Texas Public Health Committee

Patient Expectations for Health IT:

Control over Health Records Privacy Solutions for HIE

May 11, 2010
Deborah C. Peel, MD

patientprivacyrights

Reality:
massive health data
mining industry

+

ease of re-identifying
health information

2010: Top Fortune 500 Companies health data mining industry

4 General Electric (GE Centricity EHR/HIT systems, *sells clinical data*) revenue 157B

14 McKesson (*sells Rx data*) revenue 107B

18 CVS Caremark (*sells Rx data*) revenue 99B

21 UnitedHealth Group (*sells RX data*) thru

   Ingenix subsidiary) revenue 87B

31 WellPoint (*sells claims/clinical data* via BHI) revenue 65B

http://money.cnn.com/magazines/fortune/fortune500/2010/full_list/

Clinical Data Services

GE Healthcare

About 15,000 MDs (primary care, specialty)

De-identified, standardized data

Warehoused nightly

15 million unique patients, growing at 30% a year

The **de-identified data is collected from members using GE Centricity® Electronic Medical Record**. The database is used by outcomes researchers, members of the pharmaceutical industry, and academic institutions in the hopes of improving clinical care and outcomes throughout the healthcare industry.

The CDS Advantage Patient encounters since 1996; average 3 years data

Expanded demographic information (+ 3 digit zip code), insurance

Pharmaceutical therapeutic class, and brand name

Clinical Data Services
GE Healthcare's Clinical Data Services Business provides **access to de-identified ambulatory electronic medical record data. It is one of the largest anonymized clinical databases in the United States** providing access to real-world longitudinal patient information.

Clinical Data Services
The CDS Advantage

GE Healthcare

Disease Counts in Database

Hypertension 2,284,249

Hyperlipidemia 2,212,629

Depression 1,185,828

Cardiovascular Disease 1,004,214

GERD 984,864

Diabetes 922,169

Asthma 750,963

Osteoarthritis 602,043

COPD 319,310

ADD/ADHD/HKD 188,424

Rheumatoid Arthritis 85,757

Alzheimer's 35,790

Parkinson's 22,017

Note: Data reported as of February 28th, 2010

Codified Medical Problems

Prescriptions/Historical Meds

Patient Allergies, Medical Orders and Events

Vital Signs and Physical Findings

Lab Values

https://www2.gehealthcare.com/portal/site/usen/menuitem.b399d8492e44a6765c09cbd58c829330/?vgnextoid=ae0f4fb9efff5210VgnVCM100000382b3903RCRD&fromChannel=7e0f4fb9efff5210VgnVCM100000382b3903____

GREENWAY

Prime*Research* *part of an* EHR/ Practice Management Suite

"Key Benefits for Physicians"

"Make clinical research participation a revenue source"- ie, doctors sell access to their patients and patients' records for:

clinical trials
Phase II - IV trials
Post marketing surveillance
Patient registries
Peri- and post-approval research
Safety and surveillance monitoring
evidence-based medicine
pharmaceutical research
clinical and financial benchmarking services.

The result – increased practice revenues and access to patient care improvements."

GREENWAY HELPS PHYSICIANS SELL:

- "de-identified" clinical data
- "de-identified" financial data
- population data
- information **on 19 million active patients**
- data on **18 million encounters/yr**
- data on **8 million prescriptions/yr**
- data on **8,000 active providers**
- data on **30 specialties and subspecialties in 48 states**

http://www.greenwaymedical.com/solutionprime-research/

athenahealth

athenahealth strives to be the best at getting doctors paid

By combining award-winning software, services, and a proprietary rules database, we've created a network that is used by **19,500 physicians and medical providers** nationwide.

http://www.athenahealth.com/strategic-alliances/index.phphttp://www.athenahealth.com/strategic-alliances/index.php

Xconomy | Boston
Business + Technology in the Exponential Economy

Athenahealth Paying Dearly to Take on Larger Rivals **by** Ryan McBride 5/6/10

"Athena might be able to halve the amount that physicians pay to use its EHR if they participate in "AthenaCommunity."**

Athena's EHR customers who opt to share their patients' data with other providers **would pay a discounted rate to use Athena's health record software**.

Athena would be able to make money with the patient data by charging, say, a hospital a small fee to access a patient's insurance and medical information from Athena's network.

http://www.xconomy.com/boston/2010/05/06/athenahealth-paying-dearly to-take-on-largerrivals/3/

Healthcare **IT** News

Practice Fusion expands, shows signs of rapid growth

By **Diana Manos, Senior Editor**
12/31/07

Practice Fusion subsidizes its free EMRs by selling de-identified data to insurance groups, clinical researchers and pharmaceutical companies.

