REVIEW ARTICLE

THE CHANGING FACE OF CLINICAL TRIALS
Jeffrey M. Drazen, M.D., David P. Harrington, Ph.D., John J.V. McMurray, M.D., James H. Ware, Ph.D., and Janet Woodcock, M.D., *Editors*

# Integrating Randomized Comparative Effectiveness Research with Patient Care

Louis D. Fiore, M.D., M.P.H., and Philip W. Lavori, Ph.D.

From the Department of Veterans Affairs (VA) Cooperative Studies Program, Massachusetts Veterans Epidemiology Research and Information Center, VA Boston Healthcare System, Boston (L.D.F.); and the Department of Biomedical Data Science, Stanford University School of Medicine, Stanford, CA (P.W.L.). Address reprint requests to Dr. Lavori at the Department of Health Research and Policy, HRP Redwood Bldg., 259 Campus Dr., Stanford University, Stanford, CA 94305-5405, or at lavori@stanford.edu.

N Engl J Med 2016;374:2152-8.
DOI: 10.1056/NEJMra1510057


Clinical trials that are embedded into usual care have the potential to yield outcomes of great relevance to the institutions where they are performed and at the same time to yield information that may be generalizable to the health care system at large. In this article, we review four clinical trials that were conducted in three health care systems using their extant electronic health record (EHR) systems. We find that EHR-based clinical trials are feasible but pose limitations on the questions that can be addressed, the processes that can be implemented, and the outcomes that can be assessed. We think that the current requirements for ethics review should be reconsidered for such trials, in which the risk for participants that can be attributed to research is low.

## Features of Comparative Effectiveness Trials

Randomized comparative effectiveness trials compare the effects of a number of treatments in current use on clinical outcomes in order to guide decision making. The emphasis on clinical goals and decisions distinguishes comparative effectiveness trials from trials that are designed to compare an experimental treatment to a control, to establish proof-of-concept, or to elucidate a mechanism of action.

Despite this simple definition, comparative effectiveness trials vary considerably in the degree to which they are integrated with or segregated from clinical practice. Some follow usual care closely, whereas others deviate from usual care by using sophisticated methods. Comparative effectiveness trials that are segregated from usual care have been the subject of a great deal of study and commentary, and we do not consider them in this article. The comparative effectiveness trials that concern us here seek to blend themselves unobtrusively into normal clinical operations, limiting the research perturbation to the bare minimum necessary to settle the questions that motivate them. The aims of integrated comparative effectiveness research and segregated comparative effectiveness research are shared by quality-improvement efforts, and the boundaries among them are not sharp (Table 1).

Converting naturally observed treatment variation into experimental manipulation challenges the pragmatic goals of comparative effectiveness research by disturbing normal clinical operations. Experimental comparative effectiveness research may intrude on normal clinical operations with extra baseline and outcome assessments, procedures for informing patients and obtaining their consent for research, and regulatory requirements for the training of personnel engaged in research. The integration of comparative effectiveness research into clinical practice retains the minimally intrusive effects of observational research while offering the strengths provided by the experimental method (including randomization).

**Table 1. Comparison of Quality Improvement with Integrated Comparative Effectiveness Research and Segregated Comparative Effectiveness Research.***

| Variable | Quality Improvement | Integrated Comparative Effectiveness | Segregated Comparative Effectiveness |
|---|---|---|---|
| Proponent | Clinical care providers | Researcher or clinical care providers | Researcher |
| Funding source | Health care system | Health care system or research enterprise | Research enterprise |
| Site | Typically single site or single health care system | Single site or multiple health care systems | Single site or multiple health care systems |
| Regulation | Regulated under local health care system rules | Regulated as research, as implied by creation of generalizable knowledge | Regulated as research |
| Design | Randomization typically not used | Randomization accepted as standard | Randomization accepted as standard |
| Persons conducting the study | Clinicians and administrators who are not considered to be engaged in research | Clinicians and research staff who may or may not be considered to be engaged in research | Clinicians and research staff who are usually considered to be engaged in research |
| Primary implementation tool | EHR system | EHR or specialized study software and forms | Specialized study software and forms |
| Product | Practice improvement | Practice improvement and research publication | Research publication |

* EHR denotes electronic health record.

