Using EHR Data to Conduct PBRN Research

Laura-Mae Baldwin MD MPH – University of Washington, lmb@uw.edu
Alex Fiks MD – Children's Hospital of Philadelphia, fiks@email.chop.edu
Alex Krist MD MPH – Virginia Commonwealth University, ahkrist@vcu.edu

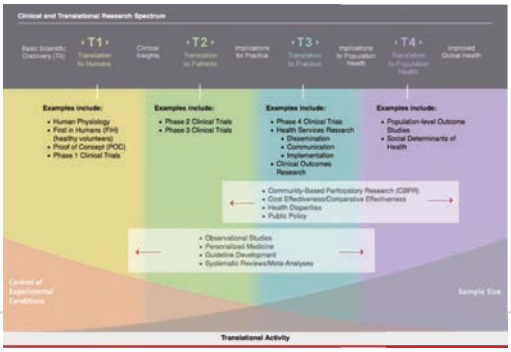
Goals

- Understand how to get started with using EHR data and HIT across the PBRN research lifecycle.
- Be aware of the possible uses for EHR data and HIT in conducting PBRN research.
- Identify the facilitators and barriers to incorporating EHR data and HIT in your research projects.
- Develop strategies to ensure that PBRN partners benefit from EHR and HIT-based projects.

Introduction

4

How EHR Data and HIT Support PBRN Research Across the Translational Science Spectrum



5

How EHR Data and HIT Support "The Research Lifecycle"

Research Lifecycle Stage	Examples of Support Provided by EHR Data and HIT
Building Collaborations	• Tailor HIT and EHR data queries to answer questions important to practices
Identifying/Choosing the Problem	• EHR data-based discovery
Stating the Research Question	
Developing a research approach	• HIT-based intervention • EHR data may provide the research data -- define outcomes, delineate mediator and moderator variables.
Select a sample	• EHR can facilitate cohort discovery
Collect high quality data	• EHR can provide discrete data fields, standardized coding (e.g., ICD, CPT, SNOMED)
Analyze and interpret data	
Disseminate findings	• EHR-based patient portals can communicate with patients

6

Research New Functions and New Technology

Creating new technology or modifying existing systems is hard…


Electronic health record functionality needed to better support primary care


SPECIAL ARTICLES

A Family Medicine Health Technology Strategy for Achieving the Triple Aim for US Health Care

Think about partners for collaboration; create a multidisciplinary team; align with your EHR vendor; find developers who have done something similar.

7

Don't Forget Dissemination & Implementation Research!

Improving how we use EHRs and technology may be more important than what we use.


The Relationship between Electronic Health Record Use and Quality of Care over Time

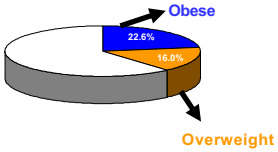
Frameworks exist for evaluating Dissemination & Implementation outcomes (e.g. www.RE-AIM.org).

8

Roles of the EHR in Research

Pilot Data – Obesity Study

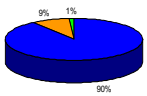
- Demographics:
  - Eligible – 3,734
  - Female – 49.5%
  - Age – 9.7 years (6.6 – 12.8)
  - Time between visits – 395 days (322 to 552)



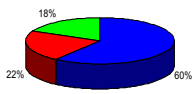
9

Longitudinal EHR Data Facilitates Further Analysis:

If "obese' at baseline, then at second visit....



If "overweight" at baseline, then at second visit...



Remained same

Returned to healthy weight

Overweight

Obese

10

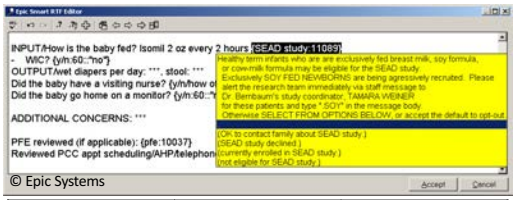
Subject Recruitment
Patient lists delivered to research assistants

