

**AMIA
2013 Summit in CRI
CRI Year-In-Review**

Peter J. Embi, MD, MS, FACP, FACMI

Assoc Prof & Vice Chair, Dept of Biomedical Informatics
Associate Professor of Medicine
Chief Research Information Officer
Co-Director, Biomedical Informatics, CCTS
The Ohio State University

San Francisco, California
March 22, 2013

Approach to this presentation

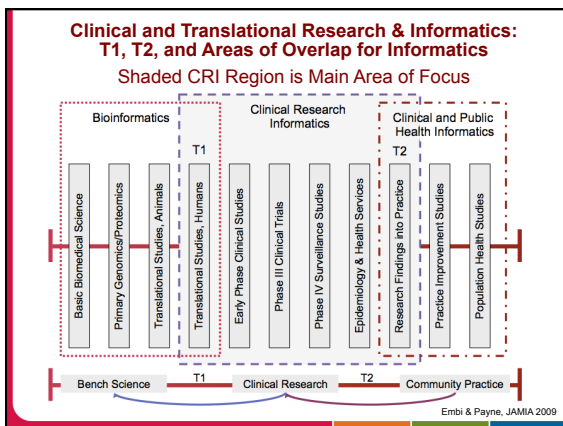
- Mixed approach to article identification:
 - Started with structured approach
 - (akin to ACP "update" sessions)
 - Augment with "what seemed interesting" approach
- Learned a lot from doing this last two years
 - Tracked manuscripts throughout the year
 - (still worked down to the wire)
- So, what was my approach...

Source of Content for Session

- Literature review:
 - Initial search by MESH terms:
 - ("Biomedical Research"[Mesh] NOT "Genetic Research"[Mesh]) NOT "Translational Research"[Mesh] AND "Informatics"[Mesh] AND "2012/01/01"[Pdat] : "2013/02/01"[Pdat]
 - Resulted in 90 articles; **33** were CRI relevant
 - Additional 483 articles found via:
 - Recommendations from colleagues
 - Other keyword searches using terms like:
 - Clinical Trials, Clinical Research, Informatics, Translational, Data Warehouse, Research Registries, Recruitment
 - Yielding **48** more CRI-relevant articles, for a total of...
 - Result = **81** total CRI relevant
 - From those, I've selected **35** representative papers that I'll present here (*briefly*)

Session caveats

- What this is not...
 - A systematic review of the literature
 - An exhaustive review
- What this is...
 - My best attempt at *briefly* covering *some* of the representative CRI literature from the past year
 - A snap-shot of excellent CRI activity over past year
 - What I thought was particularly notable



Topics

- Grouped 35 articles into several CRI categories (admittedly, not *all* CRI areas)
 - Clinical Data Re-Use for Research
 - Data/Knowledge Management & Discovery
 - Researcher Support & Resources
 - Participant Recruitment
 - Policy & Perspectives
- In each category, I'll highlight a few key articles and then given a quick "shout out" to a few others
- Conclude with notable events from the past year

Apologies up front

- I'm CERTAIN I've missed a lot of great work
- I'm REALLY SORRY about that

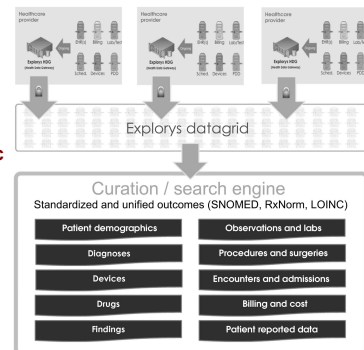
Clinical Data Re-Use for Research

“Patient characteristics associated with venous thromboembolic events: a cohort study using pooled electronic health record data.” (Kaelber DC et al, JAMIA, 2012)

- **Goal:** Demonstrate the potential of de-identified clinical data from multiple systems using different EHRs to be used for large retrospective cohort studies
- **Methods:** Data of 959,030 patients, pooled, standardized and normalized using common ontologies, searchable through a HIPAA-compliant, patient de-identified web application (Explore; Explorys Inc). Patients were 26 years or older seen in multiple healthcare systems from 1999 to 2011 with data from EHR.

“Patient characteristics associated with venous thromboembolic events: a cohort study using pooled electronic health record data.”

