



## **“Road Testing” Electronic Health Records for Effectiveness Research**

*June 4, 2008  
Washington, DC*

**Conference Report**

---

## Table of Contents

Session Title	Speaker(s)/Panelists	Page
Key Themes		3
Effectiveness Research Using EHRs: Gold Mine or Tower of Babel?	<b>Paul Tang, MD</b> <i>Palo Alto Medical Foundation</i>	5
Case Study: Geisinger Health System – "Beyond RCT"	<b>Walter "Buzz" Stewart, PhD</b> <i>Geisinger Health System</i>	7
Case Study: Veterans Administration	<b>Seth Eisen, MD</b> <i>Veterans Administration</i>	9
Promise and Pitfalls of Outcomes Research Using EHR Databases	<b>Richard Tannen, MD</b> <i>University of Pennsylvania</i>	11
Using EHRs in Decision-Making – Focusing on the "When" and "How"	<b>Stuart Altman, PhD (Moderator)</b> <i>Brandeis University</i>  <b>Philip Burstein, MD</b> <i>GlaxoSmithKline</i>  <b>Steve Pearson, MD</b> <i>Institute for Clinical and Economic Review</i>  <b>Jean Slutsky</b> <i>Agency for Healthcare Research and Quality</i>	13

*The Health Industry Forum is based at Brandeis University, chaired by Professor Stuart Altman, and directed by Robert Mechanic. The Forum brings together public policy experts and senior executives from leading healthcare organizations to address challenging health policy issues. The Forum conducts independent, objective policy analysis and provides neutral venues where stakeholders work together to develop practical, actionable strategies to improve the quality and value of the US healthcare system.*

*Conference presentations and other background materials are available at [www.healthindustryforum.org](http://www.healthindustryforum.org)*

**Health Industry Forum ♦ Heller School for Social Policy and Management ♦ Brandeis University**  
 415 South Street, MS035, Waltham, MA 02454 ♦ (781) 736-8479 (tel) ♦ (781) 736-3306 (fax)

## Key Themes

### Overview

Part of the promise of electronic health records (EHRs) has been the perceived ability to aggregate and analyze health records for clinical and effectiveness research, helping to determine which products and interventions generate better health outcomes for patients. While interest in using EHRs for research remains high, the potential is far from realized; widespread EHR use in research will require overcoming significant barriers including data inconsistency, provider buy in, and prioritization of research.

Overcoming these barriers starts with prioritizing intelligent data collection: addressing data limitations, capturing additional necessary data, and allowing health information from disparate systems to be aggregated for analysis. It also means working with providers to attain clinician buy-in and with EHR vendors to create unobtrusive data collection methods. Lastly, it requires developing research methods for the analysis of observational data that are credible to payers, manufacturers, clinicians, and academic researchers.

If these issues can be resolved—and many Forum participants are optimistic that they can be—EHRs hold the potential to significantly impact how research is currently conducted and how care is ultimately delivered.

### Context

Healthcare systems that have adopted EHRs are quickly amassing a database of valuable information on clinical treatments and patient outcomes. Although EHRs are primarily designed to assist clinicians in treating individual patients, health services researchers and policymakers are interested in aggregating these records to study the effects of treatments on patient populations.

On June 4, 2008, the Health Industry Forum's Evidence Workgroup brought together leading researchers, policy experts, and experts on electronic health records to examine when and under what conditions EHRs can produce data that are suitable for use in clinical and effectiveness research.

### Key Takeaways

- **Today EHRs are being used primarily for clinical care purposes.**

Those advocating EHRs often extol multiple potential benefits. The primary use of EHRs today is for clinical care purposes, although substantial efforts are underway to use EHR data for quality and safety reporting. Although the ability to use the data collected in EHRs for clinical and effectiveness research is exciting in concept, use of EHR data for research remains largely an afterthought.

- **HIF Forum participants believe that EHRs hold much promise as research tools.**

Although randomized controlled trials are still deemed the gold standard of evidence, several participants commented that the research community must evolve beyond an 'either/or' mindset that dismisses observational research as less credible. The reality is that different research methods and designs can be effective at answering different research questions.

EHRs have tremendous potential as a research tool, providing real-world health outcome data across a spectrum of patients, diseases, and clinical settings. These records offer large sample sizes of detailed, longitudinal, patient-level data specifying diagnoses, procedures, diagnostic information, and medications.

