Personalized Itinerary Planner and Abstract Book

AMIA 2013 Annual Symposium
November 15 - 20, 2013

To make changes to your itinerary or view the full meeting schedule, visit
http://amiaannual2013.abstractcentral.com/itin.jsp
Friday, November 15, 2013
You have nothing scheduled for this day

Saturday, November 16, 2013
You have nothing scheduled for this day

Sunday, November 17, 2013

<table>
<thead>
<tr>
<th>Time</th>
<th>Session Info</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:30 AM-12:00 PM, Jefferson East (Washington Hilton), <strong>T11: Developing an i2b2 Cell and Client Plugin</strong></td>
<td></td>
</tr>
<tr>
<td>8:30 AM-12:00 PM</td>
<td><strong>Tutorial on developing an i2b2 Cell and Client Plugin</strong> M. Mendis; I.S. Kohane; S.N. Murphy</td>
</tr>
<tr>
<td>3:30 PM-5:00 PM, International Ballroom West (Washington Hilton), <strong>S04: Panel - Query Health: Toward a Learning Health System</strong></td>
<td></td>
</tr>
<tr>
<td>3:30 PM-5:00 PM</td>
<td><strong>Query Health: One Step Toward a Learning Health System</strong> J.G. Klann; M. Buck; J. Brown; S.N. Murphy; D.B. Fridsma</td>
</tr>
</tbody>
</table>

Monday, November 18, 2013

<table>
<thead>
<tr>
<th>Time</th>
<th>Session Info</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:30 AM-12:00 PM, Lincoln East (Washington Hilton), <strong>S22: Podium Presentations - Natural Language Processing</strong></td>
<td></td>
</tr>
<tr>
<td>10:30 AM-12:00 PM</td>
<td><strong>A Rule-based Natural Language Processing System in Tagging and Categorizing Phenotype Variables in NCBI's database of Genotypes and Phenotypes (dbGaP)</strong> S. Doan; K. Lin; R. Walker; S. Farzaneh; N. Alipanah; H. Kim</td>
</tr>
<tr>
<td>1:45 PM-3:15 PM, Georgetown (Washington Hilton), <strong>S26: Panel - How Fit is Electronic Health Data for its Intended Uses? Exploring Data Quality across Clinical, Public Health, and Research Use Cases</strong></td>
<td></td>
</tr>
<tr>
<td>1:45 PM-3:15 PM (Conflict)</td>
<td><strong>How Fit is Electronic Health Data for its Intended Uses? Exploring Data Quality across Clinical, Public Health, and Research Use Cases</strong> S.J. Grannis; B.E. Dixon; S.T. Liaw; M.G. Kahn; H. Fraser</td>
</tr>
<tr>
<td>1:45 PM-3:15 PM, Lincoln West (Washington Hilton), <strong>S33: System Demonstrations - Clinical Informatics</strong></td>
<td></td>
</tr>
<tr>
<td>1:45 PM-3:15 PM (Conflict)</td>
<td><strong>eTACTS: an Eligibility Tag Cloud-based Clinical Trial Search Engine</strong> R. Miotto; C. Weng</td>
</tr>
<tr>
<td>Time</td>
<td>Location</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>1:45 PM-3:15 PM</td>
<td>International Ballroom Center (Washington Hilton)</td>
</tr>
<tr>
<td>3:30 PM-5:00 PM</td>
<td>Jefferson West (Washington Hilton)</td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td>Columbia Hall (Washington Hilton)</td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td></td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td></td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td></td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td></td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td></td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td></td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td></td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>Session Info</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td>A Randomized Controlled Trial to Assess the Effects of a Mobile SMS-Based Intervention on Temporomandibular Disorder Treatment Adherence C.S. Prado; J. Duarte; M. Carvalho; C.F. Ortolani; E.S. Ruiz; I.T. Pisa</td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td>Integrating the CCDA for Real-Time Patient Data in the i2b2 Platform N. Wattanasin; M. Mendis; J. Mandel; R.B. Ramoni; K. Mandl; I.S. Kohane; S.N. Murphy</td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td>Highlights on the Inconsistency in Encoding Race and Ethnicity in database of Genotypes and Phenotypes (dbGaP) S. Feudjio Feupe; R. Walker; H. Kim</td>
</tr>
</tbody>
</table>

