COMMENTARY

Negative spillover due to constraints on care delivery: a potential source of bias in pragmatic clinical trials

Sean Mann[1]*

Abstract

Background Pragmatic clinical trials evaluate the effectiveness of health interventions in real-world settings. Negative spillover can arise in a pragmatic trial if the study intervention affects how scarce resources are allocated across patients in the intervention and comparison groups.

Main body Negative spillover can lead to overestimation of treatment effect and harm to patients assigned to usual care in trials of diverse health interventions. While this type of spillover has been addressed in trials of social welfare and public health interventions, there is little recognition of this source of bias in the medical literature. In this commentary, I examine what causes negative spillover and how it may have led clinical trial investigators to overestimate the effect of patient navigation, AI-based physiological alarms, and elective induction of labor. Trials discussed here are a convenience sample and not the result of a systematic review. I also suggest ways to detect negative spillover and design trials that avoid this potential source of bias.

Conclusion As new clinical practices and technologies that affect care delivery are considered for widespread adoption, well-designed trials are needed to provide valid evidence on their risks and benefits. Understanding all sources of bias that could affect these trials, including negative spillover, is a critical part of this effort. Future guidance on clinical trial design should consider addressing this form of spillover, just as current guidance often discusses bias due to lack of blinding, differential attrition, or contamination.

Keywords Clinical trial, Study design, Bias, Health system capacity, Resource utilization, Patient safety

Background

Pragmatic clinical trials are used to evaluate the effectiveness of health interventions in real-world settings. In these trials, patients, providers, facilities, or other groups are randomly assigned to receive either usual care or an intervention [1]. Randomization serves to produce balanced groups of patients, so that differences in study outcomes can be attributed solely to the effect of the intervention [2]. Randomized trials are considered the "gold standard" of evidence and their results can lead to changes in clinical practice [3]. Yet, one aspect of these trials' real-world settings is often overlooked and could lead to biased results.

Clinical resources in real-world settings are sometimes scarce, meaning that their supply is insufficient to meet the needs of all patients [4–6]. During a trial, if an intervention increases patients' utilization of scarce medical resources (such as specialist appointments or hospital beds), this can negatively affect availability of care for patients in the control group. This problem is referred to by economists as "negative spillover" or "crowding out" and is sometimes addressed in trials of social welfare and

---

*Correspondence:
Sean Mann
smann@rand.org
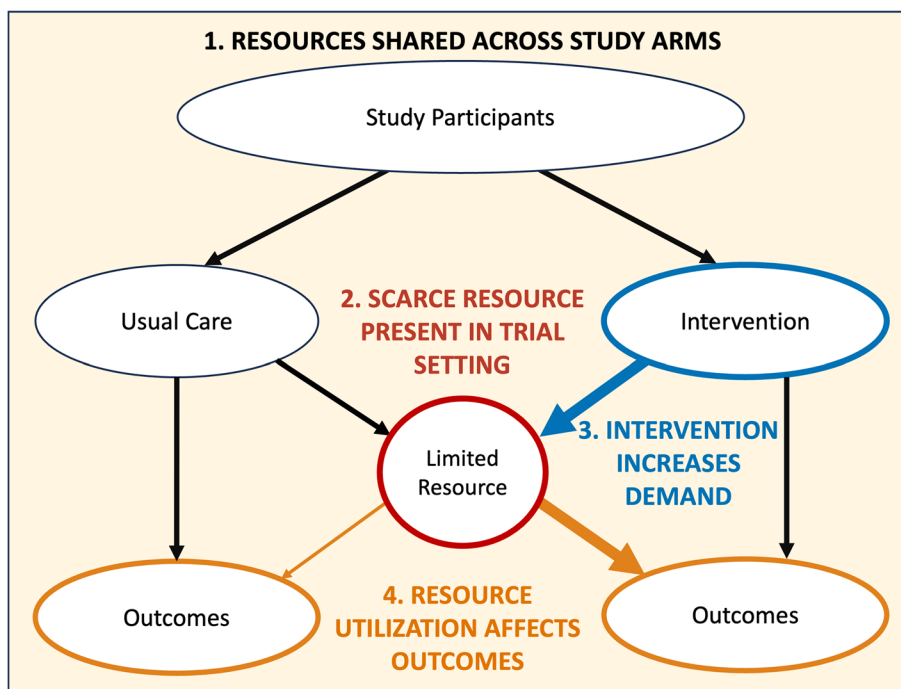[1] RAND, 1776 Main St, Santa Monica, CA 90401, USA

Fig. 1 Conditions that lead to negative spillover. There are four conditions that lead to negative spillover, each of which concerns a different portion of the trial pathway. All these conditions relate to the presence of a scarce resource that represents a constraint on care delivery. The four conditions are (1) treatment and control arms draw from a shared resource pool, (2) the presence of a scarce resource in the trial setting (red circle), (3) the study intervention increases resource demand (thick blue arrow), and (4) resource utilization affects study outcomes (thin orange arrow on left, thick orange arrow on right)

public health interventions [7–10]. However, there is little recognition of this potential source of bias in the medical literature. As a result, some clinical trials may have unknowingly overestimated the effect of a diverse set of health interventions.

For example, clinical trials evaluating patient navigation have randomly assigned patients to receive help obtaining clinical appointments. One such trial found that navigation led these patients to receive earlier diagnoses than comparison patients assigned to usual care [11]. However, the supply of appointments in the trial setting was likely constrained [12]; navigation's effect may have been due to its directing scarce appointment slots towards the treatment group at the expense of the comparison group. If so, navigation might have no effect if provided equally to all eligible patients, even in the same health system in which the trial was conducted.