**Table 2. Comparison of Critical Elements in the Trials.***

| Trial | Patient Eligibility and Recruitment | Ascertainment of Outcomes | Regulation and Governance |
|---|---|---|---|
| Chlorhexidine bathing trial[1] | Cluster randomization with automatic enrollment on ICU admission | Inpatient chart review of all participants by infection-control personnel | Waiver of informed consent, with clinicians who were not considered to be engaged in research |
| Retropro and eLung trials[2] | Alert to clinician for possible eligibility by flagging software or the enrollment of the patient by the clinician directly through the trial website | System developed to aggregate outcomes from outpatient records | Informed consent obtained by clinicians who were considered to be engaged in research |
| Insulin administration trial[3] | Randomization allowed by physician at the insulin-order entry screen in EHR | Computed by structured data from the EHR and corporate data warehouse | Informed consent obtained by the study nurse; patient referred by clinician who was not considered to be engaged in research |

* ICU denotes intensive care unit.

This article describes the successes and failures of various attempts to integrate comparative effectiveness research into practice.

We use four case studies to show the major features of integration of research into practice. For each trial, we examine the effects of the research on the eligibility and recruitment of patients, on the follow-up of patients for outcomes and safety monitoring, and on the regulatory and governance issues it raises (Table 2). The examples also show the different challenges that are encountered by an attempt at the integration of research into practice that starts with a research approach and seeks to map research operations onto the clinical context, as compared with a practice that starts with a quality-improvement approach and borrows strength from research methods. We chose these case examples because they represent the critical approaches to research integration. This review is

not meant to be a comprehensive review of all such research methods.

## EXAMPLE VIGNETTES

### CHLORHEXIDINE BATHING TRIAL

This trial is an example of "quality improvement first" integration, leading to a generalizable research result. Noto et al. used a cluster-randomized crossover trial to compare daily bathing with the use of chlorhexidine-impregnated cloths with daily bathing with the use of nonantimicrobial cloths with regard to the incidence of infection among critically ill patients.[1] Each treatment was performed for a 10-week "bathing period," and after a 2-week washout period, the alternate bathing treatment was performed for 10 weeks. Each intensive care unit (ICU) crossed over between bathing assignments three times during the trial.

Patients who were admitted to the ICU during the bathing period were enrolled unless they had a known allergy to chlorhexidine or they were admitted with burns, toxic epidermal necrolysis, or the Stevens–Johnson syndrome or "if the treating physician felt bathing would be unsafe." The authors reported that the "study was conceived as an institutional quality improvement project" and was reviewed as such by the institutional review board, which approved a waiver of informed consent. After completing the enrollment of the patients, but before analyzing the data, the authors reported realizing the possible external interest in the results, which caused them to treat it as a research study from that point on.

### RETROPRO AND ELUNG TRIALS

These two trials, one regarding statins and the other regarding antibiotic agents, show the challenges of an attempt to use the full range of standard research methods in a clinical context. The two trials[2] were intended to test the feasibility of integrating patient-level randomized, comparative effectiveness research into the practice of general practitioners in the United Kingdom, with the use of a system that was based on EHRs from the clinician's office. Qualitative and quantitative methods were used to study the effect of the research integration on patients and providers and to assess the quality of the trial data and conduct. In both trials, the treatment assignment was revealed only after randomization. General practitioners verified eligibility and obtained informed consent from the patients, accessing the trial website to record both and to obtain the randomized assignment.

Retropro compared simvastatin with atorvastatin in patients older than 40 years of age who had a risk of cardiovascular disease that was more than 20% over a 10-year period; the trial excluded patients with a history of statin use or a diagnosis of cardiovascular disease or liver disease. Clinical outcomes included cardiovascular disease during the trial period. Data from the EHRs were used to calculate baseline risk, with the use of the Framingham score and variations, and to ascertain outcomes.

The eLung trial compared immediate (prophylactic) use of antibiotics with deferred use in patients older than 40 years of age who had an acute exacerbation of underlying chronic obstructive pulmonary disease (COPD) that was characterized by increased nonpurulent sputum volume but who, in the opinion of the practitioner, did not require immediate referral to a specialist. Clinical outcomes included the forced expiratory volume in 1 second and quality of life.