Tuesday, November 19, 2013

<table>
<thead>
<tr>
<th>Time</th>
<th>Session Info</th>
</tr>
</thead>
<tbody>
<tr>
<td>7:00 AM-8:30 AM</td>
<td>Morgan (Washington Hilton), Velos Corporate Roundtable - Opportunities for Disruptive Improvement in Clinical Research Information Infrastructure</td>
</tr>
<tr>
<td>10:30 AM-12:00 PM</td>
<td>S51: Papers - Case Studies to Improve HIT</td>
</tr>
<tr>
<td>10:30 AM-12:00 PM</td>
<td>Electronic End-of-Life Care Registry: the Utah ePOLST Initiative J. Duncan; P. Taillac; B. Nangle; M. Henry; J. Houston; S. Talebreza; L. Finch; C.P. Brunker; D. Rajeev; H.J. Smith; C. North</td>
</tr>
<tr>
<td>10:30 AM-12:00 PM</td>
<td>(Conflict) The NIH Office of Rare Diseases Research Patient Registry Standard: A Report from the University of New Mexico's Oculopharyngeal Muscular Dystrophy Patient Registry S.R. Daneshvari; S. Youssof; P.J. Kroth</td>
</tr>
<tr>
<td>10:30 AM-12:00 PM</td>
<td>Lincoln West (Washington Hilton), S56: Systems Demonstrations - Clinical Research Informatics</td>
</tr>
<tr>
<td>10:30 AM-12:00 PM</td>
<td>(Conflict) Clinical Phenotyping with the Analytic Information Warehouse A. Post; T. Kurc; R. Willard; H. Rathod; M. Mansour; A. Kalsanka Pai; W.M. Torian; S. Agravat; S. Sturm; J.H. Saltz</td>
</tr>
<tr>
<td>1:45 PM-3:15 PM</td>
<td>S63: Papers - Using NLP for information extraction A Natural Language Processing Algorithm to define a Venous Thromboembolism Phenotype E.R. McPeek Hinz; J.C. Denny; L. Bastarache</td>
</tr>
<tr>
<td>3:30 PM-5:00 PM</td>
<td>S76: Late Breaking Research Abstracts - Natural Language Processing and Factors Influencing Data Quality in Clinical Settings</td>
</tr>
<tr>
<td>Time</td>
<td>Session Info</td>
</tr>
<tr>
<td>-------------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>3:30 PM-5:00 PM</td>
<td>CTSI Pilot: Evaluating EHR Flowsheet Data Quality for Secondary Use in Research S. Johnson</td>
</tr>
<tr>
<td>3:30 PM-5:00 PM</td>
<td>Lincoln East/Monroe (Washington Hilton), S74: Podium Presentations - Data Repositories and Secondary Data</td>
</tr>
<tr>
<td>3:30 PM-5:00 PM</td>
<td>Amassing Pediatric Brain MRI's to Understand &quot;Normal&quot; using Mi2b2</td>
</tr>
<tr>
<td></td>
<td>S.N. Murphy; C. Herrick; V.M. Castro; R. Gollub; N. Reynolds; E. Grant</td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td>Northwest (Washington Hilton), Summit on Clinical Research Informatics</td>
</tr>
<tr>
<td></td>
<td>Scientific Program Committee Meeting</td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td>Using patient laboratory measurement values and dynamics to deconvolve EHR bias and define acuity-based phenotypes D. Albers; R. Pivovarov; G. Hripcsak; N. Elhadad</td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td>Novel Phenotype Development Using Non-Traditional Multilevel Population-Based Attributes M.K. Breitenstein; K.A. Monsen</td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td>Using SMART and i2b2 to Efficiently Identify Adverse Events J.G. Klann; R.B. Ramoni; S.N. Murphy</td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td>Secondary Use of a Statewide Electronic Health Record System for Patient-Centered Outcome Research: An Application to Anticoagulation Agents K. Ravvaz; M. Michalkiewicz; P.J. Tonellato</td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td>Creating a Secure, Easily Accessible Environment for PHI Data Exports within FURTHeR utilizing REDCap D. Schultz; B. LaSalle; S. He; R. Gouripeddi; R. Butcher; J.C. Facelli</td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td>Clinical Drug Extraction and Normalization from Clinical Notes S. Sohn; C. Clark; S.R. Halgrim; S.P. Murphy; C.G. Chute; H. Liu</td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td>Analysis of Clinical Decision Support Use in a Trial to Decrease Inappropriate Antibiotic Prescribing for Acute Respiratory Infections Y. Kleyner; S. Jones; J.N. Doctor; M.W. Friedberg; S.D. Persell; J.V. Falcone; D. Meeker; J.A. Linder</td>
</tr>
<tr>
<td>5:00 PM-6:30 PM</td>
<td>Impact on Immunization Registry Reporting Following the Adoption of an Electronic Health Record J. Merrill; A. Phillips; J. Keeling; R. Kaushal; Y. Senathirajah</td>
</tr>
<tr>
<td>5:30 PM-7:00 PM</td>
<td>International Ballroom West (Washington Hilton), Clinical Research Informatics Working Group Meeting</td>
</tr>
</tbody>
</table>

Wednesday, November 20, 2013
<table>
<thead>
<tr>
<th>Time</th>
<th>Location</th>
<th>Session Details</th>
</tr>
</thead>
</table>
| 8:30 AM-10:00 AM    | International Ballroom West     | **S77: Panel - Towards Computational Reuse of Clinical Research Eligibility Criteria with Collaboration across Academia, Industry, and Standardization Organizations**  
C. Weng; M. Cantor; A. Taweel; T.N. Arvanitis; R. Kush |
| 8:30 AM-10:00 AM    | Georgetown (Washington Hilton)  | **S78: Panel - Research Informatics : Re-engineering the Research Enterprise**  
M. Weiner; P. Payne; P.J. Embi; S.N. Murphy |
| 8:30 AM-10:00 AM    | Cabinet (Washington Hilton)     | **S84: Papers - Information Retrieval from Clinical Notes**  
C. Bejan; L. Vanderwende; H.L. Evans; M. Wurfel; M. Yetisgen-Yildiz |
| 8:30 AM-10:00 AM    | Lincoln West (Washington Hilton)| **S86: Late Breaking Research Abstracts - Machine Learning In Relation to EMRs**  
R. Xu |
| 10:30 AM-12:00 PM   | Lincoln West (Washington Hilton)| **S96: Late Breaking Research Abstracts - Machine Learning & Analysis Techniques in the Clinical Setting for Chronic Illness**  
C. Weng |
| 10:30 AM-12:00 PM   | Cabinet (Washington Hilton)     | **S93: Papers - Data Mining and Exchange for Clinical Applications**  
C. Jayapandian; C. Chen; A. Bozorgi; S.D. Lhatoo; G. Zhang; S.S. Sahoo |
Tutorial on developing an i2b2 Cell and Client Plugin

M. Mendis; I. S. Kohane; S. N. Murphy;

1. Massachusetts General Hospital, Boston, MA, United States.
2. Partners Healthcare, Boston, MA, United States.
3. Children's Hospital, Boston, MA, United States.