[Howard](#) said he does not expect data-sharing will be a concern to physicians who use Practice Fusion's EMRs. ***"Every healthcare vendor is selling data."***

DARTNet

Distributed Ambulatory Research in Therapeutics Network

- extracts "de-identified" Critical Care Record (CCR) from EHRs of 400K patients treated by 500 primary care docs
- patient consent not obtained –research uses physician consent instead
- physicians prompted to obtain specific information *during* patient visits
- 2nd study on Depression needs 2.4 M patients, will add a RHIO

8 DARTNet orgs/EHR vendor

Medical Clinic of North Texas	NextGen®
WellMed Medical Group (TX)	SmartClinic®
Tiena Health (TX)	Allscripts Professional®
Wilmington Health Asso.	Allscripts Professional®
University of Colorado	Allscripts Enterprise®
University of Minnesota	Allscripts Enterprise®
Cranford Family Medicine (AK)	e-MDs®
Family Health Center of Joplin	e-MDs®

http://www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-andreports/ ?pageaction= Displayproduct&productID=151

Patient info available via DARTNet but not through claims data

Medication allergies
Reason for appointment
Family history
Findings (BP, weight, height, etc.)
Social history (alcohol and tobacco use, etc.)
Laboratory orders and results
Prescribed medications
Past medical history
Date of onset of disease
Referrals
Provider-level data
Practice-level data
Data collected/prompted for collection at point of care

2010: Top Fortune 500 Health Care: Pharmacy and Other Services (health data mining industry)

Rank	Company	500 rank	Revenues($ billions)	
1	Medco Health Solutions	#35	59.8	(sells Rx data)
2	HCA (largest US hospital chain)	#77	30	(?? sells hospital and Rx data)
3	Express Scripts	#96	25	(sells Rx data)
4	Quest Diagnostics	#303	7	(sells data/sends data to HIEs)

"transforms millions of test results into valuable information products"

http://www.questdiagnostics.com/brand/careers/index.html#services

| 5 | Omnicare | #347 | 6.3 | (???) |

(leading Rx provider for seniors)"we capture a tremendous amount of data"

..combines data with outcomes algorithm technology

| 6 | Lab Corp. of America | #442 | 4.7 | (sells data??/sends data to HIEs) |

What is BHI® (Blue Health Intelligence)?

share critical health information with employers

premier health intelligence resource in the nation

unmatched detail about healthcare trends and best practices while protecting individual privacy

BHI sets the new standard for healthcare data aggregation, reporting and analysis

Size and Value

1) **longitudinal data on 54 million BCBS members** [used for this purpose without consent]

2) 36 months of historical information

3) reporting not only by MSA, industry and product type, *Diagnosis Related Groups (DRGs)* code, *age group* and *gender* [allows re-identification]

How does BHI ensure the privacy and security of members' healthcare information?

1) **adheres to HIPAA regs** [no consent for use and sale of data] throughout the collection and processing of **company data** [your health information is BCBS' corporate asset]

2) Use a *system-generated identifier,* allowing *longitudinal analysis* [allows re-identification]

3) **fully de-identified in accordance with HIPAA** [17 identifiers removed, still allows re-identification of .04%]

http://www.bcbs.com/innovations/bhi/bhi-faqs-1-12-09.pdf

Health Research Data for the Real World: the MarketScan Data Bases

David M. Adamson, PhD
Stella Chang, MPH
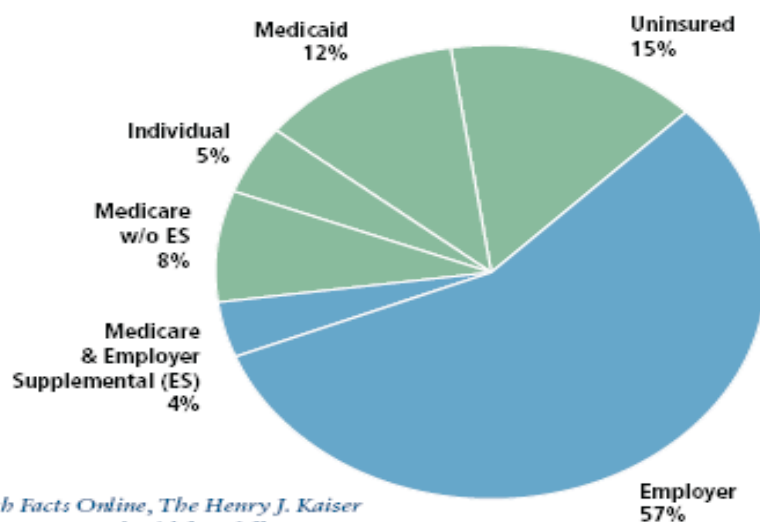Leigh G. Hanson, MS, MBA