- **But realizing EHRs' research potential won't be easy due to a host of barriers.**

While many researchers and EHR experts are optimistic about the role that EHRs can play in clinical and effectiveness research, everyone acknowledged several significant barriers, including:

- *Making research a priority.* Today clinical care is the priority for users of EHR, and research is an afterthought. Gathering the data needed for research purposes must be viewed as a necessary component to enhance clinical care.
- *Data issues.* Currently data is highly fragmented and is often incomplete. Standards are lacking for language and terminology, and different EHR systems often don't contain the same data elements. In addition, these systems are often not interoperable, making it difficult or impossible to aggregate data from different providers and users of different systems.
- *Provider buy-in.* While many providers recognize the dearth in medical evidence, and see potential for EHRs to help address this gap, they will be resistant to gathering data if doing so obstructs their normal clinical workflow.

- **Potential solutions for overcoming these barriers require early and targeted policy efforts.**

As the implementation of EHRs increases, policy efforts can improve the feasibility of EHR-based clinical and effectiveness research, including:

- *Making research a priority from the outset.* Researchers need to be in contact with users of EHRs and with EHR vendors to establish research as a priority, and to define the questions they want to answer. Knowing the necessary variables upfront can allow EHR vendors to create standard and unobtrusive data extraction modules.
- *Securing funds from comparative effectiveness research.* If a national commitment to comparative effectiveness research materializes in the near future, some of these funds should be

invested in innovative research methods for analyzing observational data through EHRs.

- *Developing data standards.* While not easy, part of the data problem could be addressed by developing consistent language and terminology for use in EHRs. The best approach may be for clinical specialty societies to prioritize research questions and data elements.
- *Creating a select network of research-focused EHR users.* Large provider systems may contain enough health care information to conduct valid effectiveness research based solely on their own population. Instead of trying to convince all providers to collect data suitable for research, a more practical idea may be to create a network of EHR users who agree

to collect expanded data for research purposes and who might be paid for doing so.

- *Aligning with the quality movement.* Many of these steps are consistent with other goals like quality reporting and safety monitoring using EHR records. Just as national committees have formed to develop data standards and quality metrics, similar efforts are needed if EHR data are to be used for research. The quality movement makes a potentially compelling business case for using EHRs to measure and reward providers based on their results. It may be possible for the research community to piggyback on these quality efforts to ensure that the right data are being collected and aggregated.

# Effectiveness Research Using EHRs: Gold Mine or Tower of Babel?

Presenter: **Paul Tang, MD**, Palo Alto Medical Foundation

## Overview

Reusing data captured in EHRs for purposes other than care delivery (such as clinical and effectiveness research) is an appealing idea. Dr. Tang believes this vision can be achieved, but major impediments must be overcome: lack of data standards, paucity of fields for capturing relevant data, and challenges in getting providers to enter needed information.

Tapping into the potential of EHRs for research requires a change in mindset. Instead of mining data for research as an afterthought, it must be a priority that is thought of proactively. The systems and processes must be created to capture standardized data as part of providers' normal workflow. Ideally, there will be a harmonization between the data being captured for clinical care with data needed for research.

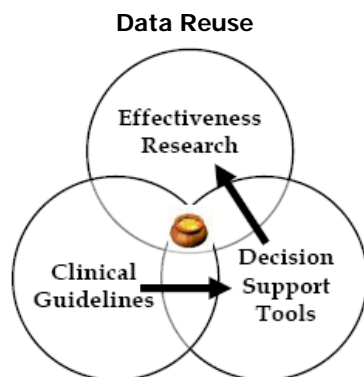
## Context

Dr. Tang assessed the issues surrounding the reuse of EHR data for conducting effectiveness research.

## Key Takeaways

- **In an ideal world, clinical data from electronic health records would be used for multiple purposes.**

Dr. Tang shared a vision where clinical data in EHRs would be used to shape real-time patient-care decisions. As a by-product of delivering care, this data would be reused for purposes including measuring quality performance, providing physicians with performance feedback, supporting public health reporting, and facilitating clinical and effectiveness research.



*"Use of EHRs provides significant opportunity to efficiently conduct effectiveness research."*

— Paul Tang, MD

- **Experience using EHR systems to measure quality shows the potential of EHR data as a research tool.**

The health care quality movement currently employs metrics and benchmarks derived from administrative claims data for use within new reimbursement policies and pay-for-performance programs. EHR systems are showing that they can generate superior quality measures and could provide essential clinical data upon which to base these quality measures. EHR data can more accurately identify patients in the target population and are less subject to "gaming" than are claims data.

Dr. Tang presented a retrospective analysis of the care for diabetic patients; electronic records correctly identified virtually all (98%) of the clinical cases, while claims data missed nearly a quarter of them and provided an inaccurate interpretation on the quality of care this population had actually received.

Dr. Tang suggested that policy makers ought to change pay-for-performance incentive measures from administrative data to clinical data from EHR, and develop a transition plan to migrate the use of administratively-based quality measures to clinically-based quality measures.

