NIH PRAGMATIC TRIALS COLLABORATORY
Rethinking Clinical Trials®

Critical Questions for Pragmatic Clinical Trialists: Insights From the NIH Pragmatic Trials Collaboratory's First Decade

WORKSHOP SUMMARY

June 15-16, 2022

Prepared by: NIH Pragmatic Trials Collaboratory Coordinating Center
Version: December 15, 2022

Executive Summary

In June 2022, the NIH Pragmatic Trials Collaboratory held a workshop over 2 half-days that explored questions critical to designing and implementing pragmatic clinical trials (PCTs) conducted within healthcare systems. PCTs differ from traditional, more-explanatory clinical trials, as they test interventions or practices delivered in as part of routine care. The workshop included a series of moderated discussions on the design and conduct of successful PCTs. Topics included changes and current best practices regarding the use of electronic health records for pragmatic research, considerations for PCT study design, how to optimize a 1-year planning phase for a PCT, and tips for implementing interventions in complex health systems. Panelists shared lessons learned, solutions to challenges, and directions for the future.

Key lessons from each of the panels are described below.

PCTs: The Role of Nursing Science

Nurses, who are a trusted and large segment of the healthcare workforce are ideal partners for the conduct of PCTs, as nursing science spans multiple disciplines and patient populations. The National Institute of Nursing Research (NINR) supports research across clinical, community, and policy settings to optimize health for individuals, families, communities, and populations. This helps advance health equity by removing structural barriers from research, cultivates diversity in perspectives and ideas, and fosters inclusion in research.

Can We Squeeze More Value Out of Electronic Health Records?

Healthcare systems are dynamic and ever changing, and as such, trial investigators must pay attention to the details and respond to challenges with agility and flexibility. Each trial encountered unique problems, and panelists described several desirable EHR enhancements for the future, including:

- Automated intervention delivery without an expensive build
- Standardization of data across multiple health systems
- A way to incorporate standardized and customized patient-reported outcome measures for research and clinical care, and a way to administer those at scheduled visits and at regular follow-up intervals
- Parameterizable questionnaires and automated tools for responding to patient information
- Parameterized tools for cohort identification
- A good way to assess and record negative medication experiences

Should We Be Considering a Different Study Design?

Because of the nature of PCTs, there are various internal and internal forces that can impact a study, underscoring the need for monitoring during all phases of the trial. One of the major lessons learned is that differential changes over time and site variability can create serious challenges, leading the panelists to suggest that individually randomized trials are favored over cluster designs.

What Is the Most Effective Use of the 1-Year Planning Phase?

Although each project used their planning phases differently to tackle what they thought would be the toughest pieces, all teams recommended figuring out as many details as possible during this time, including merging and aggregating data across health systems, evaluating the population for diversity, and preparing for nimble data collection during the study. Importantly, the trialists recommend starting with the sites that are anticipated to be the most challenging.

How Do We Implement Interventions in Complex Healthcare Systems?

Implementation science provides a framework that allows researchers to pay attention to the details while balancing the tension between being flexible enough to conduct the trial as part of care in a complex environment while maintaining the trial's core components. One reason complexity may be underestimated is that it is hard to measure. The NIH Pragmatic Trials Collaboratory is developing a tool to share with trialists that can helps to identify and measure the complex components of PCT interventions in the planning stage.

Conclusion

Ten years ago, there was an assumption that PCTs would be easy. In the intervening years, the NIH Pragmatic Trials Collaboratory has learned that along with the flexibility and opportunities that a PCT offers, there are also many challenges. Now, many other funders across and outside of NIH are leveraging the experience of the NIH Pragmatic Trials Collaboratory and conducting trials in healthcare systems and communities. Ultimately, the goal of all this work is to get better at gathering evidence that helps improve care for patients. There is a bright future, we have learned a lot, and there is a lot more to learn.

PCTs: The Role of Nursing Science

The workshop opened with a keynote presentation by Shannon N. Zenk, PhD, MPH, RN, Director of the National Institute of Nursing Research (NINR). NINR is focused on health solutions for people in the context of their lives and living conditions. NINR's mission is leading nursing science to solve pressing health challenges and inform practice and policy, optimizing health and advancing health equity into the future.

Key Takeaways

- Nursing science can contribute to the conduct of PCTs embedded in health systems, as nursing research aims to discover solutions across clinical, community, and policy settings to optimize health for individuals, families, communities, and populations.
- Nurses are the largest segment of the healthcare workforce and are the most trusted. Researchers should consider including them in research to build trust and engage community partners.

According to the NINR 2022-26 Strategic Plan, all NINR-supported research tackles today's pressing health challenges and stimulates discoveries to prepare for, prevent, or address tomorrow's challenges. The research discovers solutions across clinical, community, and policy settings to optimize health for individuals, families, communities, and populations. It also advances equity by removing structural barriers from research, cultivating diversity in perspectives and ideas, and fostering inclusion and accessibility in designing, conducting and participating in research. NINR-supported research is innovative, develops or applies the most rigorous methods, and has the potential for the greatest impact on health. NINR approaches research through 5 lenses: health equity, social determinants of health, population and community health, prevention and health promotion, and systems and models of care.

Nursing science can make significant contributions to PCTs. Nurses make up the largest segment of the healthcare workforce and are already in every healthcare setting. Nurses are in a wide range of critical community settings and know patients in the context of their whole lives. They are also familiar with the electronic health record. Nurses are decision-makers and serve as a vital link between patients, policymakers, and other providers. They bring to research a knowledge of the social determinants of the health conditions in which people are born, live, work, and age. PCTs in healthcare settings are more likely be successful if they address social needs and social determinants of health.

