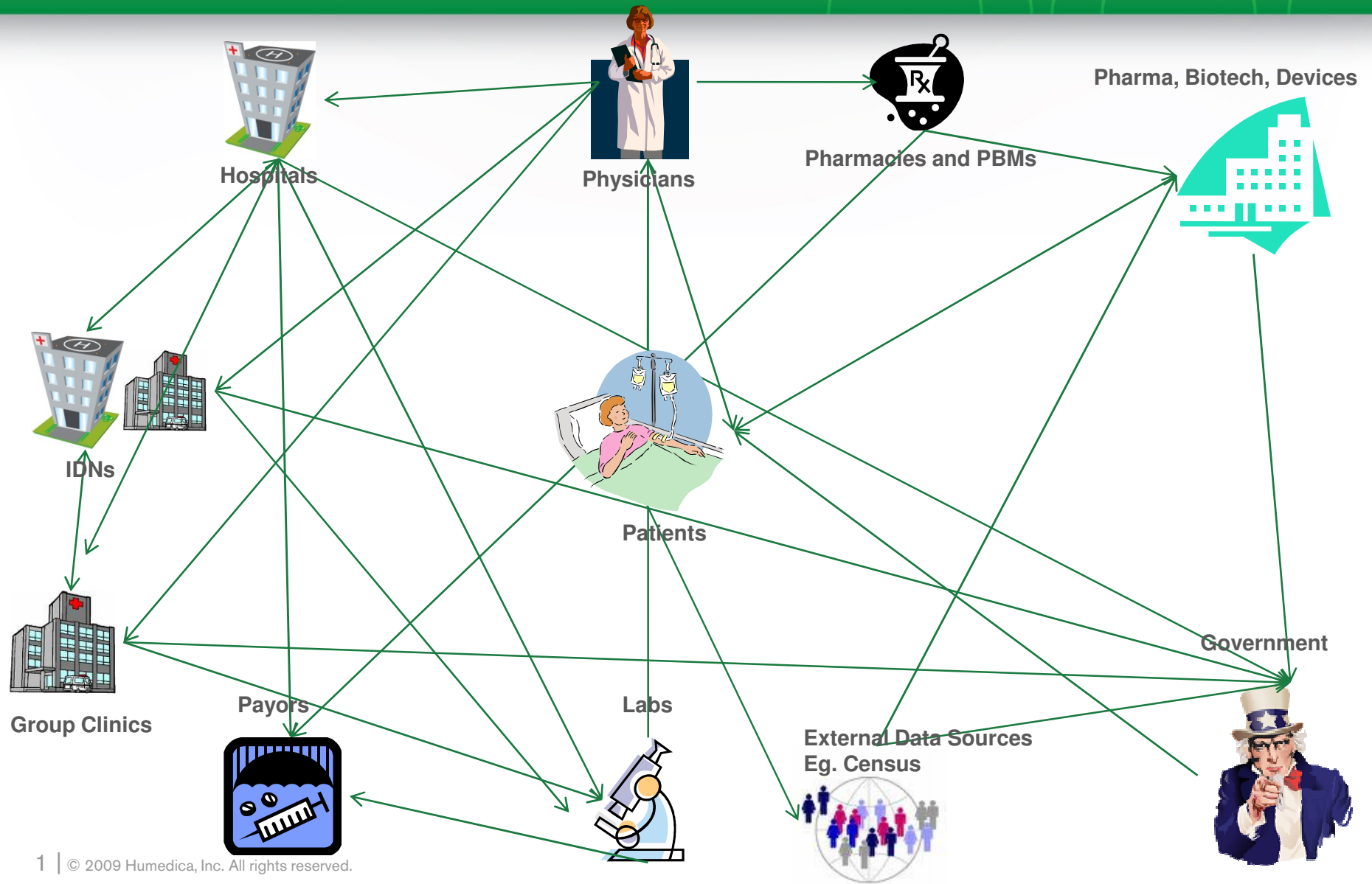


# Healthcare Informatics: Creating Value and Defining Challenges

Paul Bleicher, MD, PhD  
Chief Medical Officer  
Humedica  
Boston MA



# The Information Environment: Healthcare and Life Sciences

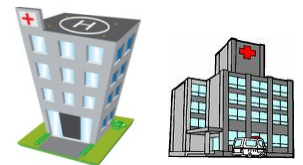


# Who Wants Data, and Why?

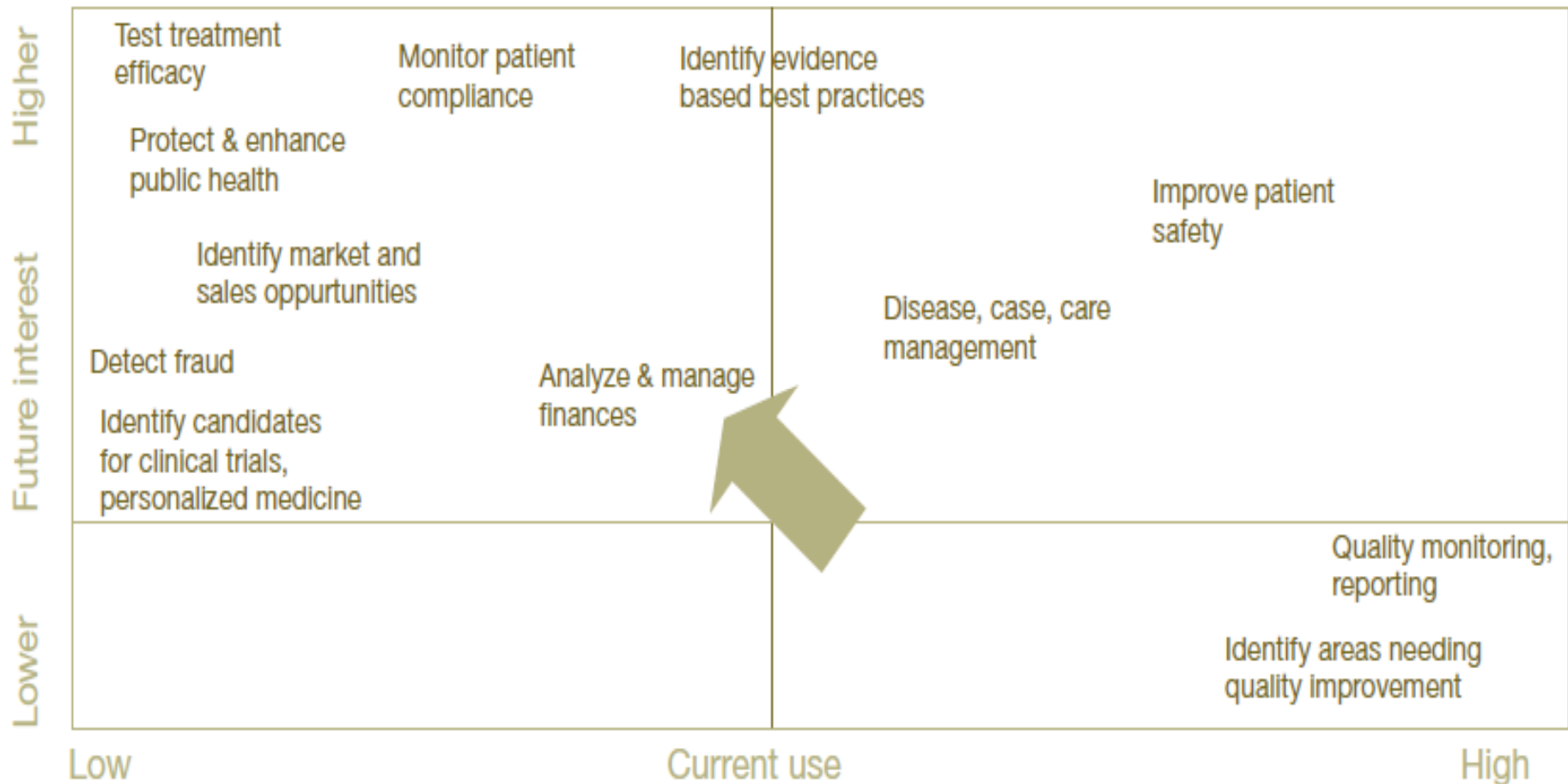
- Pharmaceutical, biotechnology, medical device companies
- Hospital and ambulatory provider organizations, and their physicians
- Government
- Financial services
- Payors
- Consumers
- Employers

# Health Care Organizational Use of Data

- Quality Management
  - Outcomes
  - Staffing and Resources
  - Physician and care team accountability
  - Accreditation and Pay for Performance
    - JCAHO, PQRI, HEDIS, etc.
  - Compliance with care maps, order sets, etc.
- Patient Safety
  - Risk profiles
  - Medication and procedural errors
  - Sentinel events
- Resource and Cost Analysis
- Research and Hypothesis Generation



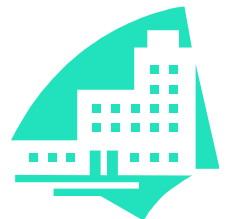
# Healthcare Needs for Electronic Data



Source: PricewaterhouseCoopers survey.