- **While the potential in using EHRs for research purposes is great, many obstacles exist.**

Dr. Tang is optimistic about the role that EHRs can play in aiding clinical and effectiveness research, viewing the glass as "half full." However, he acknowledges many practical challenges. These include:

- *Fragmentation of data.* Data is currently all over the place; getting all of a patient's (and a population's) data in one central repository represents a tremendous undertaking.
- *Lack of interoperability.* Currently there aren't standards for aggregating data across disparate electronic health record systems. Not being able to aggregate data hinders the ability to pool EHR data for research purposes.
- *Accuracy of the data.* It is necessary to know/confirm that an entry in the EHR came from an authoritative data source and that the data is accurate. This concern is exacerbated with the growth of personal health records (PHRs); patients entering information into their own file might be more likely to include inaccurate information.
- *Availability of EHRs.* A barrier is whether there is enough critical mass of EHR use, as currently only about 10% of providers use an EHR.
- *Availability of necessary data elements.* A key challenge is whether an EHR has a specific field or data element where a certain type of data can be entered. For example, is a patient's blood pressure recorded for every visit? Are pharma-

ceutical side effects systematically collected? Every EHR system is different; some have data elements that others lack.

- *Lack of standard definitions for clinical variables.* In many instances there are not standard data elements for consistent coding. For example, no ICD9 code exists for the diagnosis of "persistent" asthma. Different providers may use different terminology in entering the diagnosis, which creates a lack of consistency in the data and makes it more difficult to use the data for research.
- *Fit with providers' workflow.* A major barrier exists if providers (especially physicians) are expected to enter data into an EHR outside of their normal workflow. If it is difficult to get data into the EHR, the likelihood of capturing the necessary data decreases.
- *Auditability of data.* Processes must be in place so that the data in an EHR can be tracked over time to ensure accuracy.

▪ **To tap into the potential of EHRs for research purposes, a new approach is required.**

The traditional approach to deriving evidence from EHRs has involved asking:

- *What data are available?*
- *Are they standardized and combinable?*
- *Based on these data, what important effectiveness questions can be answered?*

This approach entails looking at the data that exists and then trying to figure out how to make use of it. Dr. Tang characterized this approach as analogous to "a hammer looking for a nail."

A more effective approach is to start not with the existing data but with the questions to be answered, and then work backwards to arrive at the data needed. The appropriate questions that stakeholders should ask:

- *What are the high-priority research questions to be answered?*
- *What critical data are needed to answer these questions?*
- *Are these data standardized and combinable?*
- *Can they exist within EHRs?*

---

*"We might have to backtrack our way out of the question . . . we need to understand how to use EHRs to get the most bang for the buck."*

— Paul Tang, MD

---

This approach represents a new way of thinking about using EHRs to collect data for research. It requires making the reuse of data for research purposes a priority, not an afterthought; it requires being proactive, not reactive.

---

*"We tend to think of research as an afterthought; we really need to [be thinking about it] upfront."*

— Paul Tang, MD

---

Making research a priority will require cultural and political changes among providers, along with redesigned systems. Specifically:

- EHR systems will need to be redesigned to capture relevant research data that are both reusable and useful to clinicians.
- Key data will need to be captured using standardized definitions and codes that facilitate data aggregation. This will mean advocating for better standards (e.g., ICD 10, SNOMED).
- The buy-in of care providers and their staffs will be required, which will happen only if they can see tangible benefits for changing the ways in which they currently work.
- Physicians will have to be shown how reusing data to develop decision support tools can streamline their workloads, automating some processes, as well as improve care.
- Data entry should fit into the standard clinical workflow and not be overly arduous or interruptive. Data must be entered only once, by the right professional.
- Data will need to be made more structured, with neither too little structure nor too much.
- Critical data needs will have to be harmonized with the needs of clinical care and quality measurement.

---

*"Once we build the right system that can lock down [the right] information, we'll have a very rich environment."*

— Paul Tang, MD

---

## Suggested Steps

- Prioritize clinically important problems.
- Leverage normal care delivery workflow when capturing data.
- Give providers a reason to capture data for research purposes. Make these data reusable and then utilize it.
- Ensure that critical data elements are on the standards development roadmap.
- Work with EHR vendors to ensure that their systems capture clinical trials data as part of their routine user workflow.

# Case Study: Geisinger Health System – "Beyond Randomized Controlled Trials"

Presenter: **Walter "Buzz" Stewart, PhD**, Geisinger Health System

## Overview

Huge gaps exist in the body of medical knowledge. The collection of data through EHRs has the potential to address some of these gaps, and provide valuable real-world information about how well specific treatments work for specific patients.