It is important to include community members from the outset of a research study so that researchers and community members can co-create interventions that respond to issues that are important to the community and will be sustainable. Because nurses make up the largest part of the healthcare workforce and are the most trusted, researchers could consider including nurses in research to build trust and engage community partners. The best way to engage nurses in research is from the start. Nurses know what interventions might work and where challenges may exist.

Can We Squeeze More Value Out of Electronic Health Records?

This panel, moderated by Keith Marsolo, PhD, discussed the research benefits of electronic health records (EHRs) that have emerged over the last 10 years, including better support of research processes such as cohort identification, recruitment, intervention delivery, and outcomes monitoring. Four trials were highlighted, each presenting specific uses of EHRs, limitations discovered, and lessons learned.

Key Takeaways

- Longstanding challenges of using EHRs remain, including modifying the EHR to be fit-for-purpose for every trial; collecting new data elements at scale; and aggregating data across sites or EHR platforms.
- Some clinics in one trial were virtual and had a changing list of providers, requiring consistent updating and monitoring during the trial.
- The transition from ICD-9 to ICD-10 coding led to unpleasant surprises for some trials.
- Among the EHR enhancements desired by the investigators were automated intervention delivery without an expensive build; a way to capture data across multiple health systems (existing tools are incomplete); a way to incorporate standardized and customized PRO measures for research and clinical care; and parameterizable questionnaires and automated tools for responding to patient information.

Panelists and Case Studies

- Improving Chronic Disease Management with Pieces (ICD-Pieces™)
  Miguel Vazquez, MD
- Lumbar Imaging with Reporting of Epidemiology (LIRE)
  Jeffrey (Jerry) Jarvik, MD, MPH
- Nonpharmacologic Options in Postoperative Hospital-based and Rehabilitation Pain Management (NOHARM)
  Andrea Cheville, MD
- Suicide Prevention Outreach Trial (SPOT)
  Greg Simon, MD, MPH

ICD-Pieces

ICD-Pieces is a cluster-randomized PCT that aimed to identify patients diagnosed with multiple comorbidities (chronic kidney disease, diabetes, and hypertension) and deliver evidence-based interventions. There were 4 participating health systems with 3 different EHRs (Epic, Allscripts, and CPRS) . The trial used EHRs to identify candidates, deliver interventions, and capture outcomes. Primary outcomes were captured from claims data.

Eligible patients were identified using the EHR by objective laboratory and clinical data. Each week, a list of potential candidates was delivered to participating sites to review and enroll the patients. An unfortunate surprise was that 2 health systems would not allow their data to be stored in the cloud, meaning the algorithm had to be recreated inside the firewall, leading to delays. Another was that creatinine-albumin ratio measurements had to be separated by at least 3 months. Therefore, the trial implemented a manual review to prevent patients being missed.

Intervention delivery via the EHR worked well. Although structured Smart Sets were available, most providers preferred to adjust care plans for the patients themselves. The trial used a softer touch whereby guides, protocols, and Smart Sets were available but with no requirement for providers to use a structured plan. Interventions were affected by changes to care caused by practitioners retiring, changes to management of chronic kidney disease, new hypertension guidelines, and new medications for diabetes.

Some metrics for monitoring intervention delivery were not always available if patients had not received a certain test. It was also not possible to automate capturing metrics when the EHR-based tools were used. Manual reviews were needed early in the study.

Capturing outcomes worked well for 3 sites. For the fourth, the intermediary supplying the data changed, which delayed allowing data sharing for the purposes of research.

LIRE

LIRE is a stepped-wedge, cluster-randomized PCT that tested the insertion of benchmark prevalence data into routine radiology reports for people who had received lumbar spine imaging. The trial took place at 4 integrated health systems using Epic. It used EHRs and other information systems to identify clinics and primary care clinicians, insert intervention text into the radiology reports, tailor intervention text by imaging modality and patient age, verify intervention implementation, and assess outcomes.

A problem arose when recruiting patients because the systems were not as static as assumed. Some clinics were virtual and could have a changing list of providers. Consistent updating and monitoring were therefore required throughout the trial.

The EHR worked well for delivering the clinical decision support intervention, which was automated and customized to imaging modality and patient age. There was, however, a lot of implementation heterogeneity among sites. One preferred to insert interventions as pop-up alerts that clinicians received, which were not permanently attached to the report. Some inserted interventions using the radiology reporting system, which allowed the dictating radiologist to remove intervention text if desired.

To monitor interventions, the research team used regular data pulls to ensure text was inserted appropriately into the radiology reports. This led to the discovery that an upgrade to the dictation software at one of the sites had broken the link that inserted the intervention text, causing unintentional crossover. As a result, the frequency of monitoring was increased.

A lot of effort was needed to assess the final outcomes due to the ICD-9 to ICD-10 transition, which occurred during the trial. Additionally, coding practices were not identical across sites, and one had their own proprietary set of codes.

NOHARM

NOHARM aims to provide validated nonpharmacologic pain care options, while attenuating opioid use, in patients undergoing qualifying surgeries. It is being conducted at 6 sites using a singular centralized Epic EHR. The EHR allowed delivery of 3 intervention components. First, a conversation guide is automatically sent to the patient portal to describe the risk of opioids, norms for pain, and efficacy of modalities for nonpharmacologic pain care. The guide educates patients on 13 modalities to allow them to make a choice concordant with their preferences. Second, a collection of clinical decision support tools prompts clinicians to support the use of these modalities. Third, a suite of self-management tools supports use of the modalities. The trial used the EHR to identify participants, support transitions as clusters went live, deliver tailored messaging, prompt care teams, and collect outcomes.