This problem resembles another better-known source of bias: lack of blinding can also result in different levels of care provided to treatment and comparison groups [13, 14]. Yet, even a study that is double-blinded, in which trial subjects and clinicians are unaware of group assignment, could suffer from spillover if an intervention, by its very nature, affects use of a scarce resource shared across study arms.

In this commentary, I examine how sharing scarce resources can lead to negative spillover and how it might affect outcomes in trials of patient navigation, elective induction of labor, and physiological alarms. I also suggest ways to detect negative spillover and design trials that avoid this potential source of bias.

Sharing scarce resources can lead to negative spillover

Four conditions are necessary for scarcity-related spillover to affect trial results, as shown in Fig. 1. Each of these conditions concerns a different aspect of a study: design, setting, intervention, and outcomes.

Condition 1. Resources are shared by intervention and comparison groups

Group assignment is a key aspect of trial design. In many trials, individual patients are randomly assigned to either the intervention or comparison group. When these patients are all served by the same clinicians or hospital, the care they receive draws upon a shared pool of medical resources.

Condition 2. A scarce resource is present in the trial setting

Scarcity is most apparent during health emergencies or in cases of "chronically limited" tangible items such as organs for transplantation [4, 6]. Scarcity is less visible when it concerns an intangible resource and is an accepted part of clinical practice.

Clinician availability is a primary example of a resource whose scarcity is often overlooked [4, 5]. In primary care, intensive care, and other health settings, clinicians must routinely decide how much time to spend caring for one patient rather than another [15, 16].

Condition 3. The study intervention affects utilization of the scarce resource

Pragmatic trials often study the effectiveness of interventions that influence how care is delivered to patients [1, 17]. The intervention under investigation might be a new health information technology, care coordination process, or medical procedure. Compared to usual care, these interventions can vary in their utilization of clinical resources. A new surgical procedure might take longer and thus require greater use of an operating room. A new approach to managing chronic conditions might involve more visits to specialists. Study interventions may affect demand for care resources in multiple and sometimes unanticipated ways.

Condition 4. Availability of the scarce resource affects study outcomes

The availability of resources such as clinical staff or intensive care unit (ICU) beds can affect patient outcomes such as mortality, satisfaction with care, hospital length of stay, or the time it takes to receive a definitive diagnosis [18–21]. Such outcomes, which are important to patients and relevant to clinical decision-making, often serve as study endpoints in pragmatic trials [1, 17].

Negative spillover may have affected a diverse set of studies

Trials evaluating a diverse set of health interventions may have met the four conditions that lead to negative spillover, as shown in Table 1. This table is a partial list only; it is provided to illustrate how a range of trials may have been affected and is not the result of a systematic search or literature review.

Patient navigation

Patient navigation for cancer diagnosis is designed to help patients schedule and receive diagnostic services following abnormal screening results. One randomized trial reported that navigation helped patients obtain diagnostic resolution earlier and at higher rates for breast and colorectal cancers [11]. Another randomized trial reported similar results for breast cancer and additionally found that patients without cancer assigned to navigation received faster diagnostic resolution than patients with cancer assigned to usual care [22]. Both trials have been held up as among the best evidence in support of patient navigation in systematic reviews [48, 49].

In both trials, patients in the intervention and control arms received care in the same hospital or health system. Neither study reported on availability of diagnostic specialists during the trial. Yet, concerns over appointment availability for cancer diagnostic services are common [19], and wait times may lengthen due to increased demand [23] or fewer available staff [50]. Navigation might have helped patients obtain more appointments with specialists at earlier dates; this would be consistent with the trials' findings. Greater utilization of appointments by patients assisted by navigators could have reduced availability of appointments for others and led to delays in diagnostic resolution for patients assigned to usual care.

Induction of labor

A multicenter trial randomly assigned low-risk pregnant women to labor induction at 39 weeks gestation or to usual care in the form of expectant management. The study found that induction of labor did not increase the frequency of adverse perinatal outcomes [27]. The American College of Obstetricians and Gynecologists issued a clinical practice advisory stating that "based on the findings demonstrated in this trial, it is reasonable for obstetricians and health-care facilities to offer elective induction of labor to low-risk nulliparous women at 39 weeks gestation." [51].

During the trial, patients in the induction and expectant management groups received care in the same hospitals. Labor and delivery rooms were not always available, which affected some patients' care [52]. The median patient in the induction group stayed over 40% longer in the labor and delivery unit, which could have reduced clinicians' availability to care for patients assigned to expectant management. Given evidence that higher clinician workload negatively affects perinatal outcomes [30, 31], negative spillover could have biased trial findings [53].

Physiological alert systems

Current guidance on sepsis management notes growing interest in algorithm-based physiological alert systems that could support timely treatment of sepsis [54, 55].

