shepherd L, Dreger A, Asch A, Baylis F, Brody H, Churchill LR, Coleman CH, Cowan E, Dolgin J, Downie J, Dresser R, Elliott C, Epright MC, Feder EK, Glantz LH, Grodin MA,

- Hoffman W, Hoffmaster B, Hunter D, Iltis AS, Kahn JD, King NM, Kraft R, Kukla R, Leavitt L, Lederer SE, Lemmens T, Lindemann H, Marshall MF, Merz JF, Miller FH, Mohrmann ME, Morreim H, Nass M, Nelson JL, Noble JH Jr, Reis E, Reverby SM, Silvers A, Sousa AC, Spece RG Jr, Strong C, Swazey JP, Turner L. The OHRP and SUPPORT--another view. *N Engl J Med*. Jul 11.2013 369(2):e3. doi: 10.1056/NEJMc1308015. Epub 2013 Jun 26. [PubMed: 23803134]
51. DeMauro SB, Foglia EE, Schmidt B. The ethics of neonatal research: A trialists' perspective. *Semin Fetal Neonatal Med*. Dec; 2015 20(6):431–5. [PubMed: 26394826]
 52. DeMauro SB, Cairnie J, D'Illario J, Kirpalani H, Schmidt B. Honesty, trust, and respect during consent discussions in neonatal clinical trials. *Pediatrics*. Jul; 2014 134(1):e1–3. [PubMed: 24913792]
 53. Peto R, Baigent C. Trials: the next 50 years. *BMJ*. 1998; 317:1170. [PubMed: 9794846]
 54. Schmölder GM, Kumar M, Pichler G, Aziz K, O'Reilly M, Cheung PY. Noninvasive versus invasive respiratory support in preterm infants at birth: systematic review and meta-analysis. *BMJ*. Oct 17.2013 347:f5980. [PubMed: 24136633]
 55. Foglia EE, Nolen TL, DeMauro SB, Das A, Bell EF, Stoll BJ, et al. Short-term outcomes of infants enrolled in randomized clinical trials vs those eligible but not enrolled. *JAMA*. 2015;313, 2377e9.
 56. Nunes AT, Trahms C, D'Angio CT. Informed consent for research: a cross-sectional survey on the views of parents of sick newborns. *IRB: Ethics & Human Research*. 2015; 37(6):9–14. [PubMed: 26783592]
 57. Zupancic JA, Gillie P, Streiner DL, Watts JL, Schmidt B. Determinants of parental authorization for involvement of newborn infants in clinical trials. *Pediatrics*. 1997; 99:E6.
 58. Lane Fox, Robin. Alexander the Great. 1973:149–51.
 59. Putnam, R. Making democracy work: Civic traditions in modern Italy. Princeton University Press; New Jersey: 1993.
 60. Abelson JI, Forest PG, Eyles J, Smith P, Martin E, Gauvin FP. Deliberations about deliberative methods: issues in the design and evaluation of public participation processes. *Soc Sci Med*. Jul; 2003 57(2):239–51. [PubMed: 12765705]
 61. Street J, Duszynski K, Krawczyk S, Braunack-Mayer A. The use of citizens' juries in health policy decision-making: a systematic review. *Soc Sci Med*. May.2014 109:1–9. [PubMed: 24657639]
 62. Carman KL, Mallery C, Maurer M, Wang G, Garfinkel S, Yang M, Gilmore D, Windham A, Ginsburg M, Sofaer S, Gold M, Pathak-Sen E, Davies T, Siegel J, Mangrum R, Fernandez J, Richmond J, Fishkin J, Chao Siu. Effectiveness of public deliberation methods for gathering input on issues in healthcare: Results from a randomized trial. *A.Soc Sci Med*. May.2015 133:11–20. [PubMed: 25828260]
 63. Degeling C, Stacy M, Carter, Lucie Rychetnik. Which public and why deliberate? e A scoping review of public deliberation in public health and health policy research. *Social Science & Medicine*. 2015; 131:114e121. [PubMed: 25770463]
 64. Street, Jackie M., PhD, BSc(Hons); Dip, Grad, PHC1; Callaghan, Peta; Braunack-Mayer, Annette J., PhD, BMedSci(Hons)1; Hiller, Janet E. Citizens' Perspectives on Disinvestment from Publicly Funded Pathology Tests: A Deliberative Forum.
 65. McWhirter, Rebekah E., 1,2; Critchley, Christine R., 2,3; Nicol, Dianne, 1,2; Chalmers, Don, 1,2; Whitton, Tess, 1,2; Otlowski, Margaret, 2; Burgess, Michael M., 4; Dickinson, Joanne L. Community Engagement for Big Epidemiology: Deliberative Democracy as a Tool. *J. Pers. Med*. 2014; 4:459–474. [PubMed: 25563457]
 66. Rychetnik L, Carter SM, Abelson J, Thornton H, Barratt A, Entwistle VA, Mackenzie G, Salkeld G, Glasziou P. Enhancing citizen engagement in cancer screening through deliberative democracy. *J Natl Cancer Inst*. Mar 20; 2013 105(6):380–6. [PubMed: 23378639]
 67. Thomas R, Glasziou P, Rychetnik L, Mackenzie G, Gardiner R, Doust J. Deliberative democracy and cancer screening consent: a randomised control trial of the effect of a community jury on men's knowledge about and intentions to participate in PSA screening. *BMJ Open*. Dec 24.2014 4(12):e005691.
 68. Olson M, Cottoms N, Sullivan G. Engaging Underrepresented Minorities in Research: Our Vision for a "Research-Friendly Community". *Prog Community Health Partnersh*. 2015; 9(4):595–8. Winter. [PubMed: 26639386]

69. Carman KL, Maurer M, Mangrum R, Yang M, Ginsburg M, Sofaer S, Gold MR, Pathak-Sen E, Gilmore D, Richmond J, Siegel J. Understanding An Informed Public's Views On The Role Of Evidence In Making Health Care Decisions. *Health Affairs*. 2016; 35(4):566–574. [PubMed: 27044953]