Abstract: Informatics for Integrating Biology and the Bedside (i2b2) (http://www.i2b2.org), is an open source software suite to construct and manage the clinical research chart in the genomic age. With it, query tool become generally available to researchers to search and work with pretention populations. This workshop will focus on the mechanics of setting up and populating an i2b2 database, and the more advanced topic of extending i2b2 software for custom uses needed at a site.
Abstract: "Measuring population health is critical to understanding the healthcare landscape. The Query Health initiative, a public-private collaboration convened by the Office of the National Coordinator for Health Information Technology (ONC) in 2011, is developing a standards-based, secure, distributed approach for measuring population health through electronic health record data, without sharing clinical data from individual sites. To date, Query Health has: developed a standards-based methodology; developed a reference implementation; and, launched several pilots to gauge its effectiveness for the purpose of population health management. This panel will present these accomplishments and future directions of the initiative. Content is appropriate for anyone with interest in population health. Prior experience with data analytics is helpful, but the presentations assume no prior knowledge of the technologies and standards used. Learning objectives: •Understand challenges in distributed clinical queries and how the Query Health methodology and reference implementation are enabling these goals. •Become familiar with the reference implementation, including how its components can be freely downloaded and implemented. •Discover how Query Health is enabling three different population health scenarios at three pilots. •Understand how Query Health fits into larger US initiatives like Meaningful Use. •Explore how Query Health could enhance attendee’s population-level healthcare analyses. "
A Rule-based Natural Language Processing System in Tagging and Categorizing Phenotype Variables in NCBI's database of Genotypes and Phenotypes (dbGaP)

S. Doan; K. Lin; R. Walker; S. Farzaneh; N. Alipanah; H. Kim;

1. Division of Biomedical Informatics, University of California, San Diego, La Jolla, CA, United States.

Abstract: We developed a new system for tagging and categorizing phenotype variables from NCBI’s database of Genotypes and Phenotypes (dbGaP) based on natural language processing (NLP). The system contains two components: 1) a tagger, which maps the terms in variable descriptions to a Unified Medical Language System (UMLS)’s Concept Unique Identifier (CUI) and then tags the CUIs with “topic” and “subject of information”, and 2) a categorizer, which groups tagged phenotype variables into predefined semantic classes. Evaluation of 500 phenotype variables showed that the system achieved 69.0% accuracy in semantic tagging and 83.28% in categorizing among the correctly tagged variables. This system will be integrated into PhenDisco (Phenotype Discoverer) – an information retrieval system for dbGaP, which is being developed in the Division of Biomedical Informatics at UC San Diego.
How Fit is Electronic Health Data for its Intended Uses? Exploring Data Quality across Clinical, Public Health, and Research Use Cases

S. J. Grannis; 1, 2; B. E. Dixon; 3, 2, 4; S. T. Liaw; 5; M. G. Kahn; 6, 7; H. Fraser; 8;

1. Department of Family Medicine, Indiana University, School of Medicine, Indianapolis, IN, United States.
2. Center for Biomedical Informatics, Regenstrief Institute, Indianapolis, IN, United States.
3. Department of BioHealth Informatics, Indiana University-Purdue University Indianapolis, School of Informatics and Computing, Indianapolis, IN, United States.
4. Center for Implementing Evidence-Based Practice, Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development Service, Indianapolis, IN, United States.
5. University of New South Wales, Sydney, NSW, Australia.
6. Department of Pediatrics, University of Colorado School of Medicine, Aurora, CO, United States.
7. Children's Hospital Colorado, Aurora, CO, United States.
8. Division of Global Health Equity, Brigham and Women's Hospital, Boston, MA, United States.

Abstract: Given growing adoption of electronic systems in health care, many providers and nations are looking to apply large volumes of electronic data across a wide spectrum of uses. Leveraging large-scale clinical and administrative data sets assumes that the data are of sufficient quality to enable valid, generalizable conclusions. Yet studies outside health care show that data quality (DQ) in information systems is often poor. In this panel, we explore the definition of DQ and how to measure it across research, clinical, and population health use cases. Each speaker provides unique context and approaches for approaching what could be a showstopper for efforts that aim to address health system challenges using Big Data from electronic health records and other informatics systems.
eTACTS: an Eligibility Tag Cloud-based Clinical Trial Search Engine

R. Miotto; 1 C. Weng; 1
1. Biomedical Informatics, Columbia University, New York, NY, United States.

Abstract: We present eTACTS, a Web search engine that allows users to select eligibility tags to filter clinical trial search results. A controlled vocabulary of frequent tags is automatically mined by cross-processing clinical trial eligibility criteria and used to index all the trials in ClinicalTrials.gov. After an initial search, these tags are presented to the users as an interactive tag cloud for iterative reduction of resulting trials. The cloud is dynamically updated at every tag selection with respect to the tag frequency and co-occurrences in the remaining trials.
Temporal Abstraction-based Clinical Phenotyping with Eureka!

A. Post; 1; T. Kurc; 1; R. Willard; 1; H. Rathod; 1; M. Mansour; 1; A. Kalsanka Pai; 1; W. M. Torian; 1; S. Agrawal; 1; S. Sturm; 1; J. H. Saltz; 1;

1. Biomedical Informatics, Emory University, Atlanta, GA, United States.

Abstract: Temporal abstraction, a method for specifying and detecting temporal patterns in clinical databases, is very expressive and performs well, but it is difficult for clinical investigators and data analysts to understand. Such patterns are critical in phenotyping patients using their medical records in research and quality improvement. We have previously developed the Analytic Information Warehouse (AIW), which computes such phenotypes using temporal abstraction but requires software engineers to use. We have extended the AIW’s web user interface, Eureka! Clinical Analytics, to support specifying phenotypes using an alternative model that we developed with clinical stakeholders. The software converts phenotypes from this model to that of temporal abstraction prior to data processing. The model can represent all phenotypes in a quality improvement project and a growing set of phenotypes in a multi-site research study. Phenotyping that is accessible to investigators and IT personnel may enable its broader adoption.
Abstract: Integrated data repositories (IDRs) are indispensable tools for numerous biomedical research studies. We compare three large IDRs (Informatics for Integrating Biology and the Bedside (i2b2), HMO Research Network’s Virtual Data Warehouse (VDW) and Observational Medical Outcomes Partnership (OMOP) repository) in order to identify common architectural features that enable efficient storage and organization of large amounts of clinical data. We define three high-level classes of underlying data storage models and we analyze each repository using this classification. We look at how a set of sample facts is represented in each repository and conclude with a list of desiderata for IDRs that deal with the information storage model, terminology model, data integration and value-sets management.
Role of ICD Granularity in Phenotyping Hematologic Malignancies for Tumor Registries
R. V. Atreya; T. A. Lasko; M. A. Levy;  
1. Biomedical Informatics, Vanderbilt University School of Medicine, Nashville, TN, United States.  
2. Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, TN, United States.  
3. Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN, United States.  