Table 1 Trials in which conditions were potentially met for negative spillover*

Study	1. Resource shared across study arms	2. Scarce resource present in study setting	3. Study intervention affects resource use	4. Study outcome affected by resource availability
Randomized trial of patient navigation for urban minority women with abnormal mammograms [22]	Patients individually randomized to patient navigation or control groups, which received care at the same breast clinic	Resource: Specialist appointments Concerns over appointment availability and wait times for diagnostic services following abnormal cancer screening results are common [19, 23]	Navigators helped patients schedule appointments. The navigation group obtained care leading to diagnostic resolution earlier than control patients	Outcomes: Time to diagnosis, patient anxiety, patient satisfaction Increased demand for specialist appointments has been linked to increased wait times for cancer screening procedures [23]. Scheduling difficulties can delay breast and colorectal cancer diagnosis [24, 25]
Randomized trial of patient navigation in an underserved population [11]	Patients individually randomized to patient navigation or control groups, which received care in the same health system	Resource: Specialist appointments Concerns over appointment availability and wait times for diagnostic services following abnormal cancer screening results are common [19, 23]	Navigators helped patients schedule appointments. The navigation group obtained care leading to diagnostic resolution earlier than control patients	Outcome: Time to resolution of an abnormal screening finding Increased demand for specialist appointments has been linked to increased wait times for cancer screening procedures [23]. Scheduling difficulties can delay breast and colorectal cancer diagnosis [24, 25]
Cluster randomized trial of EHR trigger-based reminder system to reduce diagnostic delays [26]	Primary care providers randomized to intervention or control groups. Their patients received care in the same health systems	Resource: Specialist appointments Concerns over appointment availability and wait times for diagnostic services following abnormal cancer screening results are common [19, 23]	Intervention involved reminder calls to primary care providers to schedule patients for diagnostic services. The navigation group obtained care leading to diagnostic resolution earlier than control patients	Outcome: Time to diagnostic evaluation Increased demand for specialist appointments has been linked to increased wait times for cancer screening procedures [23]. Scheduling difficulties can delay breast and colorectal cancer diagnosis [24, 25]
Randomized Trial of Induction Versus Expectant Management (ARRIVE) [27]	Patients individually randomized to induction or expectant management groups, which received care in same hospitals	Resource: Clinicians; labor and delivery rooms Delivery rooms unavailable for a few trial participants. Other studies of perinatal inpatient settings report lower than recommended nurse staffing levels [28] and that some nurses are routinely "overwhelmed" [29]	Elective induction group spent 45% longer in labor and delivery but had fewer cesarean deliveries and shorter postpartum stay Some inductions rescheduled when rooms were unavailable	Outcomes: "Composite of perinatal death or severe neonatal complications," cesarean delivery, duration of stay Other studies report adverse birth outcomes increase on days with more patients and fewer clinicians [30] and that higher workloads may lead to higher risks to patients [16] or lower rates of other procedures [31]
Randomized trial of machine learning-based sepsis prediction algorithm [32]	Patients individually randomized to usual sepsis detection system or usual system plus alerts from an additional sepsis prediction algorithm. Patients received care in the same ICUs	Resource: Clinician time and attention Concerns over clinician-to-patient ratios, staffing shortages, and care rationing in ICU settings are common [15, 33]. Clinicians' lack of time is cited as a barrier to providing patient care and following infection-control procedures in the ICU [34, 35]	Patients monitored by the added physiological alarm system received earlier tests and treatment for sepsis No information provided on number of added alarms for intervention patients. Other studies have reported that responding to alarms adds to clinician workload and can disrupt patient care [36, 37]	Outcomes: Duration of ICU stay, in-hospital mortality ICU nurse availability, in terms of workforce size and workload, has been linked to patient mortality [21]. Exposure to excessive numbers of alarms can overload clinicians and lead to missed alarms and potentially patient deaths [38]

Table 1 (continued)

Study	1. Resource shared across study arms	2. Scarce resource present in study setting	3. Study intervention affects resource use	4. Study outcome affected by resource availability
Randomized trial of artificial intelligence-enabled electrocardiographic alarm system (AI-ECG) [39]	In-hospital patients individually randomized to usual care or monitoring by AI-ECG alarm system. Patients received care from same clinicians and hospital units	Resource: Clinician time and attention; ICU beds Concerns over clinician-to-patient ratios and staffing shortages in hospital and ICU settings are common [15, 21, 33] ICU beds are not always available to meet the needs of all patients [4, 15, 20, 40]	Intervention effect "may be attributed to the increased attention of the enrolled physicians." AI-ECG system monitored 8001 patients and generated 709 alarms Patients assigned to AI-ECG alarm system were more likely to receive medical interventions or be admitted to the ICU	Outcomes: All-cause mortality, medical care utilization Clinician availability and workload has been linked to in-hospital mortality [18, 21]. Patients who are not admitted to an ICU due to lack of beds face higher mortality risks [20, 40]
Randomized trial of rapid whole-genome sequencing (NSIGHT1) [41]	Critically ill infants were individually randomized to receive standard genetic tests or standard tests plus a new genomic test. All tests were performed in same laboratory	Resource: Laboratory staff and other resources Turnaround of test results took multiple days. Others have raised concerns that staffing in clinical laboratories is sometimes inadequate [42]	The intervention consisted of performing a new genomic test in addition to standard tests, which increases laboratory workload NICU genetic sequencing and interpretation can take multiple hours or days of laboratory staff time [43, 44]	Outcome: Number of infants receiving genetic diagnosis within 28 days Turnaround time for standard and additional genomic tests directly affects rate of diagnosis within 28 days. Concerns have been raised that inadequate staff resources can affect turnaround time in clinical laboratories [42]. Added workload may lead to delayed test results for all patients
Randomized trial of text-messages to encourage uptake of COVID-19 vaccination [45]	Patients invited for COVID-19 vaccination were randomized to receive text-message reminders or no reminders. Patients scheduled vaccine appointments in the same health system	Resource: Vaccine doses and appointment slots The trial began January 2021. The number of invited patients was limited to ensure sufficient vaccine supply, but appointment availability was not reported. Other vaccination programs in the same county and university system reported limited appointment availability for eligible adults at this time [46, 47]	Patients who received text-message reminders scheduled vaccination appointments earlier and at a higher rate	Outcome: Number of participants scheduling a vaccination appointment within 6 days Availability of vaccination appointments at convenient times could have affected whether an individual scheduled an appointment

*Most trials provided limited or no information on resource availability, utilization, or effect on outcomes. As a result, I use evidence from other studies in similar settings to discuss the potential role that a scarce resource may have played during the trial

One study cited in this guidance randomly assigned ICU patients to monitoring by the usual sepsis detection system or the usual system plus a machine learning-based alert system [32]. The trial reported that in-hospital mortality decreased by over 50% for patients monitored by both the usual system and the additional alert system.