### INSULIN ADMINISTRATION TRIAL

This trial compared two methods (a sliding scale vs. a weight-based regimen) for determining the dose of subcutaneously administered insulin to be used in hospitalized patients.[3] This trial shows the issues that arose in mapping key research tasks onto existing informatics resources at U.S. Veterans Affairs (VA) hospitals. Order-entry screens at three VA hospitals were modified to include an option to enroll participants into the trial comparing these two regimens (Fig. 1). Election of the menu choice "no preference for insulin regimen" triggered the EHR workflow to notify the research nurse to obtain informed consent, automatically place a note of participation in research in the medical record, and randomly assign treatment.[4] The primary outcome of this ongoing trial was the length of hospital stay, and secondary outcomes included measures of glycemic control, all of which were ascertained from the EHR database.

## ELIGIBILITY AND RECRUITMENT OF THE PATIENTS

All four trials used the locally collected clinical data to determine whether a given patient met

the overarching criteria for inclusion. The treating clinician made the final decision regarding eligibility.

In the chlorhexidine bathing trial, recruitment was automatic once it was approved by the treating clinician. In fact, no patients met the exclusion criteria, so all the patients who were admitted to the ICU during one of the four bathing periods were included in the primary analysis.

In the Retropro and eLung trials, flagging software that was developed before the start of the trials was intended to notify clinicians of patients who were eligible for the trial during a patient visit. During the development of this software, the investigators studied attitudes of physicians and primary care nurses toward recruitment for randomization, as well as the use of software tools developed for that purpose. Not surprisingly, the top concern was the lack of time for recruitment during a patient visit. Nurses were more comfortable than physicians with regard to obtaining informed consent during the consultation. Physicians, more often than nurses, thought that financial incentives to the participating health care providers might help. Physicians also indicated that clinical uncertainty would help motivate the participation of providers. Both groups of providers believed that the doctor–patient relationship would facilitate recruitment, although it could be a barrier. In these two trials, the flagging software proved to be the most problematic part of the implementation, since it required considerable training, troubleshooting, and tailoring to the individual trial. However, once the clinicians knew how the recruitment system worked (enrollment of the first patient was considered to be the most difficult part of the process), recruitment was simple.

The insulin administration trial identified as eligible any patient for whom an order for insulin was to be placed. The patient's physician made the final determination of eligibility on the basis of his or her assessment of the relevant characteristics of the individual patient.

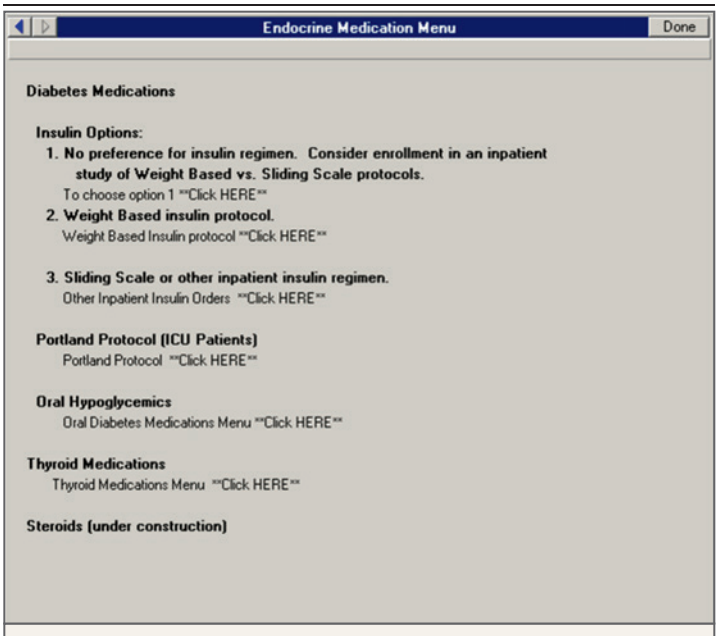


**Figure 1. Screen Shot of the Endocrine Medication Menu.**

Order-entry screens at three hospitals were modified to include an option to enroll participants in a trial comparing two regimens for the administration of insulin. Election of the menu choice "no preference for insulin regimen" triggered the electronic health record workflow to notify the research nurse to obtain informed consent, automatically place a note of participation in research in the medical record, and randomly assign treatment. ICU denotes intensive care unit.