(Kaelber DC et al, JAMIA, 2012)



“Patient characteristics associated with venous thromboembolic events: a cohort study using pooled electronic health record data.” (Kaelber DC et al, JAMIA, 2012)

- **Results:** Comparing obese/tall subjects with normal weight/short subjects, the venous thromboembolic events (VTE) OR was 1.83 for women and 1.21 for men. Weight had more effect than height on VTE. Compared with Caucasian, Hispanic/Latino subjects had a much lower risk of VTE (female OR 0.47; male OR 0.24) and African-Americans a substantially higher risk (female OR 1.83; male OR 1.58). This 13-year retrospective study of almost one million patients was performed over approximately 125 h in 11 weeks, part time by the five authors.
- **Conclusion:** Informatics approaches to pool EHR-derived clinical will provide opportunities for CRI to transform the scale and resources needed to perform certain types of clinical research. Potential to achieve results similar to much more costly, time-intensive study approaches.

“QT interval and antidepressant use: a cross sectional study of electronic health records.” (Castro VM, et al, BMJ, 2013)

- **Objective:** To quantify the impact of citalopram and other selective serotonin reuptake inhibitors on corrected QT interval (QTc), a marker of risk for ventricular arrhythmia, in a large and diverse clinical population.
- **Method:** Cross-sectional study using EHR-data (ECGs, Rx, Clinical) to explore relation between antidepressant dose and QTc. 38,397 adult patients with an electrocardiogram recorded after prescription of antidepressant or methadone between February 1990 and August 2011.

“QT interval and antidepressant use: a cross sectional study of electronic health records.” (Castro VM, et al, BMJ, 2013)

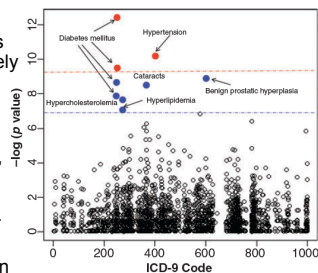
- **Results:** Dose-response association with QTc prolongation was identified for citalopram, escitalopram, and amitriptyline, but not for other antidepressants examined. An association with QTc shortening was identified for bupropion. Within-subject paired observations supported the QTc prolonging effect of citalopram.
- **Conclusions:** Confirmed modest prolongation of QT interval with citalopram, and identified additional antidepressants with similar risk. Pharmacovigilance studies using electronic health record data may be a useful method of identifying potential risk associated with treatments.

“Discovering Medical Conditions Associated with Periodontitis Using Linked Electronic Health Records” (Boland MR. et al, J of Clin Periodontol, 2013)

- **Goal:** To use linked electronic medical and dental records to discover associations between periodontitis and medical conditions independent of a priori hypotheses.
- **Methods:** This case-control study included 2475 patients who underwent dental treatment at the College of Dental Medicine at Columbia University and medical treatment at NewYork-Presbyterian Hospital. Our cases are patients who received periodontal treatment and our controls are patients who received dental maintenance but no periodontal treatment. Chi-square analysis was performed for medical treatment codes and logistic regression was used to adjust for confounders.

“Discovering Medical Conditions Associated with Periodontitis Using Linked Electronic Health Records” (Boland MR. et al, J of Clin Periodontol, 2013)

- **Results:** The method replicated several important periodontitis associations in a largely Hispanic population, including diabetes mellitus type I and II, hypertension, hypercholesterolemia, hyperlipidemia, and conditions related to pregnancy, child-birth.
- Also, previously unreported association with BPH.



“Discovering Medical Conditions Associated with Periodontitis Using Linked Electronic Health Records” (Boland MR. et al, J of Clin Periodontol, 2013)

- **Conclusions:** Describes a high-throughput method for associating periodontitis with systemic diseases using linked electronic records.
- Beyond that – is another good example of using linked electronic health records (and dental records in this case) for disease association studies, and identifying associations that may have been previously unrecognized.