Alerts from the additional system may have directed extra clinician attention to patients in the intervention group. Clinician time and attention is often limited in ICU settings, which can lead to rationing of care and decreased adherence to infection-control protocols [4, 15, 34, 35]. Exposure to excessive numbers of alarms can overload clinicians and contribute to missed alarms and patient deaths [38, 56]. If patients in the intervention group received extra attention at the expense of clinicians' availability to care for control patients, negative spillover could have occurred.

Other interventions and study designs

Table 1 also contains studies of other interventions that may have influenced allocation of scarce resources between study arms. One randomized trial evaluated the effect of text-based reminders on COVID-19 vaccine uptake in early 2021 when vaccines were not widely available [45]. Another trial evaluated the impact of a new genomic screening test on time to diagnosis as compared to standard tests conducted at the same laboratory [41]; if technicians' workload increased due to the new test, this could have delayed test results for all patients. A third trial evaluated the effect of an automated reminder system on timely follow-up of abnormal test results, though all patients were in the same two health systems and their care may have drawn from a shared pool of diagnostic resources [26].

Pragmatic trials are not the only type of study that might be affected by negative spillover. For example, cohort studies have also been used to examine the effect of patient navigation programs [57]. In these studies, if navigation affected allocation of a scarce resource across cohorts, then spillover may have occurred. The conditions that lead to negative spillover could also be present in an explanatory randomized controlled trial, though this may be less likely due to efforts to ensure such trials are held under "ideal," rather than real-world, conditions [1, 17].

Implications

Negative spillover is not just a source of bias. It also has implications for patient safety. Researchers should not simply assume that a clinical trial will do no harm to patients assigned to usual care. If negative spillover occurs, these patients are no longer receiving "usual care" as it is commonly understood. Instead, they may be receiving substandard care, if clinical resources they need have been reallocated to others via random assignment. In addition, when a study intervention affects allocation of a scarce resource, the welfare of patients who share that resource but are not direct trial participants should also be considered [58]. Clinicians can also be harmed if an intervention increases workload and contributes to burnout or moral distress [18, 58, 59].

Spillover due to sharing of scarce resources will not uniformly lead to overestimation of treatment effect and harm to patients in the comparison arm; spillover could also occur in a positive direction. Patient navigation may result in lower utilization of diagnostic resources due to fewer missed appointments. Elective induction of labor might be rescheduled to reduce clinician workload when a labor and delivery unit is crowded. A physiological alarm system that results in timely treatment could reduce overall patient acuity in an ICU. Such mechanisms could free up medical resources and thus improve outcomes for patients assigned to a comparison group, leading investigators to underestimate the beneficial effects of an intervention.

Study results that may have been affected by negative spillover can still be a valid indication that an intervention affects care utilization. For example, patient navigation for underserved populations could still be used to decrease disparities in access to diagnostic services, even if navigation simply reallocates scarce appointment slots to these groups.

Detecting spillover

It is not possible to firmly conclude whether, or to what extent, negative or positive spillover occurred in the trials discussed here. In some cases, however, study data could be reanalyzed for signs of spillover. If spillover occurred, patients assigned to usual care during a trial could have experienced changes in outcomes when compared to a pre-trial baseline. If data are also available from non-participating health providers that can serve as an external control, then difference-in-differences study designs can be used to test for spillover effects on outcomes, resource utilization, or wait times for health services [60].