## FOLLOW-UP OF PATIENTS FOR OUTCOMES AND SAFETY MONITORING

In the chlorhexidine bathing trial, trained infection-control personnel (who were unaware of the trial-group assignment) reviewed the EHR to determine the primary outcome, which was a composite of several hospital-acquired infections. Secondary outcomes such as in-hospital death and length of stay were also available. Thus, this trial took advantage of the fixed, short time horizon for outcomes and the inpatient setting (as well as the available expert staff) to minimize extra effort and expense in the ascertainment of outcomes. Outcomes that were not routinely collected in the ICU, such as the acquisition of multidrug-resistant organisms, were not recorded. By design, no special efforts were made to enhance adherence of the health care providers and patients to the protocol, and no safety monitoring beyond usual ICU practice was performed.

In the Retropro and eLung trials, data on outcomes were collected during outpatient follow-up, which required a substantial investment in a system that aggregated EHR content. Considerable work was needed to process the raw data from clinical encounters into a form that was usable for assessing outcome. These trials were conducted under Good Clinical Practice (GCP) standards, which have specific requirements concerning the quality of trial data.

The insulin administration trial was designed to study outcomes on the basis of structured data elements that were easily obtained from the EHR, such as dates of hospital admission and discharge and blood-glucose levels. More complicated outcomes such as infection (possibly due to hyperglycemia) would require more sophisticated and costly data processing and management and therefore were not studied, which shows one of the trade-off decisions that are inherent in conducting point-of-care trials.

## REGULATION AND GOVERNANCE

All four trials were conducted under standard procedures for protecting human participants from research risks. The chlorhexidine bathing trial had a single ethics review as a quality-improvement initiative, which sufficed even after the study was reconceived as a clinical trial and came under regulation for the protection of human participants from research risks. The waiver of informed consent greatly simplified the integration of the trial into clinical care. By contrast, the two U.K. trials (Retropro and eLung) and the VA trial regarding insulin administration were designed as research studies, thus requiring multiple levels of approvals and informed consent from individual patients.

The leaders of the Retropro and eLung trials began their summary of the main lessons learned by asserting that, "Electronic health record point-of-care trials are feasible, although recruitment of clinicians is a major challenge due to the complexity in trial approvals." The details of the trial-approvals process are sobering, if not particularly surprising, as described in the investigators' own words[2]:

> The overarching [National Health Service] governance review took 2 years from original application to approval, followed by local approvals (which took a further year in England, but only 2 months in Scotland). Several regions demanded local modifications of the trials, including localised consent forms and, because of prescribing guidelines, mandatory switching from atorvastatin to simvastatin in Retropro 3 months after trial entry. Several [general practitioners] were also warned that Retropro would adversely affect their statin performance targets (most regions restricted atorvastatin prescribing). Review by the ethics committee resulted in a considerable lengthening of the informed consent form.

The Retropro and eLung trials were conducted under GCP procedures. GCP requirements mandated a detailed, comprehensive review of all trial components before approval. The GCP standard includes site monitoring for compliance with the protocol, detection of threats to data integrity, a strong preference for structured data-collection forms, and extensive documentation of adverse events. The conventional EHR system could not satisfy the GCP standard without extensive reengineering. Thus, the mismatch between EHR and GCP standards imposed a substantial burden on the participating physicians, including paperwork and training before they were allowed to recruit patients for the trials.

The practitioners in the Retropro and eLung trials provided crucial input and were not compensated for their time and effort, nor were they recognized individually for academic contributions. For the Retropro trial, the number of practices that expressed an interest was 270 of 459 that were contacted (58.8%), but only 50 completed a site submission, 35 completed GCP training, and 17 (3.7%) recruited patients. The results with respect to the eLung trial were similar, with only 6 of the 459 contacted sites (1.3%) actually recruiting patients. The authors conclude[2]:

> It is unclear why point-of-care trials not intended for regulatory submission also need to comply with GCP. The fundamental question is why point-of-care trials are viewed as an activity that requires elaborate governance procedures rather than as quality improvement that is an intrinsic part of routine clinical care.