The EHR allowed timely identification of eligible patients by the entry of a surgical order. Triangulation of different data elements identified patients for the control or intervention. A difficulty was that the registry was updated after hours, meaning the EHR's knowledge transfer occurred at midnight. There was also a heterogeneity of workflows, such as use of a paper log or placing surgical orders prior to informing the patient.

Automated delivery of the conversation guide worked well. The Epic questionnaire functionality allowed HTML embedding. However, data from interactions with the HTML were not collected. Another issue was that title format was locked, which limited creating tailored experiences. However, tailored intervention delivery worked well. The guide assessed patient confidence in ability to use the intervention and tailored messaging accordingly. Modality-specific information was also tailored to patient location and local resources. An unpleasant surprise was the shift from overnight to same-day surgeries and a lot of cancellations and rescheduling, which required algorithms to identify these events.

Patient interaction with the guide was monitored to ascertain if they received education for selected modalities. Epic has a sophisticated capture functionality. If a patient had pain, alerts were used to indicate their preferences and to determine how a nurse responded. Alerts were also configured to the intervention stage. Capturing how care providers used NOHARM interfaces is difficult, as the Epic dataset that captures user engagement is limited.

A challenge when assessing outcomes was accessing rural populations. Assessing portal status could indicate if a patient was using the portal. If not, ZIP codes were used to determine if they were in a broadband inaccessible area, and if so, printed patient-reported outcome measures were automatically mailed. This increased response rates among rural and older patients. This logic, however, was vulnerable to system upgrades. Clinic-based assignment of patient-reported outcome measures was also broadened to any healthcare encounter so that portal non-users seeking care were likely to be administered the PROs.

SPOT

SPOT is a patient-randomized PCT that evaluated 2 outreach interventions for prevention of self-harm or suicide attempts among outpatients reporting suicidal ideation. It was conducted at 4 integrated health systems that used Epic. The trial used the EHR to identify people who made an outpatient visit, completed a PHQ-9 depression questionnaire, and reported self-harm or suicidal ideation. The interventions depended on messaging capabilities through the portal. Data from the EHR, external insurance claims, and state mortality records were used to identify the primary outcomes.

Potential participating sites were selected based on consistent and frequent use of PHQ-9 measures. Each site had a weekly standard program that generated a waterfall diagram to show the number of people who made a visit, filled out a PHQ-9 questionnaire, and scored high.

Two different outreach interventions were delivered. A care management intervention systematically interacted with the patient and their care team, which worked well with the portal. The other intervention involved both text-based communication and interactive material. There was discontinuity between these two ways of interacting.

When monitoring the interventions, a key question was the value of outreach to people who were not seeking care for suicide prevention. People assigned to be offered the program received it via portal messaging. Failure to respond could have been due to helping or irritating the patient. Hints for which it was could include if they read, deleted, or lingered on the message. Because the SPOT trial found that one intervention may have been harmful (patients offered dialectical behavior therapy skills training had significantly greater risk of self-harm compared to usual care), r eaching out to participants to answer questions related to why they did not respond was not possible.

Assessing outcomes using a questionnaire would have been problematic. The primary outcome included nonfatal and fatal self-harm events. This was a low frequency event (4%-5%) , and therefore even high response rates would not be enough. Also, nonresponse cannot be considered random, as willingness to participate may be related to suicidal behavior. Therefore, the SPOT team used data records for outcome assessment. They ascertained nonfatal events from encounter diagnosis in the EHR and insurance claims. An unpleasant surprise was the transition from ICD-9 to ICD-10 coding before the trial began, which led to changes in the coding of injuries and poisonings. Validation was required to identify those that may have been missed, including injuries with undetermined intent and accidents in people at risk of self-harm. There is no good way to overcome the issue that using health system records only ascertains events for which people sought care. A secondary analysis limited to severe events indicated a similar pattern of results.

Discussion Questions

Are there strategies for working with health systems to collect new data elements, such as social determinants of health, sexual orientation, and gender identity?

Health systems have the right infrastructure for collecting new data elements but are proceeding cautiously. A clear consensus about what questions should be asked, of whom, and when is needed. Importantly, we should not record anything as free text.

There are several important challenges. At Mayo, sexual orientation and gender identity questions were made permanently available to patients, should they wish to change their answers. Patients therefore assumed they had to keep reporting. Similarly, in the NOHARM conversation guide, patients thought they were required to continually reaffirm their modality selections. Epic does not have a subtle ongoing way of including these sensitive open-ended questions. Additionally, not all social determinants of health constructs are actionable by a health system.

What is on your wish list for the EHR for the next 10 years?

Panelists described several desirable EHR enhancements, including:

- Automated intervention delivery without an expensive build
- Standardization of data elements across multiple health systems
- A way to incorporate standardized and customized patient-reported outcome measures for research and clinical care, and a way to administer those at scheduled visits and at regular follow-up intervals
- Parameterizable questionnaires and automated tools for responding to patient information
- Parameterized tools for cohort identification
- A good way to assess and record negative medication experiences as the current process discourages recording useful information; for example, if you report the reason for stopping sertraline was an adverse effect, when you subsequently prescribe escitalopram, each refill prescription requires clicking through 3 warnings for class intolerance

The EHR can pull in information from other health systems, but you cannot leverage those data for research. How can research piggyback on exchange mechanisms in place?

The boundary that has been established is that in the context of a research project, once you cross the boundary to become a care provider, you should deliver the best care possible using those records. In the SPOT trial, using information from the other health systems to identify research participants or ascertain outcomes was out of bounds, but providing appropriate care management for suicide prevention was possible.

Since starting your trials, have EHR systems evolved to make data collection easier? Would it be easier to launch your trial now?