Researchers could also compare subgroup outcomes to test the hypothesis that a specific type of patient, study site, or time period included in a trial were more affected by cross-arm sharing of scarce resources. Several methods have been proposed to estimate spillover by exploiting variation in non-treated individuals' exposure to treated individuals [10, 61, 62]. Data on resource availability and utilization would be especially useful in such analyses. In some cases, however, obtaining definitive evidence on spillovers related to resource constraints
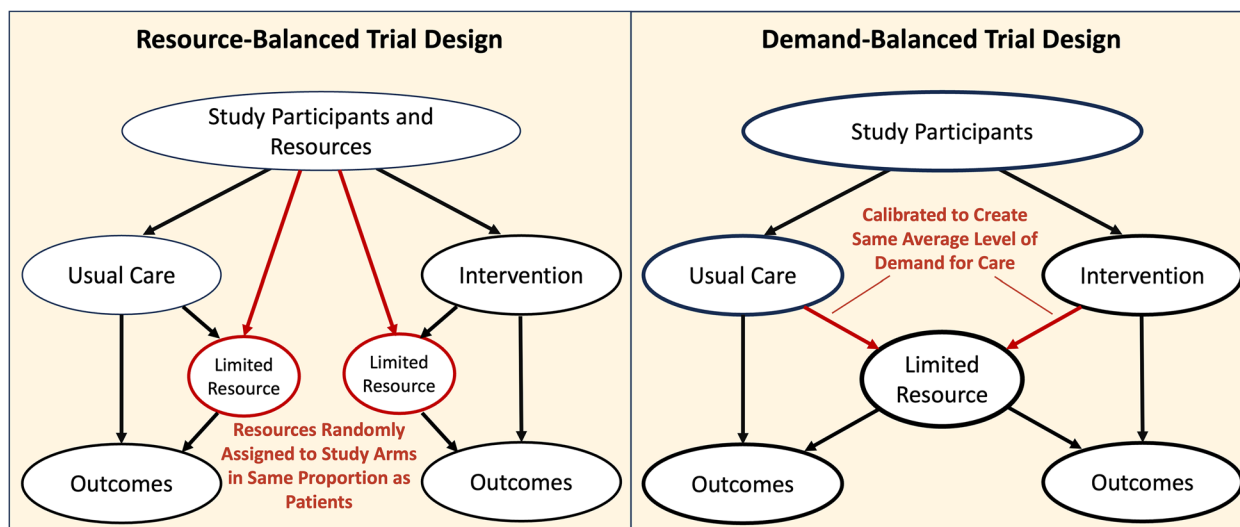
Fig. 2 Alternative trial designs that mitigate negative spillover. Two alternative trial designs might be considered to mitigate negative spillover when multi-site cluster-randomization is not feasible. In a resource-balanced trial, a resource (such as individual appointment slots) could be assigned to the control or treatment group in the same proportion as individual patients. In a demand-balanced trial design, the intervention (such as a physiological alarm system) could be calibrated to generate the same average level of demand for care per patient as is generated by usual care processes

might require conducting a two-level randomized trial like those used to identify other spillover effects [62, 63].

Designing spillover-proof trials

Investigators, research funders, institutional review boards, health system managers, patient organizations, and other stakeholders should consider whether the conditions that lead to negative spillover might be present in any proposed trial. Intangible resources, such as clinicians' time and attention, should be a particular focus given that they are easily overlooked. If a study intervention might affect allocation of a possibly scarce clinical resource, then the trial should be designed to avoid spillover.

Cluster-randomized trial designs can avoid negative spillover if the unit of randomization (such as a clinical practice or hospital) contains its own distinct pool of resources used to care for trial participants. Cluster randomization might not always be feasible, however, as it often requires higher numbers of participants and coordination across multiple study sites [2].

In some cases, a trial could be designed to balance resource availability or demand across intervention and control groups within a single site, as shown in Fig. 2. The proposed trial designs share characteristics with attention-control study designs used to limit bias in trials of behavioral interventions [64]. For example, if an intervention involves use of discrete resources such as vaccination appointments to achieve an outcome, a trial might randomize both patients and appointment slots,

in the same proportions, to the intervention and comparison arms. Similarly, an intervention that changes how patient needs are translated into a demand signal for care, such as a new physiological alarm system, might be calibrated to generate the same average number of alarms per patient as the existing system that it is compared to. Where practical, investigators might use such demand- or resource-balanced trial designs to mitigate negative spillover while also avoiding the larger costs associated with cluster randomized designs.

Finally, if trial design cannot be changed, researchers might still collect data on resource availability and utilization (ideally at baseline and during the trial) to help identify and control for potential spillover. Common data models intended to support comparative effectiveness research [65, 66] might also consider adding elements that measure availability of potentially scarce resources in study settings.

Improving methods guidance

Despite being a known problem in the field of development economics, the threat that negative spillover can pose to the validity and safety of randomized trials is not recognized in existing guidance on clinical trial methods. None of the following documents, for example, addresses the potential for spillover due to constraints on care delivery to occur in a randomized trial or that this could lead to bias or harm:

- EQUATOR Network reporting guidelines for randomized trials, including the CONSORT statement [67] and its extensions for non-pharmacological interventions [68] and AI-based interventions [69]
- Guidance on pragmatic trial design, including the PRECIS-2 framework [1] and the NIH *Living Textbook of Pragmatic Clinical Trials* [70]
- Guidance on assessing risk of bias in randomized trials, including frameworks from Cochrane [13], the Grading of Recommendations Assessment, Development and Evaluation (GRADE) working group [14], and the U.S. Preventive Services Task Force [71]
- Guidance on assessing harms to human subjects from the U.S. Department of Health and Human Services' Office for Human Research Protections [72, 73]

Future guidance on clinical trial methods should consider addressing spillover due to resource scarcity, just as current guidance discusses bias due to lack of blinding, differential attrition, or contamination [2, 3, 13, 14, 70].

Conclusion

Interest in pragmatic trials continues to grow [74]. Artificial intelligence-based interventions increasingly shape clinical decision-making, which has led to calls for more randomized trials evaluating their impact on patient outcomes [75, 76]. As new clinical practices and technologies that affect care delivery are considered for widespread adoption, well-designed trials are needed to provide valid evidence on their risks and benefits. Understanding all sources of bias that could affect these trials, including negative spillover, is a critical part of this effort.