# Pharmaceutical/Biotech Uses for De-identified Data

- Clinical research
  - Observational research – natural history, risk factors, epidemiology,
  - Registries – tumors, devices, diseases
  - Feasibility and planning
  - Recruitment
  - Treatment effectiveness, pharmacoeconomics, pharmacogenomics
- Pharmacovigilance
  - Tracking adverse events
- Market research
  - Supporting formulary decisions / pharmacoeconomics
  - Hypothesis generation
  - Understanding physician prescribing behavior
  - Understanding consumer prescribing behavior
- Sales Tools
  - Targeting physicians, systems
  - Sales incentive modeling

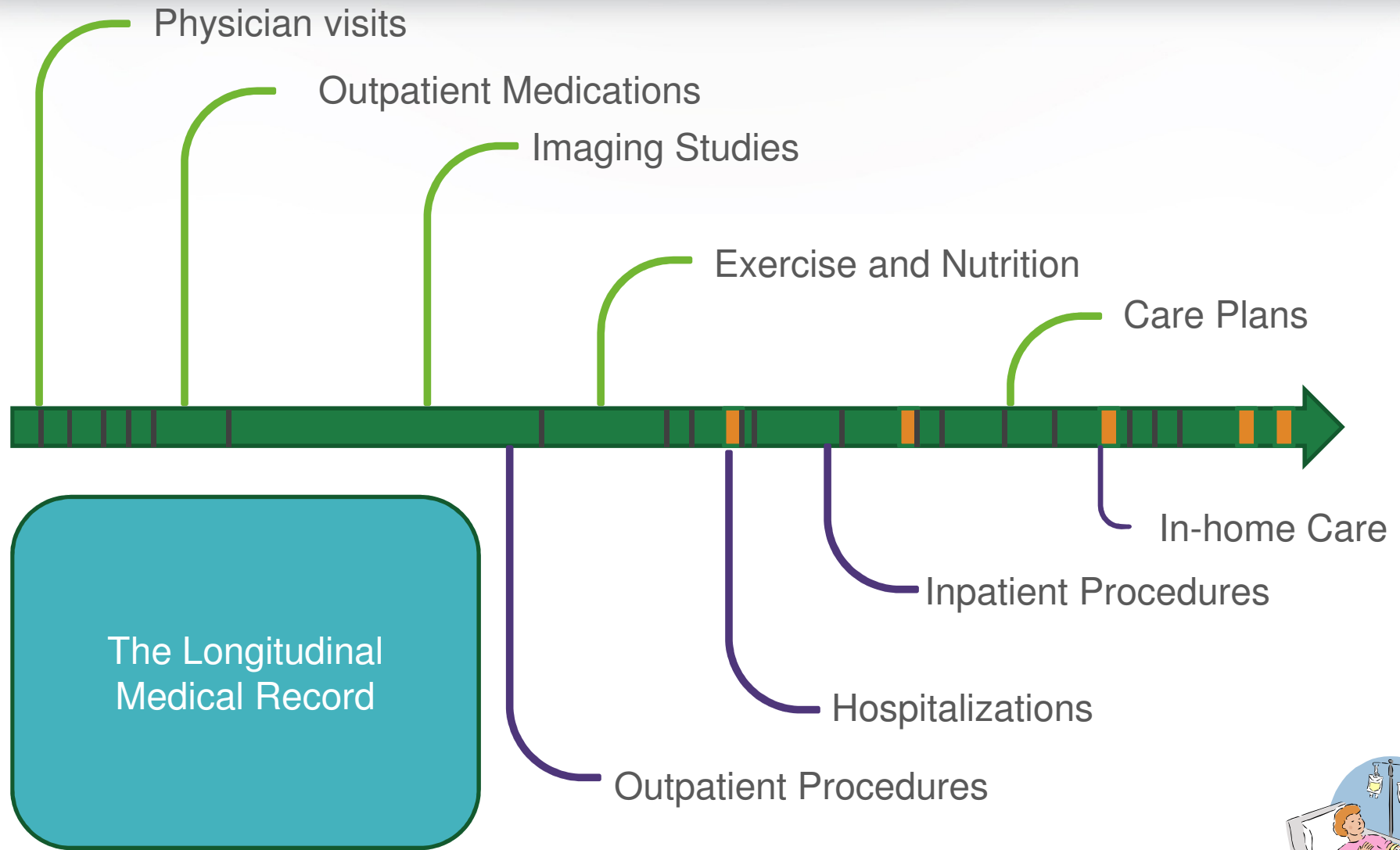


# Government Uses for De-identified Data

- Framework for reimbursement
  - Pay for performance
- Protecting public health