The insulin administration trial underwent full review by the institutional review board at each of the three participating sites and required written informed consent from the patients, which was obtained by trained research personnel, not by the treating clinician. Clinicians who referred their patients for inclusion in the trial

were therefore not considered to be engaged in research, despite the fact that they signed the randomly generated order for insulin administration, and thus they were not required to undergo research training and credentialing.

## LESSONS LEARNED

The delegation of patient recruitment and consenting processes to providers at the point of care imposes a burden that greatly reduces the enthusiasm and willingness of providers to participate.[5,6] Minimizing the time and effort required from clinicians to identify eligible patients and obtain informed consent contributes to successful embedded trials. In the chlorhexidine bathing trial, the cluster-randomization design with waiver of consent and the inclusive selection criteria (which did not exclude a single patient) facilitated enrollment.

In the Retropro and eLung trials, patient-level randomization, more complicated eligibility criteria, and a requirement for additional software (beyond the EHR application) raised barriers to clinician participation. Alternatives include centralized identification and the recruitment of patients before health care encounters, allowing telephone-based centralized processes with waiver of documentation, and allowing informed consent to be obtained early during hospital or clinic intake, followed by simple notification at the time of randomization.[7]

Other simplifying alterations of the informed-consent process may provide appropriate protections, matched to the actual risks of embedded comparative effectiveness trials, while allowing efficient learning to proceed. In a public dialogue conducted for the Health Research Authority in the United Kingdom, almost all groups expressed support in principle for simplified informed-consent processes in appropriate low-risk trials of already licensed drugs and other interventions in common use.[8] In a population-based survey conducted in the United States, when faced with the trade-off of requiring documentation of consent and allowing comparative effectiveness research to go forward, a majority of survey respondents who were asked to imagine participation in a hypothetical trial preferred to forego documentation rather than see valuable research halted.[9]

To enable embedded trials, EHR systems must have sufficient flexibility to allow the creation of workflows that support study-specific enrollment processes with minimal perturbation of clinical care. But flexible systems such as the VistA application that is used in the VA are not widely used. Realizing the full value of embedded trials requires a substantial reengineering of EHR systems.

Embedded trials work best when primary outcomes can be derived from the medical record with minimal human input. The primary prespecified outcome in the chlorhexidine bathing trial was a clinically compelling composite of central catheter–associated bloodstream infection, catheter-associated urinary tract infection, ventilator-associated pneumonia, and *Clostridium difficile* infection. But ascertainment required manual chart review for all 9340 participants.

Outcomes in the eLung trial (hospital admission for COPD exacerbation and the prescribing of oral glucocorticoids) and the Retropro trial (repeat statin prescribing, death, and cardiovascular disease during the trial) were readily ascertained from data that were initially recorded in the EHRs and then aggregated for analysis, which obviated the need for exhaustive manual chart review. Both the primary outcome (length of hospital stay) and the secondary outcome (glycemic control as determined by the glucose level) in the insulin administration trial were ascertained from structured data elements in the medical record, with the use of fully automated procedures. Thus, to avoid additional complexity and cost of outcome ascertainment outside the EHR, study planners must balance clinical relevance with technical feasibility and cost.

Care providers have three central roles in embedded research: they must be engaged as active partners in defining the objectives of the research, they must assent to randomization for each patient, and they must agree to deliver treatments in accordance with the protocol. However, they should not be considered to be engaged in research, from a regulatory standpoint, for two principal reasons. First, the regulatory burden thus introduced inhibits the recruitment of clinician participants, with little or no effect on the safety of patients or the quality of the study in embedded trials. More important, the proponents and sponsors of an embedded trial must strive to maintain the clinician's independence as the patient's advocate and final gatekeeper against inappropriate inclu-

sion in an embedded trial. The independence of providers will be bolstered by an explicit reminder that the clinician is engaged in protecting and treating the patient rather than conducting the research.