Other notable papers in this category:

- **“A pragmatic framework for single-site and multisite data quality assessment in electronic health record-based clinical research.”** (Kahn MG et al. Medical care. 2012)
- **“Oncoshare: lessons learned from building an integrated multi-institutional database for comparative effectiveness research.”** (Weber S et al. AMIA Annual Symposium proceedings 2012)
- **“Standardizing clinical laboratory data for secondary use.”** (Abhyankar S et al. Journal of biomedical informatics. 2012)
- **Mining electronic health records: towards better research applications and clinical care.** (Jensen PB. Nature Reviews Genetics. 2012)
- **“In Search of a Data-in-Once, Electronic Health Record-Linked, Multicenter Registry— How Far We Have Come and How Far We Still Have to Go.”** (Marsolo K. eGEMs (Generating Evidence & Methods to improve patient outcomes): Vol. 1: Iss. 1, 3.)

Data/Knowledge Management & Discovery

“A survey of informatics platforms that enable distributed comparative effectiveness research using multi-institutional heterogenous clinical data.”
(Sittig et al, Medical Care, 2012)

- **Goal:** To compare and contrast 6 large-scale projects that are either developing or extending existing informatics platforms for CER.
- **Methods:** Rather than compare the informatics platforms at an abstract level, focused specific CER projects that provide implementations of informatics platforms and highlight design requirements and solutions. Utilized an 8-dimension, sociotechnical model of health information technology to help guide work
- **Results:** Identified 6 generic steps that are necessary in any distributed, multi-institutional CER project: data identification, extraction, modeling, aggregation, analysis, and dissemination.

“A survey of informatics platforms that enable distributed comparative effectiveness research using multi-institutional heterogenous clinical data.”
(Sittig et al, Medical Care, 2012)

TABLE 2. Comparison of Data Sources, Types, Models, and Handling of Duplicate Patients

CER Project Name	Data Sources	Data Types	Standard Data Model(s) Used	Duplicate Patients Identified Across Organization(s)
SPAN	Health plan enrollment, pharmacy dispensing, utilization data, billing data, visits, lab results, cancer registry, death info	Local codes Standard codes No standardized text	Expanded version of the IMMO Research Network Variant Data Warehouse Version 3 (13 tables linked by a complex identity)*	No. Two organizations unlikely to have information on the same patient during a defined observation period.
WICER	Patient services, vital statistics, health literacy, socioeconomic status, inpatient, ambulatory clinic, long-term care facilities, home care agencies	Local codes Standard codes Processed free-text	Early version of the HL-7 Reference Information Model	Yes, many patients are participants in New York Care Connect RIE
CER-HUB	Ambulatory EHR, inpatient discharge summaries, billing, pharmacy dispensing, lab results; all as extracted based on project need by standard extraction mechanism	Local codes Standard codes Processed free-text	Implementation of HL-7 Clinical Document Architecture that extends the CCD (Continuity of Care Document)	No. Unlikely for sites currently involved to have overlap in patient populations. One site operates a single-source EHR for its multiple constituent member FQHC organizations.
RFOR	Demographics and lab data loaded nightly; EHR, billing and decision support systems data (including visit and encounter and ambulatory clinic data, death info and pharmacy data loaded monthly); two clinical notes available on provisioned basis	Local codes Standard codes Processed free-text	Star schema data model covers clinical events as “fact” in relational database structure with existing tables that further derive facts, along with metadata tables	Yes, Emorycare Master Patient Index
INPC COMET-AD	Multiple hospital systems, health care systems, practice organizations (eg, primary care group practices, oncology practices, sports medicine practices), laboratory organizations	Local codes Standard codes Processed free-text and structured text	Metadata (ontology, “queries” table, and “reports” table where answer term is numeric, coded, date, person, etc; or clinical, or text value. Also some “reports” table)	Yes, patients are linked across institutions in the database. Network for patient care through the global patient ID service
SCoAP-CERTN	ADT registration, laboratory, medication, test reports (eg, Doppler ultrasound report), text notes (eg, operative note)	Local codes Standard codes Processed free-text Unstructured text	HL-7 v3 in warehouse augmented by data elements within the SCoAP data collection forms	No

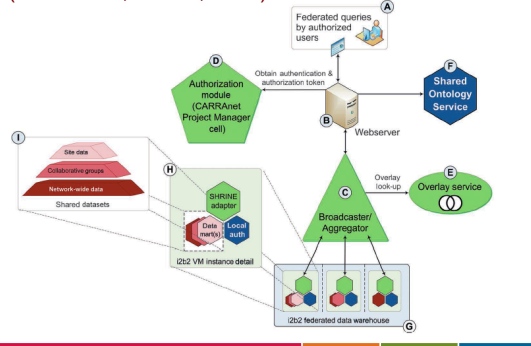
“A survey of informatics platforms that enable distributed comparative effectiveness research using multi-institutional heterogenous clinical data.”
(Sittig et al, Medical Care, 2012)