DEPARTMENT_NAME	APPT_TIME	PROV_NAME	PAT_MRN_ID	BIRTH_DATE	AGE	SEX
KDH KF DREXEL HILL	10:00:00 AM	XXX, RICHARD	########	##/##/1994	12	M
KDH KF DREXEL HILL	10:15:00 AM	XXX, RICHARD	########	##/##/1994	12	M
KDH KF DREXEL HILL	11:15:00 AM	XXX, RICHARD	########	##/##/1994	12	M
KDH KF DREXEL HILL	12:00:00 PM	XXX, RICHARD	########	##/##/1998	8	M
KDH KF DREXEL HILL	3:00:00 PM	XXX, RICHARD	########	##/##/1994	12	M
KDH KF DREXEL HILL	3:30:00 PM	XXX, RICHARD	########	##/##/1997	9	M

- **Lists of appointments for potential subjects that favor *sensitivity* over** *specificity*
- **Sent each Friday to research team**
- **Study team can target specific dates/times to be on-site to maximize their recruitment potential**

11

Subject Recruitment:
Extracting Data from Templates

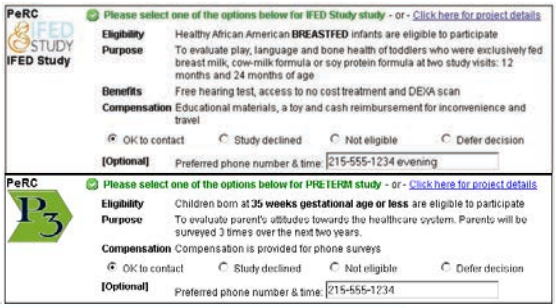


	EHR	# Enrolled
PeRC	Yes	410 in 12 months
Site B	No	86 in 18 months

Grundmeier RW et al., "Research Subject Recruitment Using the Electronic Health Record, AMIA Annu Symp Proc, 2007.

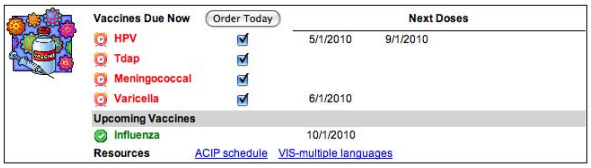
12

Enrollment: EHR-Based Forms



13

Using the EHR as the Intervention: Decision Support



Fiks AG, Grundmeier RW, Mayne S, Song L, Feemster K, Karavite D, Hughes CC, Massey J, Keren R, Bell LM, Wasserman R, Localio AR: "Effectiveness of Decision Support for Families, Clinicians, or Both on HPV Vaccine Receipt" Pediatrics 131: 1114-24, 2013.

14

Feedback Report

- Performance feedback—clinicians made aware of their own vaccination rates and how they compare to their practice and care network.
- Generated from EHR data

	Human Papilloma Virus (HPV) Vaccine Given to Eligible Adolescent								
Visit Type	You			Your Practice			Network		
	Number of Visits at which HPV Vaccine Due	Number of Visits at which HPV Vaccine Given*		Number of Visits at which HPV Vaccine Due	Number of Visits at which HPV Vaccine Given		Number of Visits at which HPV Vaccine Due	Number of Visits at which HPV Vaccine Given	
Well Visit	36	8	(22%)	216	68	(31%)	3913	1131	(29%)
Sick Visit	38	0	(0%)	253	13	(5%)	3860	85	(2%)

*Note: this is actual data from a randomly selected study clinician

Example #1 – National EHR datasets, getting to "Big Data"

16

What is Big Data Anyway?

"…a broad term for data sets so large or complex that traditional data processing applications are inadequate."

17

Data Provenance

The first law of informatics:

"Data shall only be used for the purpose for which they were collected."

van der Lei J. Use and abuse of computer-stored medical records. Methods of Information in Medicine 1991;30:79–80.

18

The law of medical information: Berg and Goorman

- "The further information has to be able to circulate (i.e. the more diverse contexts it has to be usable in), the more work is required to disentangle the information from the context of its production. The question that then becomes pertinent is; who has to do this work, and who reaps the benefits?"

Int J Med Informatics 1999; 56:51–60

19

Comparative Effectiveness Research *through* Collaborative Electronic Reporting [CER²]: Opportunities and Cautions as Data Gets Big

20

What is CER² & what is its origin?