Beyond developing better knowledge from EHRs, it is necessary to marry EHRs with decision support tools that enable providers to put this knowledge into day-to-day use in their practices.

## Context

Dr. Stewart described multiple ways that EHRs can be used to improve the evidence available to health care providers.

## Key Takeaways

- **EHRs and other tools have tremendous promise to overcome the knowledge gaps.**

A substantial portion of how medicine is delivered is derived from industry-funded randomized controlled trials (RCTs), particularly as it involves pharmaceuticals and other medical technologies. Yet most RCTs result in general claims for a broadly defined intervention class (such as: statins reduce LDL) and are confined to a single clinical domain. They lack evidence around comparative effectiveness, and what works best for which patients, especially patients with multiple comorbidities.

Electronic health records and their associated analytical tools and care processes hold promise for developing a base of medical evidence that extends beyond what is currently available through RCTs. The value of EHRs is that they contain real-world data. Great benefit will be derived from linking EHR data with claims data.

- **The value of retrospective analysis of EHR data will lie in the context of the research question.**

There are many potential uses for the retrospective analysis of data from electronic medical records. These include: comparative assessment of treatments (determining what works best for whom), new treatment indications, and assessing the relative gain in outcomes per treatment intervention.

In particular, EHRs can provide valuable data when lab and clinical measures are the primary outcomes being measured. (Examples include LDL, serum glucose, and blood pressure.) This is because the data is consistent even when different labs are involved.

In contrast, the use of EHRs for retrospective analysis is weaker when the primary endpoint relies on patient-reported outcomes, as the lack of standards in collecting and reporting the data can make them unreliable. In addition, retrospective analysis does not work well when side effects are important in optimizing treatment outcomes, such as with depression, migraines, or prostate surgery.

Regardless of the endpoints, the key limitations are issues around data quality, completeness, and specificity. Perhaps most important are concerns about using observational data collected through EHRs for making causal inferences about treatment effectiveness. For example, studying bariatric surgery outcomes, researchers cannot be completely confident that the population of surgical patients is identical to a population of medically-managed obese patients. Without randomization into treatment and control groups, such analyses may suffer from unmeasured confounding or selection bias.

- **EHR systems serve as the framework for extended data collection, as well as a more effective method of putting knowledge in practice.**

Even when EHRs are used and knowledge exists, there still is high variation in clinical decisions and outcomes. Among the reasons for this are issues related to knowledge retrieval and interpretation, as well as differences in patient preferences, motivation, and education. Thus, having knowledge is important, but even more important is how that knowledge is put to use.

Geisinger is extremely focused on how knowledge is put to use. Over the past three years they have developed a set of tools termed "EHR extenders." These tools are independent of, but interact with, EHRs.

---

*"The EHR is the starting point, but there may be a need for other tools that sit outside of it and interact with it."*

— Walter Stewart, PhD

---

- **Patient data capture tools.** To capture individual preferences and personal outcomes, patients at Geisinger are starting to input data directly in their EHR in a variety of ways. Data can be collected by mail, via computer at the patient's home, or by using touch screens in the waiting area or exam room (the exam room has been particularly effective). Some of the technologies being tested include digital pens, pentabs, touch screens, and data incorporated directly from home monitoring.

While Geisinger believes that additional data from patients will improve the care provided, they also recognize the



benefit of saving staff time by shifting work to the patient. The challenges include deciding what should be collected, where it should be collected, how it should be collected, and authenticating the patient.

- *Decision support tools.* Geisinger is developing real-time dashboards and decision support tools to reduce variation, cut the average time to input data, and improve provider efficiency. A dashboard incorporates data from EHRs as well as patient preference data, and provides an intuitive visual display. A survey of rheumatologists indicated that they need on average 15 minutes to review a patient's data; however, in reality they have just 2 to 3 minutes for this task. The hope is that a custom dashboard that pulls the key data from a patient's EHR can help these physicians get the information they need much quicker.

The main challenges in developing these decision support tools include getting all of the data into usable form (much of the information in EHRs is in unstructured text or even PDFs), ensuring that the data is consistently high quality, and is complete.

---

## Other Important Points

- **Long-term process reengineering.** Use of an EHR could significantly change how patients interact with the health care system. For example, a clinic might change the normal patient visit by having patients, upon arrival, enter information using touch-screen computers. Nurses whose sole responsibility is data gathering, could then gather and upload information into the patient's EHR. This would result in more information available to improve the quality of the patient/physician interaction as well as improving efficiency. Such changes could mean more technology in clinics or physicians' offices, and might increase the ratio of nurses to physicians.