Panelists described several improvements in EHR technical capabilities. For example, intervention delivery in the SPOT trial was clunky, but things have improved in terms of what can be embedded in messages. LIRE experienced that systems are evolving slowly. Epic has received pressure to improve capabilities to present patients with their own patient-reported outcome measure scores, to track them, and to link interventions. Another emerging capability is a functionality called Compass Rose, which is an application based on interactive care plans. Finally, the upgrade cycle for EHRs is occurs on almost a yearly basis, which brings everyone closer to a common platform.

Should We Be Considering a Different Study Design?

This panel, moderated by Patrick Heagerty, PhD, focused on the unique design considerations for PCTs: They are embedded in a healthcare context; they are pragmatic and therefore have multiple potential sources of variation; and they are longitudinal, so many changes can happen regarding time. Four trials were highlighted for the process of designing a clear study.

Key Takeaways

- Essential to study design is clarity around the research question, which leads to design fidelity, and considering peer input.
- Various intercurrent forces—both intrinsic and extrinsic—can affect any study, including changes regarding time and setting.
- Differential heterogeneity over time is a challenge for both stepped-wedge and parallel group designs. Site variability considerations favor individual-level randomization when possible.
- Cluster randomization combined with informed consent can be a big challenge. Advice was not to do a randomized trial with post-randomization consent.
- While it is important to focus on being pragmatic when feasible, overly pragmatic designs can potentially harm internal validity.
- Data and safety monitoring boards should be used for all PCTs and should include regular quality assurance checks as part of trial monitoring.

Panelists and Case Studies

- A Policy-Relevant U.S. Trauma Care System Pragmatic Trial for PTSD and Comorbidity (Trauma Survivors Outcomes and Support [TSOS])
  Doug Zatzick, MD
- Pragmatic Trial of Higher vs. Lower Serum Phosphate Targets in Patients Undergoing Hemodialysis (HiLo)
  Myles Wolf, MD, MMSc

- Pragmatic Trial of Video Education in Nursing Homes (PROVEN)
  Vince Mor, PhD
- Suicide Prevention Outreach Trial (SPOT)
  Susan Shortreed, PhD

TSOS

The research question TSOS sought to answer was, will injured patients randomized to a collaborative care intervention demonstrate significant reductions in posttraumatic stress disorder (PTSD) symptoms when compared to injured patients randomized to a u sual care control condition?

In this study of 635 survivors of both intentional and unintentional injury at 25 trauma centers, the intervention was delivered during the first 6 months post-injury. TSOS used a stepped-wedge cluster-randomized trial design with a control group and an intervention group. There were many intercurrent events during this trial, which led to the rethinking of the original design, including a regulatory pause in the final period, when they were only recruiting intervention subjects. The stepped-wedge design made this pause difficult to overcome, and researchers speculated that a parallel group randomized design might have been better.

Enrollment drift occurred in this study. Across the sites, there was a significant increase in PTSD symptoms in intervention patients, who were recruited later than control patients. Parallel group randomized design might have avoided that issue. Differential site variability over time was a key factor influencing enrollment drift.

There was also a great deal of site heterogeneity regarding number of violent injuries, leadership turnover, and more. One advantage of the stepped-wedge design is that each site contributes to both the control and the intervention group. However, differential heterogeneity over time is a challenge for both types of design. The lesson learned at the end of this trial was that regardless of stepped-wedge or parallel group randomized design, site variability considerations favor individual-level randomization when possible.

HiLo

HiLo, which is still in process, is focused on patients who require dialysis to stay alive. Patients without kidney function accumulate phosphate in their blood. Because dialysis gets rid of some of this phosphate but not enough, drugs called dietary phosphate binders are used to lower serum phosphate. However, there is no proof that taking a high phosphate and lowering it will help patient outcomes. There is only observational data suggesting that high phosphate levels are associated with worse survival outcomes.

While phosphate binder medicines lower phosphates, they also have adverse side effects. The goal of HiLo is to compare two phosphate targets—one low and one high— in patients on hemodialysis, and get clinical trial evidence about patient outcomes.

For operational reasons, the researchers agreed to conduct HiLo as a cluster-randomized trial. This trial is different from many PCTs in that the intervention goes against the guidelines, and therefore presents a more-than-minimal risk for participants and requires consent. Therefore, there is an implementation question to consider as well as a scientific question. Cluster randomization combined with informed consent can be a big challenge. Some changes later in the trial were necessary, as electronic consent was useful, and researchers worked with the Duke IRB to streamline this process.

From a pragmatic trial perspective, dialysis patients are a “dream come true” population. They come to a fixed place at a known time, and federal government mandates led to a wealth of data collection in the EHR.

At 10% enrollment, an imbalance between the 2 arms was noticed in willingness to participate. Biased enrollment was a concern, as the dieticians involved knew to which group each patient had been randomized. Although there were already more than 500 patients involved in the trial, the researchers decided to pivot to individual-level randomization. For this reason, the researchers advise others not to do a randomized trial with post-randomization consent. They also advise that while it is important to focus on trying to be pragmatic when feasible, being “pragmatic trial zealots” may end up harming internal validity.

PROVEN

PROVEN was a cluster-randomized controlled trial of an advance care planning (ACP) video intervention embedded within 2 nursing home healthcare systems. Results found that the ACP video did not make a difference in the primary outcome (hospital transfers per 1000 person-days alive among residents with advanced illness).

ACP is mandated by all healthcare systems, and observational studies have shown that patients with ACP are less likely to have burdensome interventions such as hospitalizations, tube feedings, and ICU visits. Because it is difficult to administer ACP on a large scale, ACP videos were considered as a solution.