Abstract: Tumor registries play a critical role in the epidemiological study of cancer. However, due to the manual construction process, they are very costly to maintain. There is an opportunity to develop tools to streamline the tumor registry development process and build towards a cancer learning system. In this study, we use ICD-9 billing codes to predict what kind of hematologic malignancy a patient has and seek to identify the threshold of disease classification granularity at which ICD-9 codes are no longer sufficient. Through the use of multinomial logistic regression, we demonstrate that ICD-9 code counts can acceptably classify broad, clinically relevant hematologic malignancies. However, they cannot predict more detailed schemes such as ICD-O-3 (an ICD-10 oncology derivative). This work identifies the threshold at which ICD-9 codes cannot be used to classify diseases, and could provide insight on the ICD-9 to ICD-10 transition.
The Computerized Research Record (CoRR): A Web-based Research Record System for Managing Research Support Requests Based on an EHR Metaphor

P. J. Embi; 1; M. A. Lopetegui; 1; F. Lamantia; 1; T. B. Borlawsky; 1; D. Hamon; 1; T. Nielsen; 1; R. Rice; 1;
1. Dept of Biomedical Informatics, The Ohio State University, Columbus, OH, United States.

Abstract: Researchers need access to institutional support in the form of expertise, services and resources. Providing efficient and effective intake and management of requests for such support requires handling them in a systematic and coordinated manner. However, the complexities and diversity of resources makes use of a simple solution challenging. We addressed this problem by developing a Computerized Research Record (CoRR) System. The system was designed based on an Electronic Health Record metaphor, wherein the researcher – service provider relationship is akin to a patient – healthcare provider relationship. Development of CoRR involved an iterative process, including user-input. We describe the design, development and usage statistics of the CoRR system.
Abstract: We are building an expandable framework to better understand the research aspects in clinical research informatics (CRI). The framework can be used to 1) identify areas where informatics advances can be applied to clinical research, 2) identify gaps in current technologies, and 3) assist new researchers and students in CRI to better understand values and applications of CRI.
Implementing QDAct-PC, a Continuous Learning System for Palliative Care
A. Kamal, 1, 2; M. B. Adams; 2; H. Shang; 2; J. Kelly; 2; A. Abernethy; 1, 2;
1. Duke Cancer Institute, Durham, NC, United States.
2. Center for Learning Health Care, Duke Clinical Research Institute, Durham, NC, United States.

Abstract: Palliative care continues to grow to care for the increasing number of persons with chronic and serious illness. In-step with remarkable clinical growth remains a timely need to rapidly and continuously build an ever-changing knowledge base to impact treatment decisions, near real time, and in a variety of care settings. Using a novel, hands-on multidisciplinary team spanning fields of medicine, information technology and quality; we developed the Quality Data Collection Tool for Palliative Care (QDAct-PC). Testing has proven its usability and feasibility as an electronic solution to perform rapid learning in community palliative care. Early experiences have been valuable to inform the underlying technical approach, quality improvement, decision-making, and research efforts. Further longitudinal experiences will inform iterative refinements.
Abstract: This study describes a domain analysis model (DAM) developed to represent the data stored in the database of Genotypes and Phenotypes (dbGaP). To conform to information model standards, we mapped our DAM to Biomedical Research Integrated Domain Group (BRIDG) model and the caBIG Life Science DAM (LS DAM). Our DAM was represented using BRIDG and LS DAM with a minor extension, which was necessary to cover the data use agreement class of our model.
mHealth for the CDU Electronic Disease Registry to Improve Chronic Care (CEDRIC)

S. Mukherjee; O. Ogunyemi; J. McDonough; R. de Castro;

1. Center for Biomedical Informatics, Charles R. Drew University of Medicine And Science, Lynwood, CA, United States.
2. Shorthand Mobile, San Diego, CA, United States.

Abstract: To assist Los Angeles urban safety net clinics in managing diabetes patients we created CEDRIC, then added a GIS module to contextualize factors that affect a patient’s ability to adhere to recommended behavior modifications. Now, to facilitate educating patients on diabetes, reporting behavior modification progress and medication adherence, blood glucose monitoring, appointment scheduling, and alerting to self-management tasks we have added a mHealth module to CEDRIC as well.
Final ID:
PhenotypePortal: An Open-Source Library and Platform for Authoring, Executing and Visualization of Electronic Health Records Driven Phenotyping Algorithms

J. Pathak; 1; C. Endle; 1; D. B. Suesse; 1; K. Peterson; 1; C. Stancl; 1; D. Li; 1; C. G. Chute; 1;
1. Health Sciences Research, Mayo Clinic, Rochester, MN, United States.