- CER² is a collaboration of primary care informatics researchers studying pediatric care through EHR and related electronic data
- CER² grew out of a need for large scale 21st century practice-based research
- CER² joins existing EHR-based research networks into an electronic über-network

21

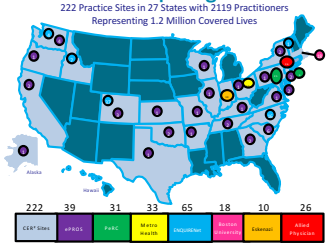
CER² Unique Appeal

- Longitudinal electronic health record (EHR) database
  - Extensive follow-up time 2000-2014
  - Clinical data supplemented by administrative data
- Contains >1.2 million children, diverse practitioners, seen by practitioners in diverse settings from across the United States
- Supported by a unique team combining, health services, informatics and pharmacoepidemiology expertise
- EHR data from CER² can be *combined* with data collected from pediatricians, parents, and children to conduct prospective interventional studies and provide a powerful 21st century research platform

22

CER² Participating States

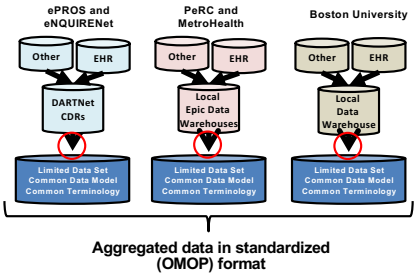
222 Practice Sites in 27 States with 2119 Practitioners
Representing 1.2 Million Covered Lives



Fiks AG, Grundmeier R, Steffes J, Adams W, Kaelber D, Pace W, Wasserman R: "Comparative Effectiveness Research through a Collaborative Electronic Reporting Consortium" Pediatrics 136(1): e215-24, July 2015

23

CER² Model



24

CER² Governance

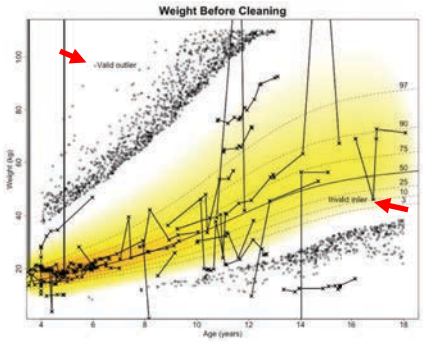
- American Academy of Pediatrics "owns" the data and has data use agreement with all data contributors
- Aggregated HIPAA-limited dataset is stored on secure server at The Children's Hospital of Philadelphia
- CER² partner members contribute and analyze data
- CER² affiliate members can help analyze data
- Non-member investigators can help analyze data
- The CER² team has begun to recruit additional partners, affiliates, and non-member investigators to write grants, contribute data and answer new questions.

25

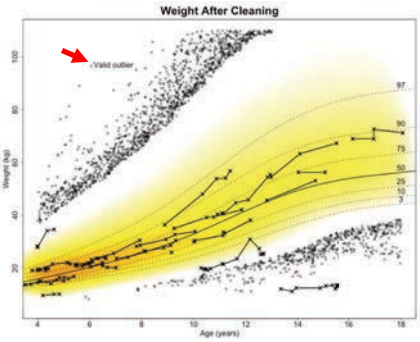
Implausible Growth Data: Data Cleaning Method

- Challenge: CER[2] has growth data on ~1.2 million children
- Errors include:
  - Substitution of metric and English system values
  - Misplaced decimal points
  - Plain old mistakes
- How to sort these out?
- The "Daymont Method"

26

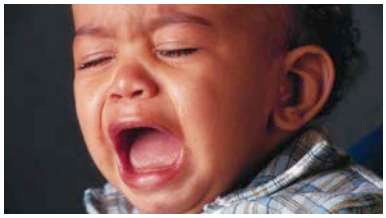


27



28

Missing Data on Race/Ethnicity



29

Missing Data on Race/Ethnicity

Missing data on race/ethnicity is common

- Could an adaptation of a method used previously in adults be helpful in pediatrics?
- We tested "Bayesian Improved Surname Geocoding" which accounts for US census geospatial and surname data in addressing missingness
- *Bottom line: The new method more correctly race/ethnicity than traditional imputation, and reduced bias*

30

Validity of Imputed Race/Ethnicity for Health Disparities Research

Table: Bias of different methods of estimate race/ethnic difference: true values and estimates by method