## Case Study: Veterans Health Administration

Presenter: **Seth Eisen, MD**, Veterans Health Administration

### Overview

The Veterans Health Administration (VHA) has a wealth of data in its electronic health records (EHR), data that is available and used extensively by epidemiologists and researchers in the VHA's Office of Research & Development (ORD).

However, the fully productive use of this EHR data for research is hampered by data access hurdles and other barriers to the free flow of medical records. Several initiatives to centralize the VHA's data, provide centralized analysis capabilities, mine textual data, and centralize IRB review will help VHA researchers realize even greater value from the VHA's EHR data asset.

As a case study, the experiences within the VHA document both the possibilities and challenges of trying to gather EHRs across multiple hospitals/clinics and making them accessible for investigators.

### Context

Dr. Eisen explained the advantages and challenges faced by researchers at the VHA's ORD. He described initiatives underway to improve research process efficiency and efforts to leverage the immense amounts of data collected by the VHA.

### Key Takeaways

- **The VHA has extensive electronically-based data available for medical research.**

Created by legislative mandate, the VHA's ORD studies medical issues of concern to America's veterans. The ORD's \$1.8 billion fiscal 2008 research budget is funded through both federal and non-federal sources. Research objectives must be "veteran centric," but the findings may be applicable to the broader population.

The VHA implemented a state-of-the-art electronic medical record system which now has over 10 years of data. It holds a vast amount of patient information on the 5.3 million vets served annually—including progress notes, doctors' orders, images, vital signs, and medication information. In just the last year, billions of new records were added.

As a result, VA researchers have access to extensive amounts of data that is available for research purposes. Further, its integrated organizational structure facilitates the system-wide implementation of findings from research.

---

*"The VHA is for researchers who love data."*  
 — Seth Eisen, M.D., MSc

---

- **The quantity of data for VHA researchers is great, but barriers in using this data affect research effectiveness.**

Although VHA researchers have access to extensive amounts of electronic data, numerous barriers inhibit the use of these data for research purposes. These barriers include:

- Data security and privacy safeguards.
- Relevant data commonly resides in multiple locations.
- Extensive restrictions often govern the use of data.
- Multicenter IRB reviews are often required (different IRBs at different sites), a process that can take many months.
- VHA databases often do not integrate data from non-VHA sources (e.g., Medicare, DoD).
- Sharing of VHA data with external entities is tightly regulated.
- Substantial health information resides in inaccessible text format — such as providers' progress notes, radiology and pathology reports, and hospital discharge summaries.

- **Three new initiatives are designed to lower the barriers that VHA researchers face in accessing data for research.**

The ORD is undertaking three initiatives to more fully realize the research potential inherent in its wealth of data. These are:

- *The Center for Scientific Computing (CSC) Research Initiative.* The CSC will centralize the VHA's data in one place, allowing for real-time mirroring of clinical data and better clinical decision support. Additional computing power will allow for complex modeling, rapid sharing and dissemination of information, and improved research progress. Further, data security and privacy will be enhanced through a remote access structure.

---

*"Ultimately, all of the VHA's clinical data will be available..."*

— Seth Eisen, MD, MSc

---

- *The Medical Informatics Research Initiative.* To address the large amount of information within clinical notes, a virtual informatics research consortium will work to develop methodologies for preparing text for data mining. Data from textual sources will be extracted, cleansed, and reformatted. Text processing will be applied to several clinical issues, and non-consortium investigators will be encouraged to develop informatics research projects using text data.

- *Central Institutional Review Board (IRB).* This planned, centralized IRB would expedite the review process, with significant time savings over the current multi-site review process. This centralized IRB would eliminate potential local conflict of interest, and would facilitate consistent ethical and scientific review.

---

## Other Important Points

- **The VHA as a national resource.** The current focus of VHA research is solely on the country's veterans. Some Forum participants felt that—given the VHA's delivery system, its EHR, and its research capabilities—a broader objective of producing research with application to the wider population would be more useful. But currently the VHA's aspirations, funding, and policies don't rise to this level.
- **Differentiating characteristics.** Several participants pointed out that the VHA patient population is sicker, poorer, and has more complex and severe medical conditions than the general population. VHA patients more closely resemble the Medicaid population.
- **Incomplete treatment data.** Although the VHA is an integrated provider, many veterans receive some of their care from non-VHA providers, like those associated with Medicare. Thus, all pertinent health and treatment information may not be recorded in the EHR, a problem that is endemic with provider-based EHR systems in this country.

# Promise and Pitfalls of Outcomes Research Using EHR Databases

Presenter: **Richard Tannen, MD**, University of Pennsylvania

## Overview

Research conducted by Dr. Tannen indicates that EHR databases can yield comparable results to RCTs. His findings support the potential value of clinical databases to investigate treatment effectiveness.