- **Conclusions:** CER stands to transform the current health care delivery system by identifying which therapies, procedures, preventive tests, and health care processes are most effective from the standpoints of cost, quality, and safety.
- State-of-the-art informatics platforms are necessary to carry out this type of research across organizations with disparate patient populations, health information systems, data types, and local governance structures.
- Such platforms and resources will continue to grow and be tested, essential to advancing CER and related research initiatives.

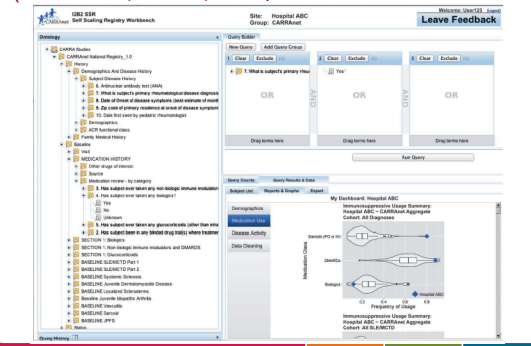
“An i2b2-based, generalizable, open source, self-scaling chronic disease registry.”
(Natter et al, JAMIA, 2013)

- **Goal:** Registries growing in importance for obtaining high-quality, disease-specific data. But, they are often highly project-specific in design, implementation and usage. This project sought to develop a “self-scaling” approach for collaborative data sharing.
- **Methods:** Leveraging i2b2 and SHRINE networking software platforms, created modular, ontology-based, federated infrastructure to give researchers access to data from multiple sites for Childhood Arthritis & Rheumatology Research Alliance (CARRA) Registry

“An i2b2-based, generalizable, open source, self-scaling chronic disease registry.”
(Natter et al, JAMIA, 2013)



“An i2b2-based, generalizable, open source, self-scaling chronic disease registry.”
(Natter et al, JAMIA, 2013)



“An i2b2-based, generalizable, open source, self-scaling chronic disease registry.”
(Natter et al, JAMIA, 2013)

- **Results:** Use of this federated, self-scaling registry platform enabled collaborative data sharing and collaboration while enabling fine-grained control over data sharing.
- The implementation of i2b2-SSR for the multi-site, multi-stakeholder CARRA Registry has established a digital infrastructure for community-driven research data sharing in pediatric rheumatology in the US.
- **Conclusions:** Such data sharing needs across sites are expected to grow, and this is a great example of leveraging current informatics platforms and approaches to accomplish this.

“Validation of a common data model for active safety surveillance research” (Overhage JM. JAMIA. 2012)

- **Problem:** Systematic analysis of observational medical databases for active safety surveillance is hindered by the variation in data models and coding systems. Translating the data from idiosyncratic data models to a common data model (CDM) could facilitate both analysts' understanding and the suitability for large-scale systematic analysis. In addition to facilitating analysis, a suitable CDM has to faithfully represent the source observational database.
- **Goal:** Before beginning to use the Observational Medical Outcomes Partnership (OMOP) CDM and a related dictionary of standardized terminologies for a study of large-scale systematic active safety surveillance, the authors validated the model's suitability for this use by example.

“Validation of a common data model for active safety surveillance research” (Overhage JM. JAMIA. 2012)

- **Approach:** To validate the OMOP CDM, the model was instantiated into a relational database, data from 10 different observational healthcare databases were loaded into separate instances, a comprehensive array of analytic methods that operate on the data model was created, and these methods were executed against the databases to measure performance.
- **Conclusion:** Acceptable representation of data from 10 observational databases in the OMOP CDM using the standardized terminologies selected, and a range of analytic methods developed and executed with sufficient performance to be useful for active safety surveillance.
- Nice example of validating data model for use with observational data - will only become more common