Participants were randomly assigned to either the intervention or the control arm. All the facilities in the intervention group adopted the new ACP videos (a waiver of consent was used), and they were applied to all nursing home residents during the enrollment period. The trial used EHRs to identify people with the greatest opportunity to benefit from ACP (i.e., those with advanced illness). In other words, everyone in the nursing home was exposed to the intervention, but the primary outcome was tested on the subgroup of patients with advanced illness. The primary outcome was calibrated on Medicare claims. Regarding fidelity, the researchers used a video status report as a user-defined assessment. There was a great deal of variation across participating facilities for various reasons, including staff turnover.

In spite of employing cluster-randomized assignment at the facility level, this trial achieved a balance on most items, with the exception of the percentage of White participants (due to facility location) . However, despite this balance, results showed no difference in the number of hospital transfers per 1000 person-days alive. There was no difference for the primary outcome or for most secondary outcomes.

Thoughts on a study redesign include the possibility of stratifying by facility implementation capacity, and/or restricting the study to a certain population. However, the investigators indicated that there are no clear advantages to using a different kind of design for this study.

SPOT

SPOT was a large trial studying 2 low-intensity suicide outreach interventions. The study sought to evaluate whether health systems should implement either of these programs (care management or dialectical behavioral therapy skills with coach) to reduce risk of suicidal behavior among outpatients reporting frequent suicidal ideation on routinely administered questionnaires. Results found that neither intervention reduced the risk of suicide, and one may actually have increased risk. While researchers had hoped that the intervention would produce more tools for suicide prevention, they do not believe changing the study design would have helped.

All of the individuals enrolled in 4 health systems responded to questionnaires about suicidal ideation. All individuals who responded that they had thoughts that they would be better off dead more than half the days in a week (or nearly every day), and used MyChart, were enrolled in this study via individual randomization. The rationale for individual randomization was that in healthcare, especially mental healthcare, patients see multiple providers, making contamination likely, and this can be an issue with clusters. Therefore, the researchers believed an individual randomization study design where the intervention is delivered in a centralized location separate from clinical care was most likely to prevent contamination.

SPOT began with 3 sites, and a fourth site was brought on to increase enrollment. This population had a wider range of baseline risk than that of other trials. The 2 interventions were dialectical behavior therapy skills training (online) and risk-based care management to facilitate outpatient care. Both interventions were lower intensity than had been previously studied, and were delivered centrally to individuals (no group component). The invitation process consisted of online messages in the patient's EHR, which included expressions of caring and concern as well as descriptions of the specific intervention services. There was also abbreviated information about consent. Reminders were sent at specific time periods.

Each intervention was compared to usual care, and all participants were included in the comparison regardless of whether they took up the intervention. Researchers were not only evaluating these 2 interventions, but also the invitation process. A hypothesis from post-hoc analysis is that repeatedly reaching out to individuals with services in which they are not interested may be more harmful than helpful.

This trial had a composite outcome: nonfatal suicide attempts and suicides. Suicides were identified through state death certificates, while nonfatal suicide attempts were identified through the EHR. The chart adjudication system was rigorous and thorough. Outcomes were examined up to 18 months after randomization.

Dr. Shortreed believes this study took place in the right setting—the healthcare system—and that health systems should be expected to reach out to people at increased risk of suicidal ideation (even if they are not seeking care) a nd then deliver an intervention. However, the researchers did not expect to see an increased risk of harm, which appeared to be related to reaching out to individuals. Although this is a possibility in many studies even when it is unexpected, in studies such as this with serious outcomes, it is important to have more than one person responsible for the reports. This result also emphasizes the value of data safety monitoring boards (DSMBs) for all PCTs. Quality assurance of data in this kind of study can be complicated, especially when the study lasts for several years. DSMBs should include regular quality assurance checks as part of trial monitoring.

What Is the Most Effective Use of the 1-Year Planning Phase?

This panel, moderated by Wendy Weber, ND, PhD, MPH, discussed ways to optimize the trial planning phase to increase successful completion of a PCT. Panelists from four trials shared examples and resources for developing the planning period.

Key Takeaways

- Figure out as many details as possible during the planning phase—for example, determining how to merge and aggregate data across several healthcare systems.
- Use the planning phase to evaluate the population for diversity and equity and adjust the recruitment plan.
- Start the trial's planning phase with sites that will be the most challenging.
- Fail fast: make sure it works, try different things.
- Do repeated, careful, and rigorous checks to ensure fidelity of the data collection process during the planning phase.
- A site readiness checklist for data collection, intervention, and IRB compliance was essential for communication across the team to make sure all milestones were reached.
- While officially there are 12 months slated for the pilot, time for writing up results is included, meaning there are really only 10 months for conducting the pilot.
- There is a need for nimble data collection that allows study leadership to know what is happening on the ground both during the planning phase and in the ongoing study.

Panelists and Case Studies

- Pragmatic Trial of User-Centered Clinical Decision Support to Implement Emergency Department-Initiated BuprenorphinE for Opioid Use Disorder (EMBED)
  Gail D'Onofrio, MD
- Personalized Patient Data and Behavioral Nudges to Improve Adherence to Chronic Cardiovascular Medications (Nudge)
  Michael Ho, MD, PhD
- Improving Advance Care Planning: Promoting Effective and Aligned Communication in the Elderly (ACP PEACE)
  James Tulsky, MD
- A Policy-Relevant U.S. Trauma Care System Pragmatic Trial for PTSD and Comorbidity (Trauma Survivors Outcomes and Support [TSOS])
  Doug Zatzick, MD

EMBED

EMBED was a 5-year project with an 18-month pragmatic parallel group randomized trial across 21 sites at 5 healthcare systems. The intervention was user-centered clinical decision support, and the primary outcome was initiation of buprenorphine in the emergency department among patients with opioid use disorder.