Abstract: Secondary use of electronic health record (EHR) data is a broad domain that includes clinical quality metrics, observational cohorts, outcomes research and comparative effectiveness research. A common thread across all these use cases is the design, implementation, and execution of EHR-driven algorithms for identifying subjects with the phenotypes of interest, and the subsequent analysis of the query results. In this abstract, we report our preliminary efforts in designing and implementing an open-source library and platform for authoring, executing, and visualization of EHR phenotyping algorithms.
Behavioral Economics-Informed EHR-Supported Interventions to Reduce Inappropriate Antibiotic Prescribing: a Cluster Randomized Trial

S. D. Persell; 1; J. A. Linder; 2; M. W. Friedberg; 3; D. Meeker; 4; E. M. Friesema; 1; A. Cooper; 1; C. R. Fox; 5; N. J. Goldstein; 5; J. N. Doctor; 6;
1. General Internal Medicine and Geriatrics, Northwestern University, Chicago, IL, United States.
2. Division of General Medicine and Primary Care, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, United States.
3. RAND, Boston, MA, United States.
4. RAND, Santa Monica, CA, United States.
5. University of California, Los Angeles, Los Angeles, CA, United States.
6. School of Pharmacy, University of Southern California, Los Angeles, CA, United States.

Abstract: Physicians often use antibiotics for acute respiratory infections (ARIs) when they are not indicated. Interventions to reduce inappropriate prescribing have been only modestly successful. We performed a pilot cluster-randomized controlled clinical trial with physician as the unit of randomization and office visit as the unit of analysis to evaluate different behavioral economic and social psychological approaches to reduce inappropriate antibiotic prescribing. Participants were randomized within a 2 x 2 x 2 factorial experiment. The three factors included: 1) Accountable Justifications 2) Suggested Alternatives and 3) Peer Comparison. The primary outcome was the rate of antibiotic prescribing for visits with an ARI diagnosis for which antibiotics are not indicated (e.g. acute bronchitis, unspecified upper respiratory infection). Secondary endpoints were prescribing rates for (acute sinusitis or pharyngitis), and prescribing rates for all other respiratory infections or symptoms of respiratory infection. Antibiotic prescribing fell substantially in the intervention year compared to the prior year for non-antibiotic-appropriate ARI diagnoses (from 24.0% to 4.6%, p<0.001). There were also declines in prescribing for sinusitis/pharyngitis (from 47.5% to 36.9%, p=0.001) and for all other respiratory infections or symptoms of infection (16.9% to 5.4%, p<0.001). The absolute number of annual ARI or respiratory encounters was similar in both years, and the proportions of encounters in each diagnosis group were similar. Substantial reductions in prescribing were observed even among control providers and changes in prescribing behavior were similar for providers receiving zero, 1, 2 or 3 of the interventions.
Abstract: The characterization of a patient, whether for clinical care or clinical research, is destined to become heavily dependent upon knowing an individual’s full genome. The development of next generation sequencing (NGS) has allowed the whole genome to be expressed in a facile manner, and informatics for integrating biology and the bedside (i2b2) is allowing the genome to be queried alongside the phenome of the patient. These queries are enabled by a data loading process that transfers the products of NGS pipelines into 10-20 million variant observations expressed in the Sequence Ontology. These observations can then exist alongside the many other observations about a patient and allow the formulation of complex queries that express the state of an individual.
A Randomized Controlled Trial to Assess the Effects of a Mobile SMS-Based Intervention on Temporomandibular Disorder Treatment Adherence

C. F. Ortolani; 2  C. S. Prado; 1  J. Duarte; 1  M. Carvalho; 1  E. S. Ruiz; 3  I. T. Pisa; 1

1. Health Informatics, UNIFESP, São Paulo, São Paulo, Brazil.
2. Orthodontics Department, UNIP, São Paulo, São Paulo, Brazil.
3. Computing and Mathematics Department, USP, São Paulo, São Paulo, Brazil.

Abstract: The use of mobile phone short message service (SMS) leveraging follow-up therapy has been studied from various areas of health care. We present a research project aimed to assess the effects of this messaging media as intervention technique on temporomandibular disorder treatment adherence during by a randomized controlled trial. The ongoing study has its initial phase completed including message system development approved for clinical study implementation.
Integrating the CCDA for Real-Time Patient Data in the i2b2 Platform

N. Wattanasin; M. Mendis; J. Mandel; R. B. Ramoni; K. Mandl; I. S. Kohane; S. N. Murphy;

1. Partners HealthCare System, Boston, MA, United States.
2. Harvard Medical School, Boston, MA, United States.
3. Children’s Hospital Boston, Boston, MA, United States.
4. Massachusetts General Hospital, Boston, MA, United States.

Abstract: The Substitutable Medical Apps, Reusable Technologies (SMART) project provides a framework of core services to facilitate the use of substitutable health-related web applications. The platform offers a common interface used to “SMART-ready” health IT systems allowing any SMART application to be able to interact with those systems. At Partners Healthcare, we have SMART-enabled the Informatics for Integrating Biology and the Bedside (i2b2) open source analytical platform and developed an i2b2 Services ETL ‘cell’ that can consume the Consolidated Clinical Document Architecture (CCDA), enabling SMART applications to interact with both retrospective patient data from i2b2 and real-time patient data from the EMR.
Final ID:

Highlights on the Inconsistency in Encoding Race and Ethnicity in database of Genotypes and Phenotypes (dbGaP)

S. Feudjio Feupe; 1; R. Walker; 1; H. Kim; 1;

1. Medicine - Biomedical Informatics, UCSD, La Jolla, CA, United States.

Abstract: This study reports on the three types of inconsistent encoding of race and ethnicity information that we observed in the database of Phenotypes and Genotypes (dbGaP). Standardized terminologies also showed inconsistencies in handling race and ethnicity concepts. Impacts that the inconsistent race and ethnicity encoding can have on data retrieval in dbGaP are described.
Final ID:

Electronic End-of-Life Care Registry: the Utah ePOLST Initiative

J. Duncan; 1, 2; P. Taillac; 1, 2; B. Nangle; 1; M. Henry; 2; J. Houston; 1; S. Talebreza; 2, 5, 6; L. Finch; 1; C. P. Brunker; 2, 3; D. Rajeev; 4; H. J. Smith; 4; C. North; 4;

1. Utah Department of Health, Salt Lake City, UT, United States.
2. University of Utah, Salt Lake City, UT, United States.
3. Intermountain Healthcare, Salt Lake City, UT, United States.
4. HealthInsight, Murray, UT, United States.
5. Veterans Administration, Salt Lake City, UT, United States.
6. Haven Healthcare, Salt Lake City, UT, United States.