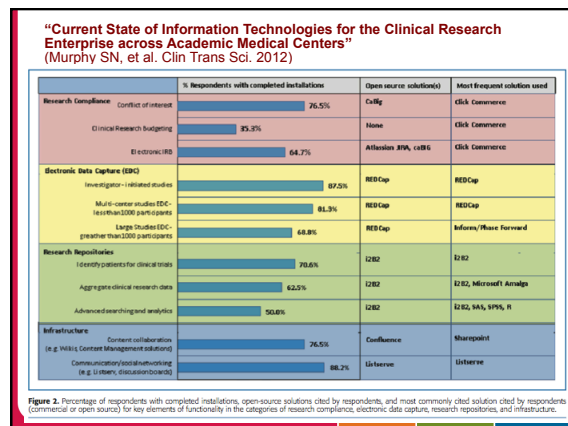
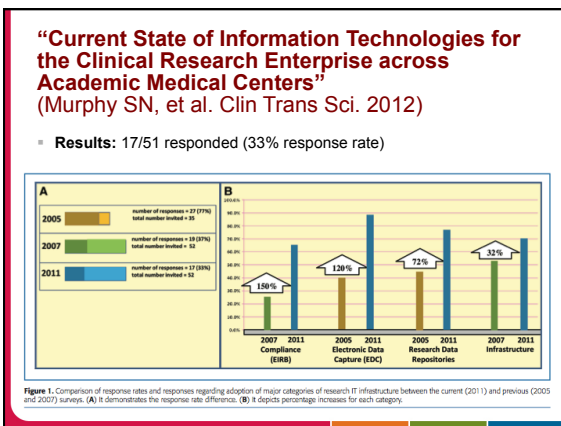
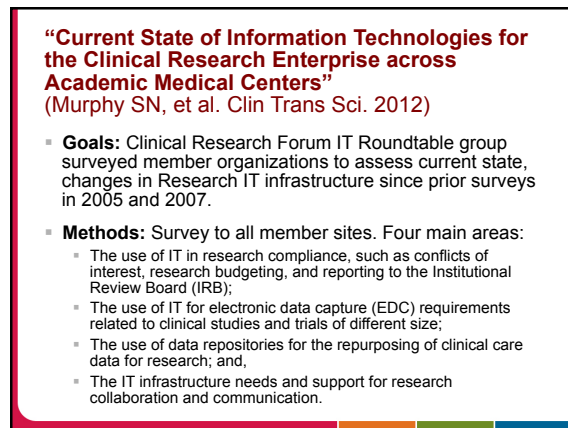
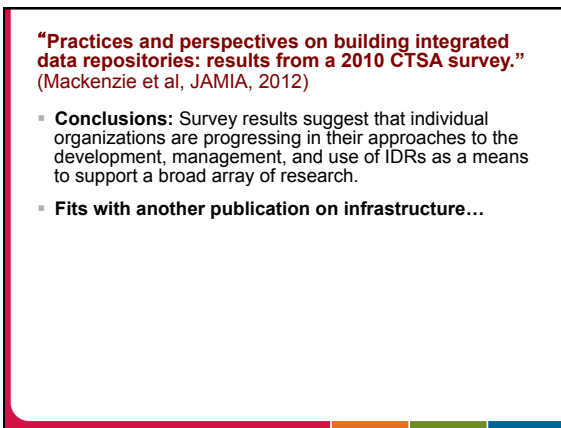
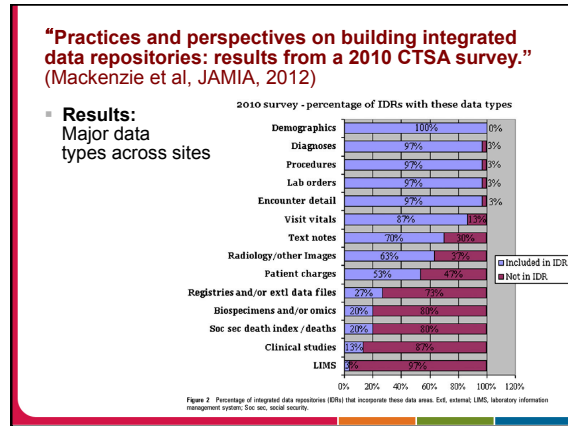
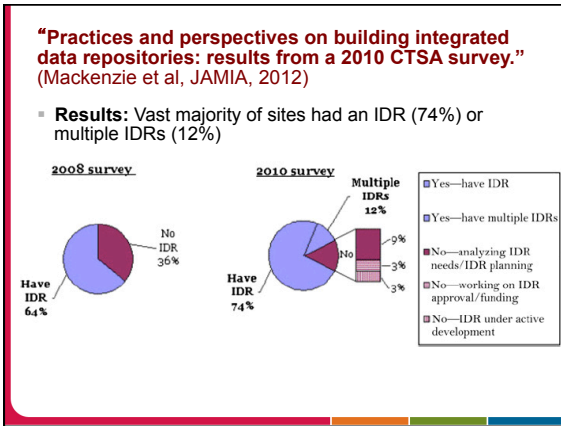
Other notable papers in this category:

- **“Identifying clinical/translational research cohorts: ascertainment via querying an integrated multi-source database.”** (Hurdle JF. JAMIA. 2013)
- **“Data model considerations for clinical effectiveness researchers.”** (Kahn MG. Medical care. 2012)
- **“Applying knowledge-anchored hypothesis discovery methods to advance clinical and translational research: the OAMiner project.”** (Payne PR. JAMIA 2012)
- **“Translating standards into practice: experiences and lessons learned in biomedicine and health care.”** (Chen ES. Journal of biomedical informatics. 2012)

Researcher Support & Resources

“Practices and perspectives on building integrated data repositories: results from a 2010 CTSA survey.”
(Mackenzie et al, JAMIA, 2012)

- **Goals:** Clinical integrated data repositories (IDRs) are poised to become a foundational element of biomedical and translational research by providing the coordinated data sources necessary to conduct retrospective analytic research and to identify and recruit prospective research subjects.
- **Methods:** The CTSA consortium's Informatics IDR Group conducted a Web-based survey of 2010 consortium members to evaluate recent trends in IDR implementation and use to support research between 2008 and 2010.
- **Results:** 74% response rate representing 28 sites and NIH clinical center.



“Current State of Information Technologies for the Clinical Research Enterprise across Academic Medical Centers”

(Murphy SN, et al. Clin Trans Sci. 2012)

- **Conclusions:** Research IS adoption across respondent sites has increased over past 7 years. The availability of more robust and available vendor-based and “open-source” solutions, coupled with new research initiatives (e.g., CTSA) and regulatory requirements, appear to be contributing to these advances.

“Access to core facilities and other research resources provided by the Clinical and Translational Science Awards.” (Rosenblum D. Clin Trans Sci. 2012)

- **Goal:** Review of 60 CTSA website offerings to assess and categorize.
- **Results:** Over 170 generic services, which this review has categorized in the following seven areas: (1) core facilities, (2) biomedical informatics, (3) funding, (4) regulatory knowledge and support, (5) biostatistics, epidemiology, research design, and ethics, (6) participant and clinical interaction resources, and (7) community engagement. In addition, many facilitate access to resources with search engines, navigators, studios, project development teams, collaboration tools, communication systems, and teaching tools.
- **Conclusion:** CTSA significantly impacting awareness of and access to research resources.

Other notable papers in this category:

- **“A model for the electronic support of practice-based research networks.”** (Peterson KA. Annals of family medicine. 2012)
- **“Approaches to facilitate institutional review board approval of multicenter research studies.”** (Marsolo K. Med care. 2012)
- **“Building the informatics infrastructure for comparative effectiveness research (CER): a review of the literature.”** (Lopez MH. Med care. 2012)
- **“Clinical and translational research studios: a multidisciplinary internal support program.”** (Byrne DW. Academic Med 2012)
- **“The feasibility of cell phone based electronic diaries for STI/HIV research”** (Hensel DJ. BMC medical research methodology. 2012.

Recruitment Informatics

“Evaluation of a prototype interactive consent program for pediatric clinical trials: a pilot study.” (Tait AR, et al. JAMIA. 2012)

- **Issue:** Standard written methods of presenting research information may be difficult for many parents and children to understand.
- **Goals:** Examine use of novel prototype interactive consent for describing hypothetical pediatric asthma trial to parents and children
- **Methods:** Parents and children interviewed to examine baseline understanding of key elements of clinical trial, reviewed age-appropriate versions of interactive program describing asthma trial, and tested for understanding.