Race/Ethnicity	True value	All data*	Complete Case analysis	Indicator variable method	Multiple imputation	BISG-enhanced imputation
Continuous Outcome†						
Black	-10	-10.0 [-10.3,-9.71]	-2.59 [-2.89,-2.23]	-0.96 [-1.26,-0.70]	-2.96 [-3.41,-2.49]	-8.05 [-8.31,-7.74]
Hispanic	+10	9.97 [9.52,10.6]	6.88 [6.09,7.47]	8.21 [7.38,8.79]	7.62 [6.86,8.29]	8.30 [7.61,8.88]
Binary Outcome‡						
Black§	0.5	0.50 [0.48,0.52]	1.35 [1.29,1.41]	1.54 [1.48,1.60]	1.30 [1.22,1.39]	0.65 [0.62,0.68]
Hispanic	2.0	1.99 [1.89,2.11]	1.33 [1.26,1.42]	1.48 [1.40,1.58]	1.34 [1.27,1.45]	1.55 [1.45,1.64]

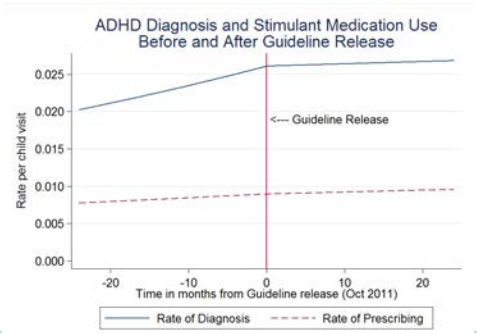
Grundmeier RW, et al., "Imputing missing race/ethnicity in pediatric electronic health records," Health Services Research, 2015 Mar 11. doi: 10.1111/1475-6773.12295. [Epub ahead of print]

31

Evaluating the Impact of New Psychotropic Medication Guidelines

- In 2011, AAP ADHD practice guidelines provided, for the first time, guidance for the diagnosis and treatment of ADHD in preschoolers
- To evaluate the impact on clinical practice, we examined changes in the diagnosis of ADHD and prescription of stimulants to children aged **4 through 5 years old** following guideline publication

32



- Patterns were similar in sensitivity analyses that included all practices.

33

Recap: Big Data

- Data is being repurposed—this has inherent challenges
- Data cleaning is critical
- Once cleaned, pooled data may facilitate the elucidation of national practice patterns and the assessment of rare outcomes
- The CER^2 team welcomes interested investigators to participate

Example #2 – From regional data sharing infrastructure to small scale EHR data projects

35

WWAMI region Practice and Research Network

- 58 clinics across 26 organizations in 5 states
- 18 clinics are members of the Data QUEST data sharing infrastructure to share EHR data for research

36

Data QUEST

WHO: A collaboration between the WPRN and CTSA biomedical informatics program

WHAT: Technology – aligns data repositories across practices

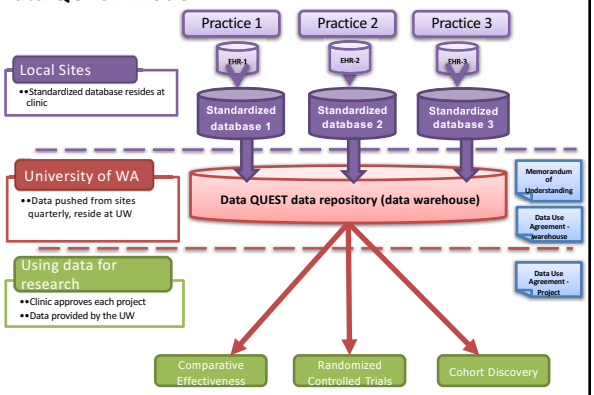
- identified clinical data repositories → at the practices
- de-identified data → in a central warehouse hosted at UW

WHY: Supporting research and quality improvement

BEYOND OUR REGION: Aligning with national efforts

- Practice-Based Research Networks – DARTNet Institute
- Patient Centered Outcomes Research Institute (PCORI) PCORNet
- NIH Collaboratory Distributed Research Network