Challenges in the planning phase included a change in participating health systems (and a change in EHR vendor) as well as a switch from a stepped-wedge study design to parallel group randomized. EMBED was able to extend the project to have more time to finish the study, analyze the results, and disseminate the findings.

During the planning phase, EMBED researchers found that coding practice related to patient race/ethnicity varied not only from system to system but also within systems, which was a challenge to reconcile. There were also variables involving non-EHR data sources, which required planning ahead and coordination. They highlight the need for repeated, careful, and rigorous checks to ensure fidelity of the data collection process during the planning phase. EMBED researchers especially recommend monthly data pulls during the trial, rather than one single data pull at the end. In addition, a site readiness checklist for data collection, intervention, and IRB compliance was essential for communication across the team to make sure that all milestones were reached.

Nudge

Nudge was a patient-level randomized study across 3 health systems, with a primary goal of improving adherence to chronic cardiovascular medications. The primary outcome was medication adherence as defined by pharmacy refill data, and the secondary outcomes were clinical measures, clinical events, and healthcare utilization. The population included patients who were diagnosed with one or more of several cardiovascular conditions (hypertension, hyperlipidemia, diabetes, coronary artery disease, atrial fibrillation) and prescribed at least one of several medication classes of interest.

The study identified patients via the EHR. They were then sent a letter about the study that included opt-out information. Patients who didn't return the opt-out letter were monitored over time for gaps in their medication refills—if they did not refill their prescription after 7 days, they were considered eligible, and were randomized to usual care, generic text messages, optimized text messages (with behavioral nudges), or optimized text messages plus an artificial intelligence chat bot. The aim of the chat bot was to identify why patients were delaying refills and provide solutions. Year 1 milestones included a trial of the patient identification process via the EHR.

Lessons learned from the 1-year planning phase included the need to be aware of the short planning time frame. While officially there are 12 months slated for the pilot, time for writing up results is included, meaning there are really only 10 months for conducting the pilot. Nudge had some administrative barriers related to contracting. For example, 2 of their health systems required Surescripts pharmacy data at more frequent intervals, and this took time to fix.

Forming collaborative partnerships at each site was helpful, especially in coming up with interim solutions when COVID-19 hit in the middle of the intervention. It was also helpful to have a project manager who had experience managing a multisite study. Finally, working with the IRB ahead of time to ensure that they approved the opt-out strategy was an efficient use of time.

ACP PEACE

ACP PEACE was a pragmatic stepped-wedge cluster-randomized controlled trial testing 2 interventions—VitalTalk (a clinician communication skills training program) and AC P Decisions (a video library educating patients on decision-making). It involved 3 d isease-focused health systems and 36 oncology clinics, with 30,000 patients enrolled. The original design included 6 steps with 12 clusters of 3 clinics each, with all data coming from EHRs. Unfortunately, COVID-19 necessitated a revised design, as there were 6 months in which research could not be conducted.

Since COVID-19 had changed institutional approaches to ACP, ACP PEACE decided to learn from this development by using the data they had from steps 1 and 2 and comparing that to steps 3-6. They had a new baseline for the trial, and ended up including more clinics. Measuring ACP rates pre- and post-COVID enabled them to complete the study as well as learn about the effect of COVID-19 on ACP.

ACP PEACE had planned to use structured data elements from the EHR to examine ACP. However, a chart review of data elements during the planning phase revealed that these were inaccurate—fewer than 60% had the correct documents. ACP PEACE had the advantage of having a laboratory that does natural language processing, which made it possible to extract the same data with far more accuracy.

TSOS

The TSOS trial tested an innovative collaborative care intervention for patients admitted nationally to Level 1 trauma centers. There were 635 patients randomized into the trial, and baseline began with an EHR PTSD screen to identify people with high symptom levels who were at risk for developing PTSD. The main outcome was PTSD symptoms assessed via interview at 3, 6, and 12 months post-injury. TSOS found significant reductions in PTSD symptoms at 6 months but not at 12.

The planning phase for TSOS was complex in that TSOS researchers wanted to recruit 25 diverse sites from national Level 1 trauma centers while simultaneously piloting the procedures. TSOS researchers were interested in balancing site characteristics and using sites beyond just academic centers. The TSOS conceptual framework was informed by implementation science, as they used the RE-AIM (Reach, Effectiveness, Adoption, Implementation, and Maintenance) framework to get a representative sample of Level 1 trauma center sites.

TSOS had planned to use an automated EHR PTSD screen, which, had it worked efficiently, would have saved time that could then be used for recruitment and the intervention. But, TSOS researchers were unable to capture variability across all sites, and many sites were unable to fully automate the screen.

Lessons learned from TSOS include the need for nimble data collection that allows study leadership to know what is happening on the ground both during the planning phase and in the ongoing study.

Discussion Questions

What is one thing you did during the planning phase that set your trial up for success?

Panelists described checking and validating data early, hiring and funding quality IT people, and having champions at each site to establish relationships with people who can help rectify issues.

If you had it to do over, are there any activities you would add to your planning phase? If you piloted elements, would you change what you piloted?

Researchers recommended more robust IT checks and training, preparing for variability, and getting better buy-in from sites by having more than one partner at each site.

What advice do you have for working with healthcare system partners during the planning phase? How can the planning phase be used to enhance diversity of participants?

Panelists described the importance of reframing study goals so that healthcare system partners see the intervention is potentially beneficial to them and their work. Furthermore, using the EHR should reduce selection bias for participants, but time must be spent relationship-building with healthcare system partners before getting the PCT grant. The grant process makes it difficult to increase diversity later on, because budget restrictions make it difficult to add sites.