Abbreviations
AI	Artificial intelligence
CONSORT	Consolidated Standards of Reporting Trials
EQUATOR	Enhancing the Quality and Transparency of Health Research
GRADE	Grading of Recommendations Assessment, Development and Evaluation
ICU	Intensive care unit
NIH	National Institutes of Health
PRECIS	Pragmatic Exploratory Continuum Indicator Summary

References
1. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. BMJ. 2015;350:h2147.
2. Mercer SL, DeVinney BJ, Fine LJ, Green LW, Dougherty D. Study designs for effectiveness and translation research: identifying trade-offs. Am J Prev Med. 2007;33(2):139-54.e2.
3. Sanson-Fisher RW, Bonevski B, Green LW, D'Este C. Limitations of the randomized controlled trial in evaluating population-based health interventions. Am J Prev Med. 2007;33(2):155–61.
4. Scheunemann LP, White DB. The ethics and reality of rationing in medicine. Chest. 2011;140(6):1625–32.
5. Scott PA, Harvey C, Felzmann H, Suhonen R, Habermann M, Halvorsen K, et al. Resource allocation and rationing in nursing care: a discussion paper. Nurs Ethics. 2018;26(5):1528–39.
6. Emanuel EJ, Persad G, Upshur R, Thome B, Parker M, Glickman A, et al. Fair allocation of scarce medical resources in the time of COVID-19. N Engl J Med. 2020;382(21):2049–55.
7. Giannella E, Homonoff T, Rino G, Somerville J. Administrative burden and procedural denials: experimental evidence from SNAP. National Bureau of Economic Research Working Paper Series. 2023;No. 31239.
8. Francetic I, Meacock R, Elliott J, Kristensen SR, Britteon P, Lugo-Palacios DG, et al. Framework for identification and measurement of spillover effects in policy implementation: intended non-intended targeted non-targeted spillovers (INTENTS). Implement Sci Commun. 2022;3(1):30.
9. Mody A, Sikazwe I, Czaicki NL, Wa Mwanza M, Savory T, Sikombe K, et al. Estimating the real-world effects of expanding antiretroviral treatment eligibility: evidence from a regression discontinuity analysis in Zambia. PLoS Med. 2018;15(6): e1002574.
10. Duflo E, Glennerster R, Kremer M. Chapter 61 Using randomization in development economics research: a toolkit. In: Schultz TP, Strauss JA, editors. Handbook of Development Economics. 4: Elsevier; 2007. p. 3895–962.
11. Raich PC, Whitley EM, Thorland W, Valverde P, Fairclough D. Patient navigation improves cancer diagnostic resolution: an individually randomized clinical trial in an underserved population. Cancer Epidemiol Biomarkers Prev. 2012;21(10):1629–38.
12. Kazarian ES, Carreira FS, Toribara NW, Denberg TD. Colonoscopy completion in a large safety net health care system. Clin Gastroenterol Hepatol. 2008;6(4):438–42.
13. Higgins JPT, Savović J, Page MJ, Elbers RG, Sterne JAC. Chapter 8: Assessing risk of bias in a randomized trial [last updated October 2019]. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. Cochrane Handbook for Systematic Reviews of Interventionsversion 6.5. Cochrane; 2024. Available from: https://training.cochrane.org/handbook.
14. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). J Clin Epidemiol. 2011;64(4):407–15.
15. Truog RD, Brock DW, Cook DJ, Danis M, Luce JM, Rubenfeld GD, et al. Rationing in the intensive care unit. Crit Care Med. 2006;34(4):958–63.
16. Simpson KR, Lyndon A. Consequences of delayed, unfinished, or missed nursing care during labor and birth. J Perinat Neonatal Nurs. 2017;31(1):32–40.