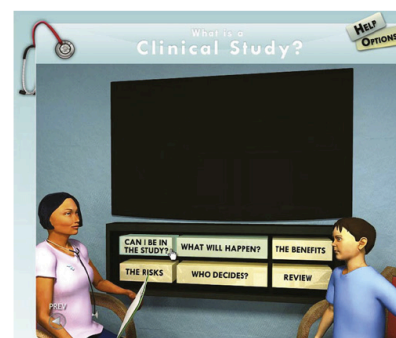


Figure 1 Screen shot of the child's version of the interactive computer consent program.

“Evaluation of a prototype interactive consent program for pediatric clinical trials: a pilot study.”
(Tait AR, et al. JAMIA. 2012)

- **Results:** Parents and children improved understanding of key research concepts following program. E.g. percentage of parents and children who could correctly define the terms clinical trials and placebo improved from 60% to 80%, and 80% to 100% among parents and 25% to 50% and 0% to 50% among children, respectively. Results also suggest that the interactive programs were easy to use, facilitated understanding of the trial among parents and children.
- **Conclusions:** Interactive media may offer an effective means of presenting understandable information to parents and children regarding participation in clinical trials. This will likely grow along with need for better ways of ensuring “informed” consent using novel, automated approaches.

“Challenges in creating an opt-in biobank with a registrar-based consent process and a commercial EHR.”
(Marsolo K. JAMIA, 2012)

- **Goal:** Implement an opt-in biobank, with auto-consent at the time of registration and the decision stored in EHR for data and bio-specimen re-use.
- **Methods:** Implemented at registration and information about decision and related data stored in system. Investigators can search for samples using i2b2 data warehouse.
- **Results:** Patient opt-in rate over 86%, with 83% requesting to be notified of any incidental research findings. In 6 months, obtained decisions from over 18,000 patients and processed 8000 blood samples for storage in our research biobank. Found some limitations in current systems that required work-arounds.

“Challenges in creating an opt-in biobank with a registrar-based consent process and a commercial EHR.”
(Marsolo K. JAMIA, 2012)

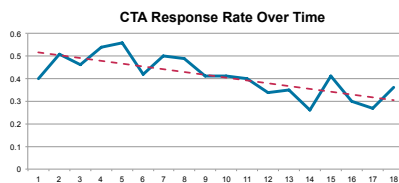
- **Conclusions:** Real-world example with impressive success in obtaining consent/assent from majority of patients requiring limited resources. Lessons that should help others advance these kinds of efforts to re-use data and specimens for research and inform improvements to IT systems to support such efforts.

“Evaluating alert fatigue over time to EHR-based clinical trial alerts: findings from a randomized controlled study.”
(Embi & Leonard, et al. JAMIA, 2012)

- **Goal:** Study whether repeated exposure to such alerts leads to declining user responsiveness and to characterize its extent if present to better inform future CTA deployments.
- **Methods:** During a 36-week study period, we systematically documented the response patterns of 178 physician users randomized to receive CTAs for an ongoing clinical trial. Data were collected on: (1) response rates to the CTA; and (2) referral rates per physician, per time unit.

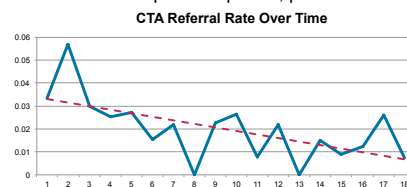
“Evaluating alert fatigue over time to EHR-based clinical trial alerts: findings from a randomized controlled study.”
(Embi & Leonard, et al. JAMIA, 2012)

- Significant declining trend in overall CTA response rate
 - 2.7% decline in each 2-week time period
 - Significantly different than zero (flat), $p < 0.0001$
- Notably – still at ~35% response rate after 36 weeks exposure



“Evaluating alert fatigue over time to EHR-based clinical trial alerts: findings from a randomized controlled study.”
(Embi & Leonard, et al. JAMIA, 2012)

- Decline in “referral” rates more pronounced than for “response” rates
 - 4.9% decrease per time period, $p = 0.0294$



“Evaluating alert fatigue over time to EHR-based clinical trial alerts: findings from a randomized controlled study.”

(Embi & Leonard, et al. JAMIA, 2012)

- **Conclusions:** CTA response rates declined over time, suggesting some contribution of alert fatigue. More so among community vs. university based MDs. Overall response rates remained relatively high, however (~35%) over the 36-week study.
- Approach to measuring alert-fatigue could be helpful when monitoring and applying in more widespread fashion.