  - Many similar clinical research uses
  - Disease statistics, outcomes
  - Disease surveillance, bioterrorism surveillance
  - Drug approval and pharmacovigilance
- Quality and clinical effectiveness
  - Benchmarking and national standards
  - Comparative effectiveness research



# The REAL "Holy" Grail

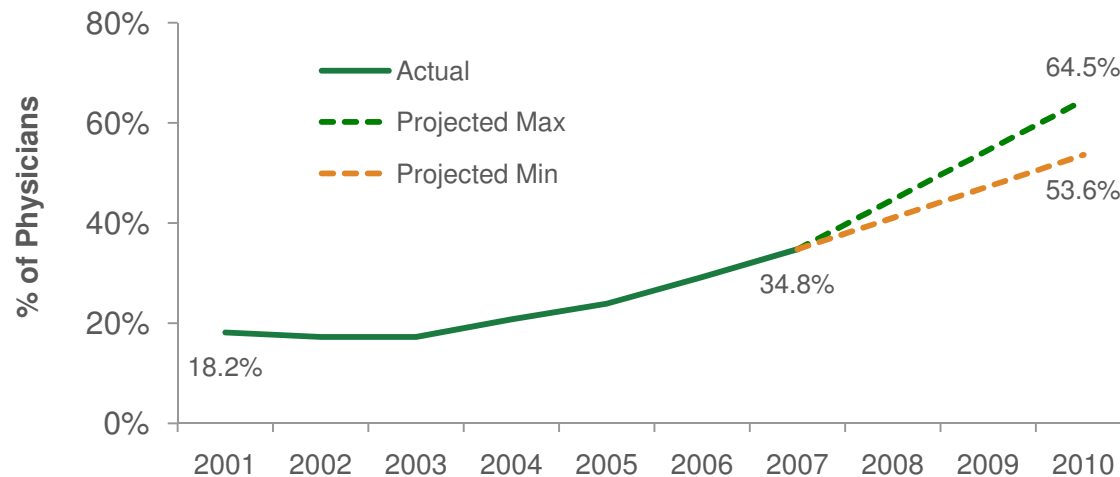




# Now EHR Adoption is Better Than Assumed

- The rate of EMR adoption has been increasing significantly over the past decade
  - Fueled by government incentives courtesy of American Recovery & Reinvestment Act (ARRA)
  - On average, government provides **\$44,000 per physician** to spur adoption

% of office-based physicians using any EMR system and projected 2010 use\*



Adoption of any EMR system reached 74.3% among offices with more than 11 physicians.\*\*