17. Ford I, Norrie J. Pragmatic trials. N Engl J Med. 2016;375(5):454–63.
18. Aiken LH, Clarke SP, Sloane DM, Sochalski J, Silber JH. Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. JAMA. 2002;288(16):1987–93.
19. Institute of Medicine. *Transforming health care scheduling and access: Getting to now*. Washington, DC: The National Academies Press; 2015.
20. Cardoso LT, Grion CM, Matsuo T, Anami EH, Kauss IA, Seko L, et al. Impact of delayed admission to intensive care units on mortality of critically ill patients: a cohort study. Crit Care. 2011;15(1):R28.
21. Needleman J, Buerhaus P, Pankratz VS, Leibson CL, Stevens SR, Harris M. Nurse staffing and inpatient hospital mortality. N Engl J Med. 2011;364(11):1037–45.
22. Ferrante JM, Chen PH, Kim S. The effect of patient navigation on time to diagnosis, anxiety, and satisfaction in urban minority women with abnormal mammograms: a randomized controlled trial. J Urban Health. 2008;85(1):114–24.
23. Hubers J, Sonnenberg A, Gopal D, Weiss J, Holobyn T, Soni A. Trends in wait time for colorectal cancer screening and diagnosis 2013–2016. Clin Transl Gastroenterol. 2020;11(1): e00113.
24. Wujcik D, Fair AM. Barriers to diagnostic resolution after abnormal mammography: a review of the literature. Cancer Nurs. 2008;31(5):E16-30.
25. Singh H, Khan R, Giardina TD, Paul LW, Daci K, Gould M, et al. Postreferral colonoscopy delays in diagnosis of colorectal cancer: a mixed-methods analysis. Qual Manag Health Care. 2012;21(4):252–61.
26. Murphy DR, Wu L, Thomas EJ, Forjuoh SN, Meyer AND, Singh H. Electronic trigger-based intervention to reduce delays in diagnostic evaluation for cancer: a cluster randomized controlled trial. J Clin Oncol. 2015;33(31):3560–7.
27. Grobman WA, Rice MM, Reddy UM, Tita ATN, Silver RM, Mallett G, et al. Labor induction versus expectant management in low-risk nulliparous women. N Engl J Med. 2018;379(6):513–23.
28. Simpson KR, Spetz J, Gay CL, Fletcher J, Landstrom GL, Lyndon A. Hospital characteristics associated with nurse staffing during labor and birth: inequities for the most vulnerable maternity patients. Nurs Outlook. 2023;71(3): 101960.
29. Roth C, Brewer MA, Bay RC, Gosselin KP. Nurses' experiences of "being swamped" in the clinical setting and association with adherence to AWHONN nurse staffing guidelines. MCN Am J Matern Child Nurs. 2020;45(5):271–9.
30. Snowden JM, Kozhimannil KB, Muoto I, Caughey AB, McConnell KJ. A 'busy day' effect on perinatal complications of delivery on weekends: a retrospective cohort study. BMJ Qual Saf. 2017;26(1): e1.
31. Maibom J, Sievertsen HH, Simonsen M, Wüst M. Maternity ward crowding, procedure use, and child health. J Health Econ. 2021;75: 102399.
32. Shimabukuro DW, Barton CW, Feldman MD, Mataraso SJ, Das R. Effect of a machine learning-based severe sepsis prediction algorithm on patient survival and hospital length of stay: a randomised clinical trial. BMJ Open Respir Res. 2017;4(1): e000234.
33. Pastores SM, Kvetan V, Coopersmith CM, Farmer JC, Sessler C, Christman JW, et al. Workforce, workload, and burnout among intensivists and advanced practice providers: a narrative review. Crit Care Med. 2019;47(4):550–7.
34. Stahmeyer JT, Lutze B, von Lengerke T, Chaberny IF, Krauth C. Hand hygiene in intensive care units: a matter of time? J Hosp Infect. 2017;95(4):338–43.
35. Fridkin SK, Pear SM, Williamson TH, Galgiani JN, Jarvis WR. The role of understaffing in central venous catheter-associated bloodstream infections. Infect Control Hosp Epidemiol. 1996;17(3):150–8.
36. Dursun Ergezen F, Kol E. Nurses' responses to monitor alarms in an intensive care unit: an observational study. Intensive Crit Care Nurs. 2020;59: 102845.
37. Ruppel H, Funk M, Clark JT, Gieras I, David Y, Bauld TJ, et al. Attitudes and practices related to clinical alarms: a follow-up survey. Am J Crit Care. 2018;27(2):114–23.
38. Sendelbach S, Funk M. Alarm fatigue: a patient safety concern. AACN Adv Crit Care. 2013;24(4):378–86.
39. Lin C-S, Liu W-T, Tsai D-J, Lou Y-S, Chang C-H, Lee C-C, et al. AI-enabled electrocardiography alert intervention and all-cause mortality: a pragmatic randomized clinical trial. Nat Med. 2024;30(5):1461–70.
40. Iapichino G, Corbella D, Minelli C, Mills GH, Artigas A, Edbooke DL, et al. Reasons for refusal of admission to intensive care and impact on mortality. Intensive Care Med. 2010;36(10):1772–9.
41. Petrikin JE, Cakici JA, Clark MM, Willig LK, Sweeney NM, Farrow EG, et al. The NSIGHT1-randomized controlled trial: rapid whole-genome sequencing for accelerated etiologic diagnosis in critically ill infants. NPJ Genomic Med. 2018;3(1):6.
42. Valenstein PN, Souers R, Wilkinson DS. Staffing benchmarks for clinical laboratories: a College of American Pathologists Q-probes study of staffing at 151 institutions. Arch Pathol Lab Med. 2005;129(4):467–73.
43. Kingsmore SF, Cole FS. The role of genome sequencing in neonatal intensive care units. Annu Rev Genomics Hum Genet. 2022;23:427–48.
44. Saunders CJ, Miller NA, Soden SE, Dinwiddie DL, Noll A, Alnadi NA, et al. Rapid whole-genome sequencing for genetic disease diagnosis in neonatal intensive care units. Sci Transl Med. 2012;4(154):154ra35-ra35.
45. Dai H, Saccardo S, Han MA, Roh L, Raja N, Vangala S, et al. Behavioural nudges increase COVID-19 vaccinations. Nature. 2021;597(7876):404–9.
46. University of California San Francisco. Novel coronavirus (COVID-19) resources: latest updates 2021 [Internet]. [updated August 4, 2021]. Available from: https://coronavirus.ucsf.edu/latest-updates.
47. City of Hidden Hills California. COVID-19 vaccination information 2021 [Internet]. [updated January 22, 2021]. Available from: https://hiddenhillscity.org/2021/01/22/covid-19-vaccination-information/.
48. Dougherty MK, Brenner AT, Crockett SD, Gupta S, Wheeler SB, Coker-Schwimmer M, et al. Evaluation of interventions intended to increase colorectal cancer screening rates in the United States: a systematic review and meta-analysis. JAMA Intern Med. 2018;178(12):1645–58.
49. Robinson-White S, Conroy B, Slavish KH, Rosenzweig M. Patient navigation in breast cancer: a systematic review. Cancer Nurs. 2010;33(2):127–40.
50. Vijayaraghavan GR, Guembou IM, Vedantham S. The current state of timeliness in the breast cancer diagnosis journey: abnormal screening to biopsy. Seminars in Ultrasound, CT and MRI. 2023;44(1):56–61.
51. American College of Obstetricians and Gynecologists' Committee on Obstetric Practice. Clinical guidance for integration of the findings of the ARRIVE trial: labor induction versus expectant management in low-risk nulliparous women 2018 [Internet]. Available from: https://www.acog.org/clinical/clinical-guidance/practice-advisory/articles/2018/08/clinical-guidance-for-integration-of-the-findings-of-the-arrive-trial.
52. Grobman WA, Rice MM, Reddy UM, Tita ATN, Silver RM, Mallett G, et al. Supplementary Appendix: Labor induction versus expectant management in low-risk nulliparous women. N Engl J Med. 2018;379(6):513–23.
53. Mann S, James KJ. Elective Induction of Labor May Have Negative Effects at the Hospital Level. J Obstet Gynecol Neonatal Nurs. Published online October 10, 2024. https://doi.org/10.1016/j.jogn.2024.09.003.
54. Evans L, Rhodes A, Alhazzani W, Antonelli M, Coopersmith CM, French C, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. Intensive Care Med. 2021;47(11):1181–247.
55. U.S. Centers for Disease Control and Prevention. Hospital sepsis program core elements. Atlanta: U.S. Department of Health and Human Services, CDC; 2023. Available from: https://www.cdc.gov/sepsis/hcp/core-elements/index.html.
56. The Joint Commission. Medical device alarm safety in hospitals. Sentinel Event Alert. 2013;50:1–3.
57. Serrell EC, Hansen M, Mills G, Perry A, Robbins T, Feinberg M, et al. Prostate cancer navigation: initial experience and association with time to care. World J Urol. 2019;37(6):1095–101.
58. U.S. Department of Health and Human Services; Secretary's Advisory Committee on Human Research Protections. The protection of non-subjects from research harm [Internet]. 2022. Available from: https://www.hhs.gov/ohrp/sachrp-committee/recommendations/tab-c-the-protection-of-non-subjects-from-research-harm.html.
59. National Academy of Medicine; Committee on Systems Approaches to Improve Patient Care by Supporting Clinician Well-Being. Taking action against clinician burnout: a systems approach to professional well-being. Washington, DC: The National Academies Press; 2019.
60. Callaway B, Sant'Anna PHC. Difference-in-differences with multiple time periods. J Econometrics. 2021;225(2):200–30.
61. Vazquez-Bare G. Identification and estimation of spillover effects in randomized experiments. Journal of Econometrics. 2023;237(1): 105237.