Research and Pharmaceutical Division
Thomson Medstat
January 2006

Medicare and Medicaid data is for sale

Figure 1: Population Distribution by Insurance Status — 2002

- Medicaid 12%
- Uninsured 15%
- Individual 5%
- Medicare w/o ES 8%
- Medicare & Employer Supplemental (ES) 4%
- Employer 57%

Source: State Health Facts Online, The Henry J. Kaiser Family Foundation, www.statehealthfacts.kff.org; U.S. residents – 285,007,110. Note: Percentages do not add to 100% because of rounding.

To address the need for better data on privately insured Americans, Thomson Medstat created the MarketScan® data collection. Since its creation, MarketScan has been expanded to include data on Medicare and Medicaid populations as well, making it one of the largest collections of claims-based patient data in the nation. MarketScan data reflect the real world of treatment patterns and costs by tracking millions of patients as they travel through the healthcare system, offering detailed information about all aspects of care. Data from individual patients are integrated from all providers of care, maintaining all healthcare utilization and cost record connections at the patient level.

Personal health information is for sale

Table 1: Sample Data Elements for Commercial and Medicare Databases

Demographic	Medical Information (Inpatient and Outpatient)	Health Plan Features	Financial Information	Drug Information	Enrollment Information
Patient ID	Admission date and type	Coordination of benefits amount	Total payments	Generic product ID	Date of enrollment
Age	Principal diagnosis code	Deductible amount	Net payments	Average wholesale price	Member days
Gender	Discharge status	Copayment amount	Payments to physician	Prescription drug payment	Date of disenrollment
Employment status and classification (hourly, etc.)	Major diagnostic category	Plan type	Payment to hospital	Therapeutic class	
Relationship of patient to beneficiary	Principal procedure code		Payments—total admission	Days supplied	
Geographic location (state, ZIP Code)	Secondary diagnosis codes (up to 14)			National drug code	
Industry	Secondary procedure codes (up to 14)			Refill number	
	DRG			Therapeutic group	
	Length of stay				
	Place of service				
	Provider ID				
	Quantity of services				

Businessweek July 23, 2008: *"They Know What's in Your Medicine Cabinet, How insurance companies dig up applicants' prescriptions—and use them to deny coverage"* http://www.businessweek.com/magazine/content/08_31/b4094000643943.htm?chan=magazine+channel_in+depth

DATA ON DEMAND | Two companies dominate the field of selling prescription information to insurance companies:

	MEDPOINT	INTELLISCRIPT
Parent	UnitedHealth Group's Ingenix	Milliman
Location	Eden Prairie, Minn.	Brookfield, Wis.
History	UnitedHealth acquired MedPoint in 2002 from a small, Utah-based health-technology company, Nex2	Milliman, a Seattle consulting firm, acquired IntelRx and its IntelliScript product in 2005
Fine print	Delivers five-year history of drug purchases, dosages, refills, and possible medical conditions	Similarly provides five-year purchase history, which includes information on pharmacies and treating physicians
Pitch to insurers	"Identify high-risk individuals, reduce costs, lower loss ratios, and increase revenue"	"Clients report financial returns of 5:1, 10:1, even 20:1"

Data: MedPoint and IntelliScript

Anonymous data *isn't*

"... a common practice is for organizations to release and receive person specific data with all explicit identifiers, such as name, address and telephone number, removed ....because the resulting data look anonymous.

However,... the remaining data can be used to re-identify individuals by linking or matching the data to other data .."*

Latanya Sweeney, PhD, Director, Laboratory for International Data Privacy, School of Computer Science, Carnegie Mellon University

*k-anonymity: a model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems,* 10 (5), 2002; 557-570.