Abstract: As a patient's end-of-life approaches, it is typical for the disease to be the focus of treatment instead of the dying patient. There is limited congruence between the care preferred by patients and the treatment actually delivered to patients during their end-of-life. The Physician Orders for Life-Sustaining Treatment Paradigm has been endorsed or is in development in all but four states and districts in the US in an effort to ensure that patients are provided with adequate opportunities to specify their end-of-life care preferences. However, most states are using paper forms to document these preferences which may be inaccessible when needed. We have developed an electronic end-of-life care registry that allows authorized users to store and retrieve information pertaining to patients' end-of-life care preferences. In this paper, we describe (a) the requirements identified for the registry from the users' perspective and (b) the design and development of the electronic registry.
The NIH Office of Rare Diseases Research Patient Registry Standard: A Report from the University of New Mexico’s Oculopharyngeal Muscular Dystrophy Patient Registry

S. R. Daneshvari; S. Youssof; P. J. Kroth; 1
1. School of Medicine, University of New Mexico, Albuquerque, NM, United States.

Abstract: Patient registries remove barriers to performing research by assembling patient cohorts and data in a systematic, efficient, and proactive manner. Consequently, registries are a valuable strategy for facilitating research and scientific discovery. Registries for rare diseases are arguably even more valuable since there is difficulty in assembling cohorts of adequate size for study. Recently, the NIH Office of Rare Diseases Research created a rare disease registry Standard to facilitate research across multiple registries. We implemented the Standard for the Oculopharyngeal Muscular Dystrophy patient registry created at the University of New Mexico Health Sciences Center. We performed a data element analysis for each Common Data Element defined in the Standard. Problems included the use of previous HL7 versions, non-structured data types, and a recent update to the Standard. Overall, the Standard is an excellent first step toward standardizing patient registries to facilitate work on broader questions and promote novel interdisciplinary collaborations.
Clinical Phenotyping with the Analytic Information Warehouse

A. Post; 1; T. Kurc; 1; R. Willard; 1; H. Rathod; 1; M. Mansour; 1; A. Kalsanka Pai; 1; W. M. Torian; 1; S. Agravat; 1; S. Sturm; 1; J. H. Saltz; 1;

1. Center for Comprehensive Informatics, Emory University, Atlanta, GA, United States.

Abstract: Healthcare quality improvement and research increasingly leverage large electronic health record (EHR) datasets to understand and compare patient populations. Much information exists in the medical record implicitly as billing codes, clinical events and observations, and concepts embedded in text reports. Clinical phenotyping promises to infer useful diagnostic, therapeutic response and process information from those data using temporal pattern finding, concept extraction from text and other techniques. The Analytic Information Warehouse (AIW) is our open source clinical phenotyping system. We have implemented a web user interface for AIW, Eureka! Clinical Analytics, for clinical investigators and IT personnel to specify phenotypes as temporal patterns, perform data processing, and extract found phenotypes and data into existing analysis tools. Eureka aims to enable clinical investigators and analysts to manage phenotype definitions and perform phenotyping in a highly automated fashion with limited IT resources. We will demonstrate Eureka’s data import, phenotype specification and data processing functionality. These features aim to support analytics techniques such as “hot spotting” and predictive modeling, as well as clinical research that uses routinely collected data to understand how interventions perform in the real world.
A Natural Language Processing Algorithm to define a Venous Thromboembolism Phenotype
E. R. McPeek Hinz; 1; J. C. Denny; 2; L. Bastarache; 2;
1. DHTS, Duke University Health System, Durham, NC, United States.
2. Department of BioMedical Informatics, Vanderbilt University Medical Center, Nashville, TN, United States.

Abstract: Deep venous thrombosis and pulmonary embolism are diseases associated with significant morbidity and mortality. Known risk factors are attributed for only slight majority of venous thromboembolic disease (VTE) with the remainder of risk presumably related to unidentified genetic factors. We designed a general purpose Natural Language (NLP) algorithm to retrospectively capture both acute and historical cases of thromboembolic disease in a de-identified electronic health record. Applying the NLP algorithm to a separate evaluation set found a positive predictive value of 84.7% and sensitivity of 95.3% for an F-measure of 0.897, which was similar to the training set of 0.925. Use of the same algorithm on problem lists only in patients without VTE ICD-9s was found to be the best means of capturing historical cases with a PPV of 83%. NLP of VTE ICD-9 positive cases and non-ICD-9 positive problem lists provides an effective means for capture of both acute and historical cases of venous thromboembolic disease.
Abstract Body: Academic health centers (AHCs) funded through Clinical Translational Science Awards (CTSA) develop infrastructure to accelerate research within the community and across CTSA sites. Clinical data repositories (CDRs) are key infrastructure components; however, the type of data included varies across CTSA sites. Only two CTSA sites were found that include electronic health record (EHR) flowsheet data. The goal of our research was to extend the University of Minnesota’s CDR with flowsheet data and determine the data quality for future research. We identified five clinical conditions, the policies for documentation, and the consistency of documentation with these policies. The five clinical conditions are: falls assessment, pressure ulcer assessment and prevention, pain management, urinary catheter management, and venous thrombosis embolism (VTE) prevention.
**Abstract:** Many neurological and psychiatric illnesses are, for mysterious reasons, not manifested until a person becomes an adult. Pre-illness data for diseases such as epilepsy and schizophrenia, limited to cohorts of high risk probands, are difficult and very expensive to collect because there are no valid predictive biomarkers. However, in the normal course of care, and now in the electronic health record, a large amount of observational data can be captured. We are focusing on the brain Magnetic Resonance Images (MRI’s) of children, acquired prior to disease onset, to find characteristics predictive of or associated with subsequent disease onset. In the course of understanding any observational clinical data that has not previously been used for research, we must characterize the normal state and its variance. We are doing this with the massive medical image data repositories that exist at the Massachusetts General Hospital and Boston Children’s Hospital with freely available software from Informatics for Integrating Biology and the Bedside (i2b2).
Using patient laboratory measurement values and dynamics to deconvolve EHR bias and define acuity-based phenotypes

D. Albers; 1; R. Pivovarov; 1; G. Hripcsak; 1; N. Elhadad; 1;
1. Biomedical Informatics, Columbia University, New York, NY, United States.