Dr. Tannen discussed his unique study design strategy and development of the "Prior Event Rate Ratio" (PERR), a statistical adjustment method that is integral to addressing unmeasured confounding.

However, this research also identified challenges and potential pitfalls: the existence in observational studies of hidden bias, especially from confounding by indication, and validity and data limitations of the information in the EHR database.

## Context

Dr. Tannen described his research examining whether studies using an EMR database would yield valid outcomes in comparison to existing RCT results.

## Background

Dr. Tannen and a team of colleagues initiated a research program to determine whether data within an EHR database could be used to retrospectively simulate the results of a randomized controlled trial. By replicating all aspects of an RCT (selection criteria, study time frame, treatment, and outcomes), Dr. Tannen acknowledged two major hurdles to using EHR data for effectiveness research: 1) unrecognized confounding in observational studies; and 2) the validity of the data within the EHR database. Consequently, he tested whether the observational, non-random nature of the patient data would generate comparable results.

Dr. Tannen and his team selected patients from the United Kingdom's General Practice Research Database (GPRD), a primary care centric EHR database that contains medical records from approximately 700 General Practices. The database contains all treatment and prescription information for approximately 5.5% of the UK's population, representing 3.5 million active patients and 25 million patient years of experience.

Using the electronic medical records from GPRD, Dr. Tannen's team proceeded to model these data for comparison with well-known randomized controlled trials (for example, the Women's Health Initiative). In doing so, the same inclusion and exclusion criteria were used; the same outcomes were used; and the same confounders and control variables were examined. The key difference was the assignment of patients into the intervention and control groups; whereas patients in the RCT were assigned randomly, patients and

physicians in the GPRD dataset chose whether an individual received treatment or not. If clinicians preferentially steered their healthier or more proactive (or some other characteristic) patients toward therapy, then any difference in health outcomes could be attributed to these unmeasured variables.

Along with standard statistical methods to control for potential bias due to the non-random assignment of treatment and control groups (propensity score analysis, covariate analysis, logistic regression, and adjusted incidence rate ratios), Dr. Tannen and his team adjusted for unmeasured confounding with the "Prior Event Rate Ratio" (PERR), described below.

## Key Takeaways

- **Observational studies using EHR data can yield comparable results to RCTs but sometimes suffer from unobservable bias.**

In general, this research showed that a simulated RCT using observational data from an electronic primary care practice database can reasonably replicate the findings on the RCT. In each of the six RCTs he replicated, Dr. Tannen showed that it was possible to achieve relatively comparable results between an EHR study and a RCT.

---

*"I think there is evidence that EHR databases can potentially have enormous power to answer research questions."*

— Richard Tannen, MD

---

When no confounding was present, the RCT outcomes were comparable to the GPRD outcomes. However, in one of the six studies, the results were not comparable. Dr. Tannen and his team concluded that there must have been an unmeasured confounding present. To adjust for this, they developed an adjustment termed the "Prior Event Rate Ratio" (PERR). The PERR creates a ratio of the events among the exposed group and the unexposed group prior to the start of the study and during the study, akin to a "differences-in-differences" approach utilized in health services research. Use of this ratio adjusts for confounding and yields comparable results between the RCT and the GPRD.

Dr. Tannen's group has published or submitted for publication papers on all six GPRD/RCT comparisons, which includes explanation of the development of PERR. It is their hope that other researchers agree that the overall findings and the development of PERR are valid.

- **Use of EHR databases could transform research.**

The validity of studies using observational data for assessing treatment effectiveness remains an issue of considerable

controversy. Furthermore, using electronic medical records is also controversial, due to the concern about the reliability of the data. While it is not possible to precisely replicate a RCT from clinical practice data, Tannen's studies show that, despite lack of randomization, sufficient similarity can be achieved to yield relatively comparable results between a GPRD study and a RCT. In addition, the EHR data would be able to be analyzed faster and less expensively than a RCT, and also would reflect "real-world" situations as opposed to more controlled RCTs.

Dr. Tannen believes that there are important limitations to this study and to the overall study design strategy. While he was able to demonstrate the feasibility of this specific analytic approach, it may not be possible to generalize these results to different and other studies. Furthermore, the results using the GRPD cannot be extrapolated to other databases, and different study characteristics, (different confounders, linkages to death

certificates, the study start time, and so on) will influence the results in other studies. Nevertheless, despite these potential shortcomings, the simulated RCT largely replicated the findings of the RCT.