62. Benjamin-Chung J, Arnold BF, Berger D, Luby SP, Miguel E, Colford JM Jr, et al. Spillover effects in epidemiology: parameters, study designs and methodological considerations. Int J Epidemiol. 2018;47(1):332–47.

63. Egger D, Haushofer J, Miguel E, Niehaus P, Walker M. General equilibrium effects of cash transfers: experimental evidence from Kenya. Econometrica. 2022;90(6):2603–43.

64. Miller LD, Laye-Gindhu A, Liu Y, March JS, Thordarson DS, Garland EJ. Evaluation of a preventive intervention for child anxiety in two randomized attention-control school trials. Behav Res Ther. 2011;49(5):315–23.

65. Beaber EF, Kim JJ, Schapira MM, Tosteson AN, Zauber AG, Geiger AM, et al. Unifying screening processes within the PROSPR consortium: a conceptual model for breast, cervical, and colorectal cancer screening. J Natl Cancer Inst. 2015;107(6):djv120.

66. Fleurence RL, Curtis LH, Califf RM, Platt R, Selby JV, Brown JS. Launching PCORnet, a national patient-centered clinical research network. J Am Med Inform Assoc. 2014;21(4):578–82.

67. Schulz KF, Altman DG, Moher D, the CG. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Med. 2010;8(1):18.

68. Boutron I, Altman DG, Moher D, Schulz KF, Ravaud P. CONSORT statement for randomized trials of nonpharmacologic treatments: a 2017 update and a CONSORT extension for nonpharmacologic trial abstracts. Ann Intern Med. 2017;167(1):40–7.

69. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, Chan A-W, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med. 2020;26(9):1364–74.

70. Heagerty P. Experimental designs and randomization schemes. Bethesda: NIH Pragmatic Trials Collaboratory; 2023. Available from: https://rethinkingclinicaltrials.org/chapters/design/experimental-designs-and-randomization-schemes/experimental-designs-introduction/.

71. U.S. Preventive Services Task Force. Appendix VI. Criteria for assessing internal validity of individual studies 2017 [Internet]. Available from: https://www.uspreventiveservicestaskforce.org/uspstf/about-uspstf/methods-and-processes/procedure-manual/procedure-manual-appendix-vi-criteria-assessing-internal-validity-individual-studies.

72. U.S. Department of Health and Human Services; Office on Human Research Protections. Draft guidance on disclosing reasonably foreseeable risks in research evaluating standards of care [Internet]. 2014. Available from: https://www.hhs.gov/ohrp/regulations-and-policy/requests-for-comments/draft-guidance-disclosing-risk-in-standards-of-care/index.html.

73. U.S. Department of Health and Human Services; Office on Human Research Protections. Reviewing and reporting unanticipated problems involving risks to subjects or others and adverse events: OHRP guidance [Internet]. 2007. Available from: https://www.hhs.gov/ohrp/regulations-and-policy/guidance/reviewing-unanticipated-problems/index.html#Q2.

74. Palakshappa JA, Gibbs KW, Lannan MT, Cranford AR, Taylor SP. Systematic review of the "pragmatism" of pragmatic critical care trials. Crit Care Explor. 2022;4(7): e0738.

75. Wynants L, Smits LJM, Van Calster B. Demystifying AI in healthcare. BMJ. 2020;370: m3505.

76. Topol EJ. Welcoming new guidelines for AI clinical research. Nat Med. 2020;26(9):1318–20.

Publisher's Note