Anonymous data *isn't*

"We must respond to the surprising failure of anonymization"

Anonymization , [is] the name for techniques for protecting the privacy of individuals in large databases by deleting information like names and social security numbers
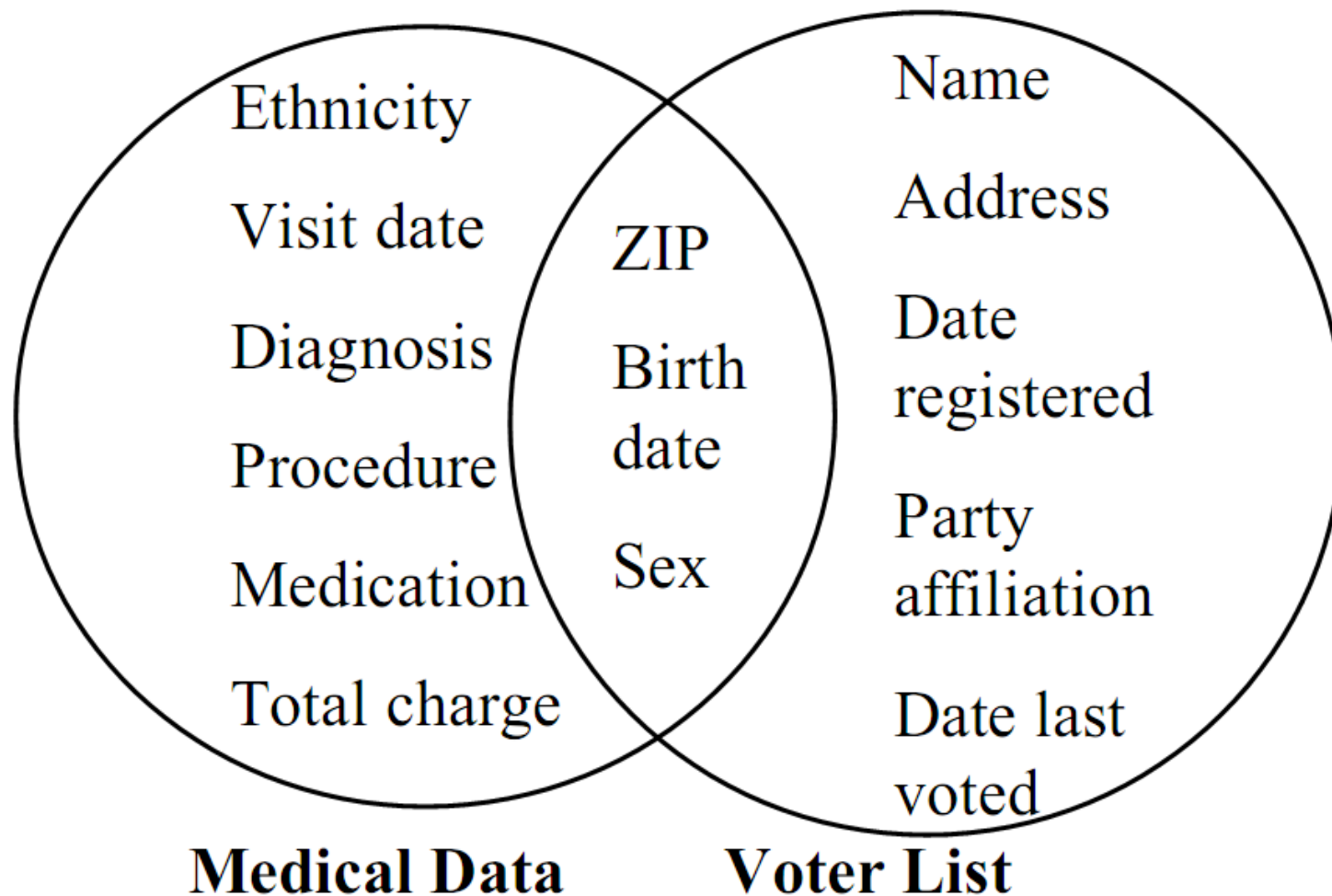
"[S]cientists have demonstrated they can often "reidentify" or"deanonymize" individuals hidden in anonymized data with astonishing ease"*

Paul Ohm, Associate Professor, University of Colorado Law School

* Broken promises of Privacy: Responding to the Surprising Failure of Anonymization, VER. 0.99
SSRN: 8/14/2009:
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1450006&rec=1&srcabs=1446862

Linking to re-identify data

Medical Data circle: Ethnicity, Visit date, Diagnosis, Procedure, Medication, Total charge

Intersection: ZIP, Birth date, Sex

Voter List circle: Name, Address, Date registered, Party affiliation, Date last voted

Medical Data   **Voter List**

L. Sweeney. Weaving technology and policy together to maintain confidentiality. *Journal of Law, Medicine and Ethics*. 1997, 25:98-110.