Abstract: Patients are measured for many reasons ranging from acute health episodes to regular check ups. Understanding the impact of these reasons on large-scale measurement analysis is an open research problem. We focus here on lab test measurement. Analysis of measurement dynamics (i.e., temporal patterns of measurements for a given lab across patient populations) against lab values informs us of ways to treat different labs when conducting EHR-based phenotyping activities.
Novel Phenotype Development Using Non-Traditional Multilevel Population-Based Attributes

M. K. Breitenstein; ¹; K. A. Monsen; ², ¹;
1. Institute for Health Informatics, University of Minnesota, Minneapolis, MN, United States.
2. School of Nursing, University of Minnesota, Minneapolis, MN, United States.

Abstract: Population-based interventions aim to increase health outcomes by solving problems at multiple social levels (individual to population) that are not traditionally healthcare related. Public Health Nurses were found to administer their population-based interventions on nontraditional healthcare attributes and across all social levels. Current informatics solutions do not appropriately synthesize nontraditional multilevel attributes, prohibiting appropriate allocation of resources despite the strength population-based intervention frameworks. Non-traditional multilevel patient attributes, as detailed in our poster, need to be accounted for in future phenotyping endeavors.
Final ID:

Using SMART and i2b2 to Efficiently Identify Adverse Events

J. G. Klann; 1, 2, 3 ; R. B. Ramoni; 4 ; S. N. Murphy; 1, 2, 3 ,
1. Laboratory of Computer Science, Massachusetts General Hospital, Boston, MA, United States.
2. Harvard Medical School, Boston, MA, United States.
4. Center for Biomedical Informatics, Harvard Medical School, Boston, MA, United States.

Abstract: New healthcare reimbursement models add urgency to monitoring patient safety. The widespread adoption of electronic health records could accelerate this, but automatic adverse event detection has had disappointing accuracy. However, the Institute for Healthcare Improvement’s manual review approach is highly accurate. We demonstrate a partial automation of this approach using the i2b2 data analytics system and SMART apps platform. Because of i2b2’s broad adoption, we believe this can widely enable accurate patient safety assessment.
Secondary Use of a Statewide Electronic Health Record System for Patient-Centered Outcome Research: An Application to Anticoagulation Agents

K. Ravvaz; 1; M. Michalkiewicz; 2; P. J. Tonellato; 1;
1. Zilber School of Public Health, Biomedical and Health Informatics, University of Wisconsin-Milwaukee, Milwaukee, WI, United States.
2. Patient Centered Research, Aurora Health Care Research Institute, Milwaukee, WI, United States.

Abstract: This multistage joint project has been designed and undertaken to develop a retrospective EMR-based longitudinal anticoagulation clinical database being used for patient-centered outcome research on most frequently prescribed anticoagulation agents such as warfarin. The developed normalized “Anticoagulation” registry is invaluable to pursue the broad objectives of individualized and thus, improved outcome patient-centric research.
Creating a Secure, Easily Accessible Environment for PHI Data Exports within FURTHeR utilizing REDCap

D. Schultz; 1; B. LaSalle; 1; S. He; 1; R. Gouripeddi; 1; R. Butcher; 1; J. C. Facelli; 1;

1. Biomedical Informatics, University of Utah, Salt Lake City, UT, United States.

Abstract: We report on the use of an open source, HIPAA compliant tool, Research Electronic Data Capture (REDCap) as a low cost, user-friendly and secure environment for storing federated data queries, which include Protected Health Information (PHI).
**Abstract:** In electronic medical records (EMR), the same drug information is often described using different words and patterns, which can cause errors in the administration of drugs and jeopardize patients' safety. Effective drug management requires standards for the drug names so that the standards can be mapped to various drug names from different sources. RxNorm provides this standard in that its unique concept ID (RxCUI) maps to variations in the drug names and it is being part of Meaningful Use of increasing functionality of health record technology. In this study, our aim was to extract complete drug information from clinical narratives and normalize it with RxCUI.
Analysis of Clinical Decision Support Use in a Trial to Decrease Inappropriate Antibiotic Prescribing for Acute Respiratory Infections

Y. Kleyner; 1; S. Jones; 2; J. N. Doctor; 3; M. W. Friedberg; 2, 4; S. D. Persell; 5; J. V. Falcone; 1; D. Meeker; 2; J. A. Linder; 1, 4;
2. RAND, Santa Monica, CA, United States.
3. University of Southern California, Los Angeles, CA, United States.
4. Department of General Internal Medicine, Brigham and Women's Hospital, Boston, MA, United States.
5. Northwestern University, Chicago, IL, United States.