Data QUEST model



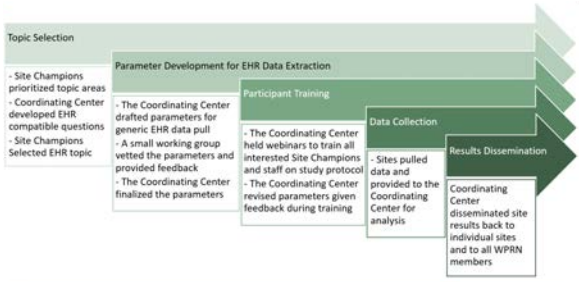
38

Data QUEST

- Strengths
  - Remarkable tool that has stimulated PBRN collaborations on:
    - 13 grants from small pilot studies to large scale pragmatic clinical trials
    - 17 manuscripts on the science of data sharing
- Challenges
  - Costs are significant for
    - Coordinating center personnel to manage and conduct operations, monitor and maintain data quality
    - Installation costs ~$7,000-$10,000 per site, maintenance is $2,500 annually. Data extraction costs ~$3800 for the first site, $750 each additional.
  - Navigating governance structures in clinical organizations can be formidable.
  - Clinical organizations may change EHRs or make upgrades that could derail the infrastructure.
  - Variation in EHR structure and how EHRs are used is substantial.
  - Not all WPRN sites can participate in Data QUEST.

39

Leveling the Playing Field: Using a Small Scale EHR Data Project to Increase Quality and Research Capacity in the WPRN



40

WPRN EHR project query data table

See definitions for bolded terms

Measure	Result
1) Denominator: Total number of **adult patients** with a visit to the clinic during the **study period**	
For Group A Sleep Medications	
2) Total patients: Number of **adult patients** with a visit to the clinic during the **study period** who had a **prescription** (during the study period) for at least 1 Group A sleep medication	
3) Male patients: Number of **adult male patients** with a visit to the clinic during the **study period** who had a **prescription** (during the study period) for at least 1 Group A sleep medication	
4) Female patients: Number of **adult male patients** with a visit to the clinic during the **study period** who had a **prescription** (during the study period) for at least 1 Group A sleep medication	
For Group B Sleep Medications	
5) Total patients: Number of **adult patients** with a visit to the clinic during the **study period** who had a **prescription** (during the study period) for at least 1 Group B sleep medication	
6) Male patients: Number of **adult male patients** with a visit to the clinic during the **study period** who had a **prescription** (during the study period) for at least 1 Group B sleep medication	
7) Female patients: Number of **adult male patients** with a visit to the clinic during the **study period** who had a **prescription** (during the study period) for at least 1 Group B sleep medication	
For Group C Sleep Medications	
8) Total patients: Number of **adult patients** with a visit to the clinic during the **study period** who had a **prescription** (during the study period) for at least 1 Group C sleep medication	
9) Male patients: Number of **adult male patients** with a visit to the clinic during the **study period** who had a **prescription** (during the study period) for at least 1 Group C sleep medication	
10) Female patients: Number of **adult male patients** with a visit to the clinic during the **study period** who had a **prescription** (during the study period) for at least 1 Group C sleep medication	
For Group D Sleep Medications	
11) Total patients: Number of **adult patients** with a visit to the clinic during the **study period** who had a **prescription** (during the study period) for at least 1 Group D sleep medication	

41

Returning Results to Sites is Critical

- One page study summary sheet
- Presented anonymously to the WPRN site champions at our annual meeting.
- Sites knew who they were – could compare their rates to similar practices

42

Small Scale EHR Data Project

- **Strengths**
  - **More WPRN sites could participate**
  - **Sites were highly engaged in the science, data collection, and results interpretation**
  - **No IRB approvals required!**
- **Challenges**
  - **No process for assessing data quality or validation**
  - **Results can be used for preliminary studies, but no individual data for research**
  - **Limited resources to support follow-on quality improvement efforts**

43

Lessons Learned

- EHR-based projects are a powerful tool for engaging PBRN sites in research
  - Small scale EHR projects
    - Highly collaborative, responsive to PBRN site interests
    - Demonstrated site capabilities for EHR queries to sites themselves and to the PBRN Coordinating Center
  - Large scale data sharing
    - Requires a greater commitment from sites – must provide information on data origins (provenance)
    - Engages them as collaborators with academic investigators on pragmatic clinical trials, implementation and dissemination research