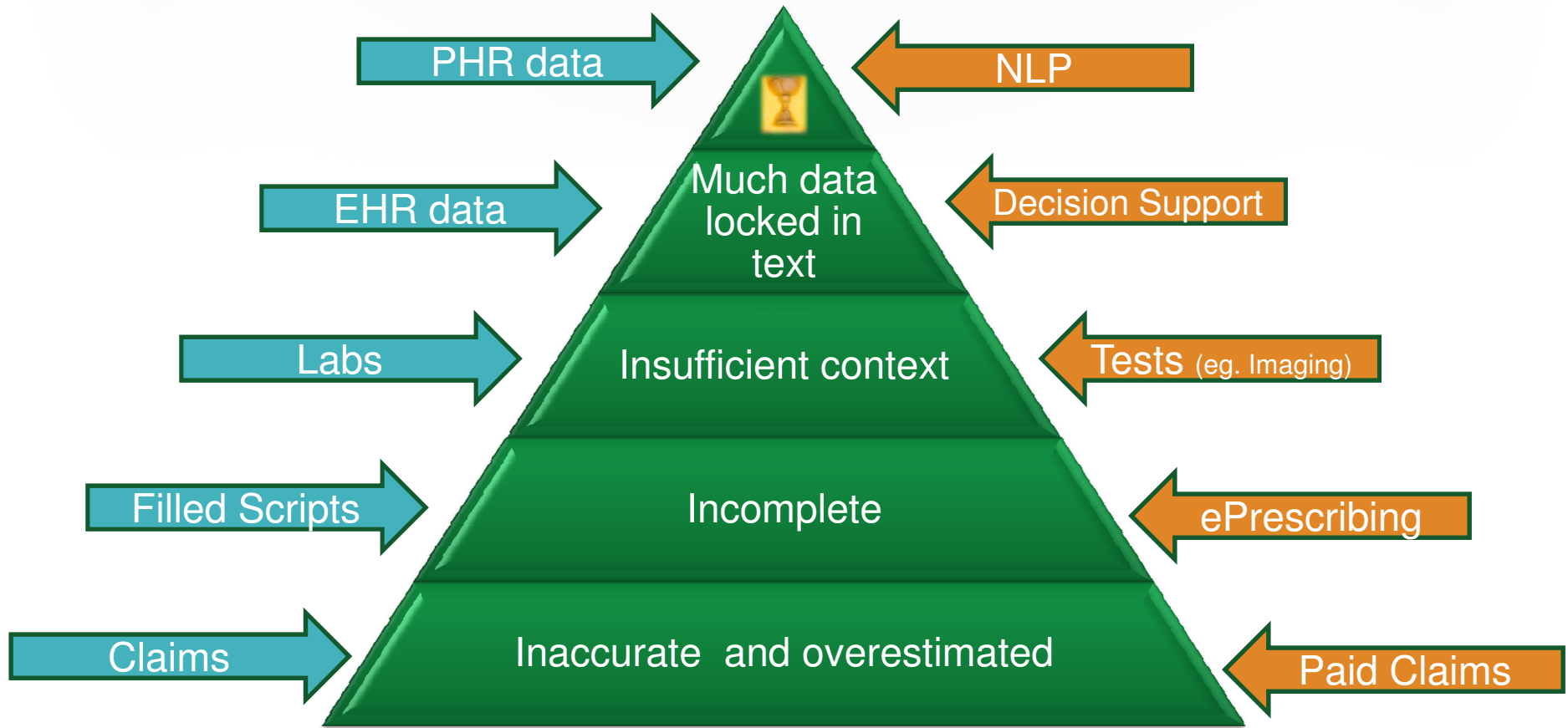
\* Source: CDC/ NCHS, National Ambulatory Medical Care Survey

\*\*Source: Electronic Medical Record Use by Office-based Physicians and Their Practices: United States, 2007

# However, EHR Market is still a Hurdle to Secondary Data

- Hundreds of vendors and platforms
- Fragmented market with different needs/solutions for each
  - Solo practices, group practices, hospital systems, independent delivery networks (hospital c/w group practices)
- Many legacy systems use older technology
  - Older languages, hierarchical databases
- Much of the data isn't structured for statistical analysis
  - Images of transcribed dictations, etc.
- In many systems newer technology (XML, web services, etc.) is very low priority for implementation
- HL-7 may be available, but isn't used in many implementations
- Local customization makes data normalization very difficult

# The Pyramid of Value



# Healthcare Data Warehousing: Areas of Concern



# Danger in Data Mining



Give me a lever long enough and a fulcrum on which to place it, and I shall move the world.

--- Archimedes

- Sample of 5M people
  - 80% power to detect a statistical difference between
    - 19.9% and 20% readmission rate!
    - (Almost) anything can be proven with a large enough dataset
    - Need to distinguish “meaningful” from significant
  - Make enough comparisons and something will be “statistically” significant
    - 1 of any 20 random comparisons will, by definition be “significant” with an  $\alpha = 0.05$

# Privacy and Security is Essential

## Extortion Plot Threatens to Divulge Millions of Patients' Prescriptions

By David Kravets  November 06, 2008 | 6:48:54 PM Categories: [Breaches](#)

A St. Louis company managing prescription benefits of 50 million people said Thursday it called the FBI to investigate an extortion plot threatening to expose personal information, including prescriptions, of millions of its clients.