---

## Other Important Point

- **50 Million Records.**

While the cohort extracted from the GRPD database was adequate for his research purposes, and contained 3.5 million patients, Dr. Tannen believes that a database with 50 million patients would be ideal. It would increase the chances of having enough patients in various exposed groups, and would allow for analysis of sub-groups and segments.

## Using EHRs in Decision-Making—Focusing on the “When” and “How”

Moderator: **Stuart Altman, PhD**, Brandeis University

Panelists: **Philip Burstein, MD**, GlaxoSmithKline  
**Steve Pearson, MD**, Institute for Clinical and Economic Review  
**Jean Slutsky**, Agency for Healthcare Research and Quality

### Overview

Participants are enthusiastic about the potential of using EHR data for clinical and effectiveness research. They believe this provides a new form of observational data that can be used in numerous ways by multiple stakeholders. EHR data may be able to provide insights that other forms of data cannot, and may be able to help answer questions better, faster, and less expensively than other research methods.

To make the use of EHR data for research a reality, early and targeted policy efforts are needed, including: agreement on the language used in data collection, investment in new research methods, and engagement of information technology vendors in designing EHRs for research purposes.

### Context

The panelists shared their perspectives on the role that EHRs can play in clinical and effectiveness research; Forum attendees then offered comments on the key issues to be addressed.

### Key Takeaways (Slutsky)

Ms. Slutsky shared several thoughts related to EHRs and the approach that the Agency for Healthcare Research and Quality (AHRQ) is taking.

- **Types of research should be viewed along a continuum.**  
Often research is described as either a randomized controlled trial or observational, with RCT seen as the “gold standard.” But this ‘either/or’ mindset is not the right approach. In reality, research methodologies should be examined across a spectrum. Different types of study designs control different variables. The key question that should be asked for any research is: What is the appropriate study design for the objectives pursued?
- **Examining EHRs’ use in research raises a larger question about the need for innovative research methods.**  
For years researchers have taken the same approaches to conducting research. What is needed is innovation in methods. In particular, funding agencies must fund research as well as invest, encourage, and promote new statistical and research tactics.

---

*“Funding agencies need to encourage and support innovative new methods [of research].”*

— Jean Slutsky

---

- **In using EHRs for research, it is important to pay attention to the cohort falling outside of the EHR.**  
There is a chance that the population enrolled in EHRs or registries is not representative. It is important to examine who has been left out. Currently, EHRs are primarily utilized by mostly urban, academically-oriented provider systems, potentially limiting generalizability.
- **It may be necessary for additional data to be collected and stored in EHRs for research purposes.**  
The data currently in EHRs is focused on patient care, as it should be. To make EHR databases more robust for research purposes, collecting additional research-focused information, such as quality-of-life measures, may be desirable.
- **AHRQ is funding pilots that use observational data from EHRs and registries.**  
AHRQ believes in the potential of using EHR data from clinical practices for research, and currently funds the Distributed Network for Ambulatory Research in Therapeutics (DARTNet), aggregating EHR data from a network of 200 primary care-based research practices. These practices collect data for research purposes, and receive physician benchmarking reports in return (an attractive proposition for practices). Similarly, AHRQ sponsors the HMO research network, a consortium of 15 HMO organizations that utilize administrative data, electronic medical records, and disease-specific registries to conduct population-based research to answer key clinical and health policy questions.  
  
A conclusion from these pilots: it's important to engage EHR vendors such as EPIC and Cerner in creating EHRs that work well for research purposes.

### Key Takeaways (Burstein)

Dr. Burstein described the potential value that pharmaceutical companies such as GlaxoSmithKline (GSK) see in EHRs.

- **For pharmaceutical companies, EHRs present an opportunity to better understand patient populations.**  
Pharmaceutical companies hope to leverage EHR data to better understand their targeted patient populations, which will

improve the effectiveness of drug discovery and development processes as well as improve and expedite clinical studies.

Already GSK pulls huge quantities of information from multiple sources into a "health care information factory," designed to help the company understand what is happening with patients in general, and those receiving GSK treatments in particular. Adding EHR data will enrich the data set and increase its value.

---

*"There is a huge opportunity for the pharmaceutical industry to leverage EHR data to better understand the population of patients, which will be beneficial in drug discovery and development."*

— Philip Burstein, MD

---

- **Pharmaceutical companies do not need special research data to be collected.**

From GSK's perspective, no additional information need be collected for research purposes; just getting access to the actual clinical data entered by clinicians holds tremendous value. Ideally these data would be searchable.

- **Pharmaceutical companies would also like to review study protocols.**

As key stakeholders in the outcome of EHR-driven research studies, pharmaceutical companies are interested in seeing and collaborating in study protocols in advance of its commencement.