Other notable papers for this section:

- **“Secondary use of routinely collected patient data in a clinical trial: An evaluation of the effects on patient recruitment and data acquisition.”** (Köpcke F, et al. Int J Med Inform. 2013)
- **“Connecting communities to health research: Development of the Project CONNECT minority research registry.”** (Green MA, et al. Contemporary clinical trials. 2013)
- **“A model for the design and implementation of a participant recruitment registry for clinical studies of older adults.”** (Dowling NM, et al. Clinical trials. 2012)
- **“Computational challenges and human factors influencing the design and use of clinical research participant eligibility pre-screening tools.”** (Pressler TR, et al. BMC medical informatics and decision making. 2012)

Policy and Perspectives

CRI Policy & Perspectives Pieces:

- **“Using EHRs to integrate research with patient care: promises and challenges.”** (Weng C. et al. JAMIA 2012)
- **“Opportunities and challenges for comparative effectiveness research (CER) with Electronic Clinical Data: a perspective from the EDM forum”** (Holve E. Medical care. 2012)
- **“Health services research evaluation principles. Broadening a general framework for evaluating health information technology”** (Sockolow PS et al. Meth Info in Med. 2012)
- **“Informatics and operations - let's get integrated”** (Marsolo K. JAMIA. 2013)
- **The SMART Platform: early experience enabling substitutable applications for electronic health records.** (Mandl KD. JAMIA. 2012)

Notable CRI-Related Events in Past Year

Release of the HIPAA Omnibus Rule – Jan 2013

- The Final Rule provisions has important implications for research:
 - Changes allowing compound authorizations should alleviate administrative burdens on clinical trial subjects and researchers and facilitate harmonization with the Common Rule and global requirements for research documentation.
 - Revised interpretation regarding authorization for future research use will remove barriers on researchers' ability to use data for future research purposes – some of which cannot even be contemplated at the time the data is gathered, but which could hold great promise to advance science and medical care.
 - The declassification as “PHI” of certain information of decedents over time will ease researchers' ability to perform research using such information.
- Researchers, research institutions and research sponsors have until September 23, 2013, to come into full compliance with the Final Rule. **Start now!**

Fiscal Cliff Legislation and Research... (Not all bad news)

- As noted, registries are growing in importance
- Incentives needed to encourage participation
- Legislation to avert the fiscal cliff (at least once) passed in January included a little known provision
 - Incentive to contribute to clinical data registries
 - Physicians get rewarded as if using current CMS PQRS reporting system
- Should help encourage adoption and contribution to specialty registries like those now being established by leveraging CRI innovations

Special Journal Issues dedicated to CRI Topics



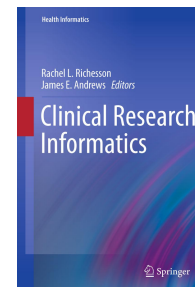
New e-Journal about Evidence Generating Activities: eGEMs

- Generating Evidence & Methods to improve patient outcomes
- Four Thematic Areas:
 - Methods
 - Informatics
 - Governance
 - Learning Health System
- Out of the EDM Forum of Academy Health
- Free, Open, Peer-reviewed



First of its kind textbook dedicated to CRI

- Editors: Richesson & Andrews
- Contributing authors from across our community
- A major achievement
- More evidence of CRI as established domain



- <http://www.springer.com/public+health/book/978-1-84882-447-8>
- <http://www.amazon.com/Clinical-Research-Informatics-Health/dp/1848824475>

In Summary...

- Maturing data infrastructure and sharing capabilities
- Increases in informatics approaches to re-use data accelerating and yielding real results
- Infrastructure advances maturing and beginning to accelerate and improve science
- Some (too few) evaluations, controlled studies
- Poised to deploy and test our approaches to realize the "learning health system"
- Exciting time to be in CRI!

Thanks!

Special thanks to:

- Philip Payne
- Eta Berner
- Rachel Richesson
- Adam Wilcox
- Shawn Murphy
- Chunhua Weng

Thanks!

Peter.Embi@osumc.edu

Slides will be posted on AMIA Website & on <http://www.embi.net/> (click on "Informatics")