Determination of the final eligibility of a patient and the recommendation for inclusion in the trial by clinical care providers — a feature of each of the four trials presented — creates a trade-off. Variation among clinicians' judgments (not rigorously constrained by a recruitment protocol that has been implemented by trained research staff) results in a mix of patients, filtered for equipoise by practitioners rather than by primary investigators who are following the instructions of trial designers. Heterogeneity that is introduced by relaxation of the rigor of inclusion criteria leads to an imprecise biologic specification of the reference population, but results may be more informative for health care system practices.

## DISCUSSION

The goals of randomized comparative effectiveness trials can be realized by the integration of randomized treatment assignment within the usual care ecosystem and by the collection of baseline and outcome data by traditional observational study methods (i.e., from the patient's EHR rather than from specific study-visit forms). As such, the limitations of these point-of-care studies are inherited from the observational substrate; trials conducted in this manner are not traditional full-strength pragmatic studies stripped of technique but rather should be considered observational comparative effectiveness studies that are enhanced by randomization.

The point-of-care method is best suited for interventions that are in common use and regarding which there are both uncertainty among clinicians and a strong desire to explore comparative effectiveness, that have well-described toxicity profiles (allowing risk-based monitoring native to clinical practice), and that are situated where study operations can be implemented with minimal perturbation of the clinical care ecosystem. Ideally, the EHR should be able to be configured to accommodate study-specific workflows (such as the creation of specialized study-based pop-ups and order-entry screens) with back-end data linked to other databases (such as outpatient record systems).[10] All study events that are driven by automatic processes and that affect clinical care (such as the determination of eligibility) must be verified by human experts.

Embedded comparative effectiveness research relies on the engagement of care providers and health care systems as active partners in defining the objectives of the research rather than as passive consumers of its product. Exploiting the full potential of point-of-care methods includes rethinking and redefining traditional ethical and regulatory standards (including informed consent and engagement in research) in this paradigm of low risk that is attributable to research.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

## REFERENCES


1. Noto MJ, Domenico HJ, Byrne DW, et al. Chlorhexidine bathing and health care-associated infections: a randomized clinical trial. JAMA 2015;313:369-78.
2. van Staa T-P, Dyson L, McCann G, et al. The opportunities and challenges of pragmatic point-of-care randomised trials using routinely collected electronic records: evaluations of two exemplar trials. Health Technol Assess 2014;18:1-146.
3. Fiore LD, Brophy M, Ferguson RE, et al. A point-of-care clinical trial comparing insulin administered using a sliding scale versus a weight-based regimen. Clin Trials 2011;8:183-95.
4. D'Avolio L, Ferguson R, Goryachev S, et al. Implementation of the Department of Veterans Affairs' first point-of-care clinical trial. J Am Med Inform Assoc 2012;19(e1):e170-6.
5. Mosis G, Dieleman JP, Stricker BC, van der Lei J, Sturkenboom MC. A randomized database study in general practice yielded quality data but patient recruitment in routine consultation was not practical. J Clin Epidemiol 2006;59:497-502.
6. Weir CR, Butler J, Thraen I, et al. Veterans Healthcare Administration providers' attitudes and perceptions regarding pragmatic trials embedded at the point of care. Clin Trials 2014;11:292-9.
7. Faden R, Kass N, Whicher D, Stewart W, Tunis S. Ethics and informed consent for comparative effectiveness research with prospective electronic clinical data. Med Care 2013;51(Suppl 3):S53-7.
8. Identifying and recruiting participants for health research: a public dialogue for the Health Research Authority. London: OPM Group, July 2015 (http://www.opm.co.uk/wp-content/uploads/2015/07/hra-sciencewise-dialogue-report-final-july-2015-recruiting-participants.pdf).
9. Cho MK, Magnus D, Constantine M, et al. Attitudes toward risk and informed consent for research on medical practices: a cross-sectional survey. Ann Intern Med 2015;162:690-6.
10. Winhusen T, Wilder C, Wexelblatt SL, et al. Design considerations for point-of-care clinical trials comparing methadone and buprenorphine treatment for opioid dependence in pregnancy and for neonatal abstinence syndrome. Contemp Clin Trials 2014;39:158-65.