Including smaller rural and community hospitals can be difficult due to funding, and some trials ran into regulatory problems at sites that were not accustomed to doing research.

However, doing trials at these sites is important because they can inform policy and payer decisions.

What activities in your planning phase were helpful to determine the quality of data or how you merged/aggregated data?

Panelists shared several key activities, including:

- Testing EHR data elements, interpretation, and data collection
- Collecting data monthly
- Creating an agreed-upon data dictionary
- Having a good IT person on board at each site

How Do We Implement Interventions in Complex Healthcare Systems?

This panel, moderated by Steven George, PT, PhD, discussed how to think about the complexity of both the intervention and the healthcare system in which it is delivered. Panelists from four trials identified implementation science concepts and highlighted strategies that supported success.

Key Takeaways

- One reason complexity may be underestimated is that it is hard to measure. The NIH Pragmatic Trials Collaboratory is developing a tool to share with trialists that can help to identify and measure the complex components of interventions in the planning stage.
- A key to success was developing a protocol that fit within normal business operations and the busy world of pediatric clinics.
- The EHR can be used to alert doctors that an incoming patient might be a good fit for the trial.
- Being nimble and flexible is important for the study team when partnering with clinics that are new to pragmatic research. Be welcoming to the clinics' ideas for adapting or improving the intervention.
- Researchers will need to continue to define and develop conceptual models (from implementation science) that inform their work.
- There are more and more training opportunities in implementation science available, including mentored programs at various universities. If a researcher's institution does not have implementation science expertise available, it is important to find a way to get it.
- Map out a workflow for the trial, so it is clear which person is doing what in each health system. This helps relieve clinics' concerns about too much work or

responsibilities for the research, and it helps with efficient site startup and potential employee turnover—a major reality for PCTs.

- Disparities can be addressed by thinking carefully about the populations you want to reach. Consider how to reduce barriers to the intervention and include reducing disparities as an endpoint. Support participation of community health centers, which are well-positioned to reach underserved populations but are also under-resourced.
- When interventions actually work in complex healthcare settings, and healthcare systems have the ability to disseminate them, there is a potential to have an impact.

Panelists and Case Studies

- Guiding Good Choices for Health (GGC4H): Testing Feasibility and Effectiveness of Universal Parent-Focused Prevention in Three Healthcare Systems
  Margaret Kuklinski, PhD
- Hybrid Effectiveness-Implementation Trial of Guided Relaxation and Acupuncture for Chronic Sickle Cell Disease Pain (GRACE)
  Miriam Ezenwa, PhD, MSN, BSN, RN, FAAN
- Strategies and Opportunities to Stop Colorectal Cancer in Priority Populations (STOP CRC)
  Gloria Coronado, PhD
- Improving Chronic Disease Management with Pieces (ICD-Pieces™)
  Miguel Vazquez, MD

Trial Descriptions

GGC4H was a cluster-randomized trial evaluating the use of an evidence-based parenting program in primary care to reduce substance use initiation in adolescents. The intervention consisted of 6 sessions delivered over 6 weeks.

GRACE is a hybrid effectiveness-implementation trial studying guided relaxation and acupuncture to manage pain in patients with sickle cell disease.

STOP CRC was a cluster-randomized trial evaluating an approach to increase colorectal cancer screening in 40,000 patients in federally qualified health centers.

ICD-Pieces was a cluster-randomized trial studying whether a novel technology platform could help primary care physicians treat patients in more effective ways to reduce hospitalizations, emergency department visits, cardiovascular complications, and deaths.

Discussion Questions

In planning your trial, how did you think about intervention and system complexity—and how did it change during trial conduct?

GGC4H described having a lot of complexity to navigate. They knew it was essential to work with research teams (including PIs, clinical staff, informaticians, biostatisticians) that had experience doing embedded research in health systems. Having existing relationships, experience, and the right mindset made GGCRH's work with the clinics go well from the beginning. A key to success was developing a protocol that fit within normal business operations and the busy world of pediatric clinics. For example, they gave doctors scripts they could use to recommend families to the program and offered them prescription pads to give parents information about the program. They also used the EHR to alert doctors that a patient was coming in who might be a good fit for the trial. The GGRH team tried to adopt a mindset of patience for the unexpected. For example, during COVID-19, they had to deal with pediatric clinics being reorganized in one of the health systems, which threatened the trial's randomization.

ICD-Pieces' intervention had the complexity of addressing 3 conditions in the context of 4 different health systems with a lot of diversity. They found it helpful to have partners who knew how to do research, but pragmatic research was not common among their collaborators. The health systems were receptive and willing to help the research team, while the researchers had to be very flexible on delivery of the intervention.

In GGC4H, the notion of a clinical intervention that becomes part of usual care was unfamiliar to the clinics. The research team had to overcome the clinics' ideas of traditional research. Researchers found being nimble and flexible to be most important, and they welcomed the clinics' ideas to adapt and improve the intervention.

How did you incorporate what is known about implementation science strategies in your trial?

STOP CRC used Plan-Do-Study-Act (PDSA) c ycles to help their clinics become partners in implementing the program. These partnerships helped the research team understand barriers. They expected that each clinic would test a reminder for the program, but some clinics wanted to work on their workflow. The PDSA cycles identified all kinds of challenges that the research team would not otherwise have known existed. STOP CRC merged its program with a standard quality improvement framework that the clinics were interested in. Organizational characteristics influenced the success of program implementation.