Abstract: Understanding how clinical decision support (CDS) influences clinician behavior can inform the development of more effective systems. We examined CDS use from an ongoing cluster randomized trial CDS based on behavioral economic principles to decrease inappropriate antibiotic prescribing for acute respiratory infections (ARIs).
Abstract: The increased adoption of electronic health records (EHRs) is anticipated to increase the use and efficiency of immunization and other public health registries. State and local registries are important to public health by consolidating often-fragmented health data. This study evaluated 1.7 million immunization records submitted to the NYC Citywide Immunization Registry from 2007 to 2011, a period of increased EHR adoption. The study found record submissions from the immunization module supplanting other submission methods and an improved timeliness of submissions.
Towards Computational Reuse of Clinical Research Eligibility Criteria with Collaboration across Academia, Industry, and Standardization Organizations

C. Weng;¹; M. Cantor;², ³; A. Taweel;⁴; T. N. Arvanitis;⁵; R. Kush;⁶;

1. Department of Biomedical Informatics, Columbia University, New York, NY, United States.
3. School of Medicine, New York University, New York, NY, United States.
4. Department of Informatics, King's College London, London, Strand, United Kingdom.
5. Department of Computer Engineering, The University of Birmingham, Edgbaston, Birmingham, United Kingdom.
6. CDISC, Austin, TX, United States.

Abstract: Central to clinical and translational research activities, clinical research eligibility criteria are perceived, defined, interpreted, and implemented by various stakeholders in a series of translations during protocol authoring, research subject recruitment, electronic screening for cohort selection for clinical research studies, systematic reviews, and evidence-based medicine. The efficiency and integrity of these research activities heavily rely on the quality and clarity of clinical research eligibility criteria. However, the free-text clinical research eligibility criteria are fraught with problems, including ambiguities and the lack of representativeness of the real-world patient population. In the past 25 years, the biomedical informatics community, from academia to industry, and from USA to UK, has spent significant efforts, combining knowledge representation and natural language processing, to improve the standardization and computational reuse of clinical research eligibility criteria across various stakeholders throughout the clinical and translational research pipeline.
Abstract: Recent incentives favoring the adoption and meaningful use of information technology in clinical practice are helping to increase the comprehensiveness and quality of routinely collected electronic patient information. Research informaticists play a vital role in capturing, organizing, visualizing, analyzing, understanding and relaying to others the full breadth, depth and longitudinal scope of these data. This panel will describe how the discipline of research informatics and the products of this work add value to multidisciplinary research teams and interinstitutional partnerships that span the full spectrum of research endeavors, from basic science innovations to translational research and clinical trials to health services research and post marketing safety surveillance.
Abstract: In this paper we describe a natural language processing system which is able to predict whether or not a patient exhibits a specific phenotype using the information extracted from the narrative reports associated with the patient. Furthermore, the phenotypic annotations from our report dataset were performed at the report level which allows us to perform the prediction of the clinical phenotype at any point in time during the patient hospitalization period. Our experiments indicate that an important factor in achieving better results for this problem is to determine how much information to extract from the patient reports in the time interval between the patient admission time and the current prediction time.
Creation and Comparative Analysis of a Novel Disease Phenotype Network Based on Clinical Manifestation

R. Xu; 1

1. Case Western Reserve University, Cleveland, OH, United States.

Abstract Body: Systematic studies of disease phenotype networks have been used to link diseases to their causing genes. Clinical manifestation is an important aspect of phenotype data and can be used to predict disease causes. We propose to create a disease manifestation network (DMN) using the highly accurate and structured manifestation data from multiple biomedical ontologies. We found significant correlations between manifestation similarities and genetic associations. Network comparisons demonstrated that DMN not only reflects genetic mechanisms of diseases, but also contains novel information to complement the existing phenotype network. When combined with other phenotype data, DMN has great potential to facilitate deeper understandings of disease etiology.
Using Electronic Health Records To Assess Generalizability of Clinical Trials

C. Weng; 1

1. Columbia University, New York, NY, United States.

Abstract Body: This study used a distribution-based method to assess the generalizability of the target populations of 1,761 Type 2 diabetes randomized controlled trials downloaded from ClinicalTrials.gov. The trial target populations and patient population in one metropolitan hospital were compared using hemoglobin A1c (HbA1c) and ages at various (e.g., latest or earliest) HbA1c measures. The results show that Type 2 diabetes clinical trials systematically target the young and sick (i.e., with higher HbA1c values) ones among Type 2 diabetes patients.
Cloudwave: Distributed Processing of “Big Data” from Electrophysiological Recordings for Epilepsy Clinical Research Using Hadoop

C. Jayapandian; 1; C. Chen; 1; A. Bozorgi; 2; S. D. Lhatoo; 2; G. Zhang; 1; S. S. Sahoo; 1;
1. Division of Medical Informatics, Case Western Reserve University, Cleveland, OH, United States.  
2. Department of Neurology, Case Western Reserve University, Cleveland, OH, United States.

Abstract: Epilepsy is the most common serious neurological disorder affecting 50-60 million persons worldwide. Multimodal electrophysiological data, such as electroencephalography (EEG) and electrocardiography (EKG), are central to effective patient care and clinical research in epilepsy. Electrophysiological data is an example of clinical “big data” consisting of more than 100 multi-channel signals with recordings from each patient generating 5-10GB of data. Current approaches to store and analyze signal data using standalone tools, such as Nihon Kohden neurology software, are inadequate to meet the growing volume of data and the need for supporting multi-center collaborative studies with real time and interactive access. We introduce the Cloudwave platform in this paper that features a Web-based intuitive signal analysis interface integrated with a Hadoop-based data processing module implemented on clinical data stored in a “private cloud”. Cloudwave has been developed as part of the National Institute of Neurological Disorders and Strokes (NINDS) funded multi-center Prevention and Risk Identification of SUDEP Mortality (PRISM) project. The Cloudwave visualization interface provides real-time rendering of multi-modal signals with “montages” for EEG feature characterization over 2TB of patient data generated at the Case University Hospital Epilepsy Monitoring Unit. Results from performance evaluation of the Cloudwave Hadoop data processing module demonstrate one order of magnitude improvement in performance over 77GB of patient data. (Cloudwave project: http://prism.case.edu/prism/index.php/Cloudwave)