## Key Takeaways (Pearson)

Dr. Pearson agreed with Ms. Slutsky's comments about the need for a continuum of research types. He said that many people view research as a dichotomy, being either RCT or observational. In his view this dichotomy isn't accurate; there are lots of creative and experimental research options. However, he is concerned that some people wishfully believe that EHR-related research is the only research that will ever be needed.

- **There needs to be a consensus on terminology.**

As Dr. Tang illustrated in his presentation, there is lack of consistency on many definitions, such as "persistent" asthma. And no entity is currently working to create consensus on terminology; but for research purposes, such consensus is necessary.

- **An opportunity exists to incorporate "research templates" into EHRs for particular research areas.**

The creation of research templates would allow for more consistent collection of specified data. The templates could differ for various specialty areas. Here, the role of technology vendors will be critical; as groups purchase off-the-shelf EHR systems, companies like EPIC can develop standard data extrac-

tion modules to unobtrusively capture data within the normal clinical workflow.

---

*"There is lots of good in EHRs, but today the primary focus is clinical. It is a struggle to use it for research."*

— Steve Pearson, MD

---

- **Forethought needs to be given to data aggregation.**

Many organizations are already working on how EHR data will be aggregated for administrative and claims purposes, as well as for clinical care purposes. But as of yet there has been little or no thinking around how data will be aggregated from EHRs for research purposes. Without research-specific data aggregation, though, leveraging EHRs for research won't take place.

## Key Takeaways (Group Discussion)

- **There is lack of agreement around whether EHRs need to be altered for research purposes.**

The question was raised about whether existing EHRs need to be modified in order to make them valuable tools for research. Some participants believe that for EHR data to be comparable to RCT data, EHRs need to be modified. Dr. Burstein disagreed, however, stating that the pharmaceutical industry doesn't want or need special data collected; the industry just wants to be able to review the clinical data already being collected. Dr. Tannen commented that the UK's GPRD—an EHR database—was not created as a research tool; it was designed purely as a clinical tool, and is now being used for research purposes.

It was noted that if providers are asked to collect additional data for research purposes, it will be a barrier. One suggestion was not to ask all health systems or providers to collect data for research, but to have targeted delivery systems—that would be paid for their efforts—collect research data.

- **Perhaps "quality" is the hook for using EHRs to collect data for research purposes.**

Efforts are also necessary to expand the penetration of EHR systems. Rob Mechanic observed that for providers to adopt an EHR, they require a business case, and currently the most compelling business case appears to be around enhancing clinical care. He wondered whether there might be a compelling case for EHR adoption built around the need to collect and report quality data. Such a case might involve financial incentives. Development of a quality-focused business case might be a hook for, and ultimately benefit, the research community.

- **Research data for specialty areas holds appeal**

Participants see the need for, and value in, customizing the EHRs and data collected for various specialty areas. They perceive that clinical specialty practices and national medical specialty societies will be receptive to this.



- **Ultimately, EHRs could be used to build predictive models on both the population and individual levels.**

Historically, data has been difficult and expensive to collect, but data in EHRs has the potential to be abundant and inexpensive to collect and analyze. A long-term goal is to use the data collected through EHRs to build predictive models, which would be tested and validated using RCTs. Most other sciences use predictive models, but this has not been the case in medicine.

Individual providers and patients will value the ability to identify other patients with profiles similar to those treated, highlighting which treatments are likely to be most effective. Review the (anonymous) history of patients who are "just like me" could improve doctor-patient communications and treatment planning.

- **New methods, and the support of the research community, are needed.**

Participants were in agreement regarding the need to pursue innovative new research methods. A barrier in pursuing new methods is the lack of participation from leading biostatisticians who, at the moment, don't appear terribly engaged or interested. This is a cultural issue among the statistician community that needs to be overcome. (But some participants believe that researchers and statisticians are beginning to become engaged in this area.)

Barbara McNeil believes that the research community could take more initiative in examining the potential value of EHRs. For example, researchers could create a list of 50 or 75 questions where EHR data might be used to make decisions. Then, researchers should go through the process of asking, "If an observational data set was available, would it be believed, would it make a difference, and would it be actionable?"

- **Patient identifier desired.**

One participant raised the issue that a unique patient identifier is needed to be able to track and identify patients.

- **Use a portion of comparative effectiveness funding to develop new research methods and technologies.**

Stuart Altman sees the likelihood at some future point of funding devoted to comparative effectiveness. He suggested that perhaps a portion of such funding should target new research methods and technologies, such as observational research using EHRs.

Participants concurred that EHR data can play a critical role in comparative effectiveness research. As Dr. Pearson pointed out, "It is just a question of which parts of the EHR will be valuable for comparative effectiveness."