GRACE's hybrid effectiveness-implementation trial sought to document barriers and find solutions. They used the Consolidated Framework for Implementation Research (CFIR) [1] to focus on different levels, such as the policy level. GRACE included people with expertise in many different areas and engaged multiple stakeholders. For example, they interviewed patients and providers to be culturally sensitive, and used videos to show patients what acupuncture is like so they could decide if they wanted to participate in the trial.

GGC4H incorporated implementation science explicitly through the Practical Robust Implementation and Sustainability Model (PRISM)[2] and RE-AIM frameworks,[3] which were a huge help for study design and measurement. They used implementation science principles and strategies early on and throughout the trial. When COVID-19 hit during the trial, GGC4H engaged parents by holding focus groups to see if parents wanted to connect and engage with the program even during the difficult time. The research team also made sure to maintain connection with pediatricians in the clinics to tell them how the trial was going and to share quotes from parents. GGC4H is doing a qualitative study using the PRISM principles involving interviews with pediatric and non-pediatric leaders and staff to understand from multiple perspectives support for preventive interventions. All of these efforts will help bridge implementation to sustainability. More attention to implementation science can increase appreciation for the systems-level factors that influence research.

Panelists shared that choosing the type of hybrid design and emphasis for a trial will depend upon the literature and foundation of knowledge to date. Researchers will need to continue to define and develop the conceptual models that are informing this work.

Where did you find implementation science expertise for your trial?

Panelists described two main ways they were able to access implementation science expertise for their research: investigators trained in implementation science on their research team or at their institution, and partnerships with collaborators at other institutions. The STOP CRC co-PI had participated in an implementation science training program, and they were also fortunate to work with the developer of RE-AIM framework, which bolstered their emphasis on implementation science. The research team also partnered with collaborators to do some more state-of-the-art analyses for the trial. Similarly, GRACE shared that having partners they had worked with at different institutions in the past was critical. Through these partnerships, they were able to have expertise in anthropology, Chinese medicine, and acupuncture for their trial.

The panel recognized that there are more and more training opportunities in implementation science available, including mentored programs at various universities. If a researcher's institution does not have implementation science expertise available, it is important to find a way to get it.

What is the value of mapping out the workflow before you start a trial? And how were those maps used during the trial?

ICD-Pieces described the extreme value of mapping for their trial. They learned the importance of working together and individually with each health system to create individualized mapping. Being clear about which person is doing what in each health system helped relieve practices' concerns about too much work or responsibilities for doing the research. Additionally, having clear workflows helped with efficient site startup and working through employee turnover, which is a major reality for PCTs.

STOP CRC and GGC4H described data mapping for their trials. Even when all clinics use the same EHR system, they do so in different ways. Mapping helped to identify data available in the EHR, ensure the accuracy of reports, and verify of who is doing what from the perspective of the data.

Flexibility was a common theme. Mapping helped the GRACE trial give their clinics options; for example, they could accommodate differences in the timing of blood draws during patient encounters. The GRACE trial had to be flexible because patients with sickle cell disease might have too much pain to attend their scheduled visit, or come to their visit and have too much pain to complete the intervention. However, trialists must be mindful in considering when an implementation strategy becomes an adaptation. Questions to think about include what is too flexible and what adaptations need to be tracked. The STOP CRC trial made sure to be clear about what the core components were and what could be adapted. Whatever the clinics did was considered an adaptation, and the research team got a lot of good information with periodic reflections. GGC4H had a member of the research team tracking departures from what was originally planned. The tracking will be helpful when interpreting their outcomes data.

How can PCTs be conducted to reduce disparities?

Panelists agreed implementation science provides a great opportunity to address disparities in PCTs, especially when working with health systems. They made several suggestions, including:

- Think carefully about the populations you want to reach.
- Include reducing disparities as an endpoint.
- Consider how to reduce the barriers to the interventions.
- Provide transportation for patients or have the trial come to the patients.
- Support participation of community health centers, which are well-positioned to reach underserved populations but are also under-resourced.
- Consult relevant stakeholders on the study goals and any modifications needed for the intended population (e.g., staffing, language, cultural sensitivity).
- Consider new implementation science models that look at disparities in terms of reach.

STOP CRC chose an intervention that was more effective in Latino patients, which have a low uptake for colorectal cancer screenings. They worked with their most diverse clinic as a pilot, which helped them understand how to engage Latino patients, identify issues, and develop effective patient instructions (Spanish language and wordless). By starting with that clinic, they could develop culturally relevant tools for use by all of the trial’s clinics.

What are the advantages of a hybrid design? What is the definition of a well-established intervention?

Panelists described that all PCTs intersect with implementation science. When interventions actually work in complex healthcare settings, and healthcare systems have the ability to disseminate them, there is a potential to have an impact. if we can crack the implementation barriers. Researchers have to weigh the pros and cons of different study designs in light of the body of evidence for an intervention.

Sustainability is a critical question. STOP CRC brought together an advisory board, including community health center clinics, to talk about barriers of colorectal testing. A major barrier identified was that patients might face co-pays when they need a colonoscopy procedure as a follow up to fecal testing. The researchers had to advocate for policy and coverage changes to make the study sustainable.

References

1. Consolidated Framework for Implementation Research. Available at: https://cfirguide.org/.
2. McCreight MS, Rabin BA, Glasgow RE, et al. Using the Practical, Robust Implementation and Sustainability Model (PRISM) to qualitatively assess multilevel contextual factors to help plan, implement, evaluate, and disseminate health services programs. *Transl Behav Med.* 2019.
3. Glasgow RE, Harden SM, Gaglio B, et al. RE-AIM Planning and Evaluation Framework: Adapting to New Science and Practice With a 20-Year Review. *Front Public Health.* 2019;7:64.