Example #3 – Integrating patient reported data into the EHR

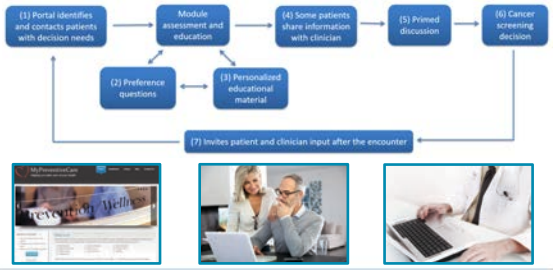
45

Informed Decision-Making Module

- Focus on three cancer screening decisions
- Specific aims:
  - **Aim 1** – to reach patients outside the clinical setting – before clinical encounters – to explore their preferred approach to decisions about cancer screening
  - **Aim 2** – to follow patients into the clinical encounter to study the assistance offered by clinicians, its congruence with the patient's stated preferences, and the effect of website exposure on the conversation and decision outcomes

46

How Information Technology Could Help



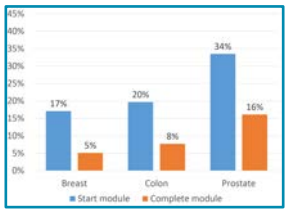
47

Patients Facing Cancer Screening Decisions (n=72,000 portal users)

	Any	Breast	Colon	Prostate
Overall	11,094 (13.5%)	3,615 (4.4%)	6,115 (7.5%)	1,364 (1.7%)
Phase 1 (6 weeks) - Prompt when using	1,010	297	542	171
Phase 2 (14 weeks) - Invite before visit	610	171	354	85
Phase 3 (12 weeks) - Invite outside of visit	9,436	3,220	5,136	1,080

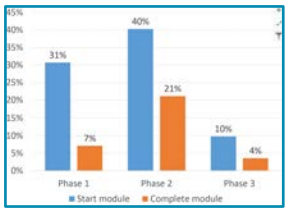
48

Module Use by Cancer Screening and Phase



P ≤ 0.001 (start and complete)

- Breast vs. colon
- Breast vs. prostate
- Colon vs. prostate

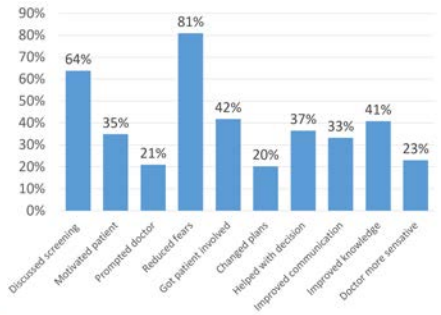


P ≤ 0.001 (start and complete)

- Phase 1 vs. phase 2
- Phase 1 vs. phase 3
- Phase 2 vs. phase 3

49

Impact of the Module on the Visit



50

Getting HIT to Work for this Research Study

Facilitators

- We controlled the design and programming of the patient portal
- Patient and clinician engagement supported the research
- HIT automated a repetitious process that occurred frequently in practice

Barriers

- We had limited control over how the module interacted with the EHR
- More time is needed to cause the culture change needed to redesign care
- Sometimes it is difficult to balance research versus clinical data collection

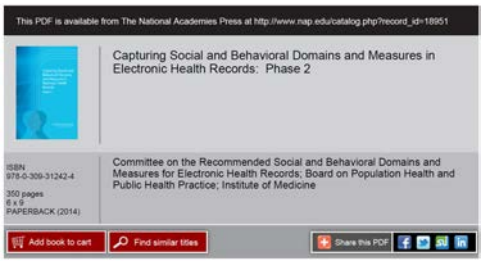
51

The Benefits of Patient Reported Data

- Can be used to better engage patients in their care and wellbeing
- Can extend care outside of traditional office visits
- Can serve as a new data source – the answers to which only patients know

52

The IOM Has Prioritized Capturing Patient Reported Data in Electronic Health Records



53

Thank you for your interest

54

Key References

- Electronic health record functionality needed to better support primary care. *J Am Med Inform Assoc*. 2014; 0:1-8. [PMID 24421225].
- A family medicine health information strategy for achieving the triple aim for U.S. healthcare. *Ann Fam Med.* 2015; 47(8):628-35. [PMID 26382121].
- Capturing social and behavioral domains and measures in electronic health records: phase 2. Available at http://iom.nationalacademies.org/Reports/2014/EHRdomains2.aspx.
- Implementation of a health data-sharing infrastructure across diverse primary care organizations. *J Ambul Care Manage.* 2014;37(2):164-70. PMCID: PMC4065306
- https://dataquest.iths.org