Express Scripts said it has received an anonymous letter containing the names of some 75 clients that includes dates of birth, Social Security numbers and their prescriptions. The letter threatens to expose millions of patient records if Express Scripts does not pay an undisclosed amount of money.

"We are cooperating with the FBI and are committed to doing what we can to protect our members' personal information and to track down the person or persons responsible for this criminal act," George Paz, the company's president, said in a [statement](#).

Paz added that, "as security experts know, no data system is completely invulnerable."

Of course, this needs to be balanced with...

## **Mass. General paperwork for 66 patients lost on Red Line train**

**Personal, billing data are missing**

**By Milton J. Valencia, Globe Staff | March 24, 2009**

And...

- One in five Americans report that they or a family member have experienced a medical error
- Conservatively, medical errors are the 8<sup>th</sup> leading cause of death in the US with 44-98,000 deaths per annum
- 195,000 people in the USA die each year because of potentially preventable, in-hospital medical complications

# Healthcare Data Analytics





# Principles of Secondary Health Data Analytics

- Patient focus for all data activities
  - All uses should benefit patients
  - Minimal disclosure of data to meet need
  - Never allow re-identification of patients
- Data uses must be transparent
  - overseen by honest brokers or stewards
  - everyone in the process is a data steward
- Data must be collected **THROUGH** the process of care, not in addition to it
- Data analytics for analysis of outcomes, value and comparative effectiveness must be the goal. This will require:
  - new data architectures
  - new expertise for experts in data visualization and predictive analytics
  - new training for others in understanding the output of these efforts

# Extracted data from EHR systems to develop a complete view of patient care

## De-identified Patient Records

**EHR\***

### Patient Demographics

- Geography
- Date of birth (Age)
- Gender
- Health system MPI ID
- Provider
- Primary facility ID
- Race
- Health System
- Education level
- Income level
- ...

### Patient History

- Physical examination
- Problem lists
- BMI
- Smoking status
- Co-morbidities
- ...

### Diagnosis

- ICD9 code
- ...

### Physician Notes

- Chief complaint
- History of illness
- Medical history
- Family history
- Medication history

### Lab / Radiology

- Test Order / Result
- Regular / Specialty
- Test order name / code / date
- HTS code
- Result
- Reference high / low / range
- Specimen collected
- Numeric / non-numeric result
- ...

### Procedure / Treatment

- Procedure code / name / date
- CPT code
- HTS code
- HCPCS code
- Medical devices used
- ...

**Prescribing**

### Medications

- Prescribed
- Dispensed
- Med name / code
- Prescribing provider ID
- GPI, HTS, NDC
- Cost
- Prescription date, dose, route of admin
- Quantity per fill
- Sig (directions)
- ...

**Practice Management**

**Claims**

### Insurance

- Medical / Pharmacy benefit
- Financial class
- Insurance order/ plan
- Insurance policy number
- Formulary status
- ...

### Claims

- Local Claim Code
- BETOS code
- Claim charge
- Claim CPT, HCPCS, ICD9 code
- Post / service date
- Total / Work RVU\*\*
- Work RVU
- (Mal)-Practice Expense RVU
- ...

### Provider / Facility

- Specialty
- NPI
- Payer mix
- Care area
- Appointments / scheduling
- Date of birth
- Gender
- Years of practice
- ...

\* EHR= Electronic Health Record, including contributing reporting systems \*\* RVU = Relative Value Unit

# Health Data Integration Framework

## Program Stewardship

## Health Intelligence Solutions

Clinical Research  
and Optimization

Patient Safety

Health Outcomes  
and Economics

Health  
Management

Market Intelligence

## Business Intelligence Core

Optimization

Predictive  
Analytics

OLAP

Process  
Metrics

Reports and  
Dashboards

Queries

## Data Architecture

### Longitudinal Data Structure

Standards

Ontologies

Metadata

Statistical Normalization

Mapping

Structural Normalization

NLP

Data Warehouse

## Data Sources

EHR/EMR

Claims

Patient PHR

Prescription

Labs and Tests

Telemedicine Devices

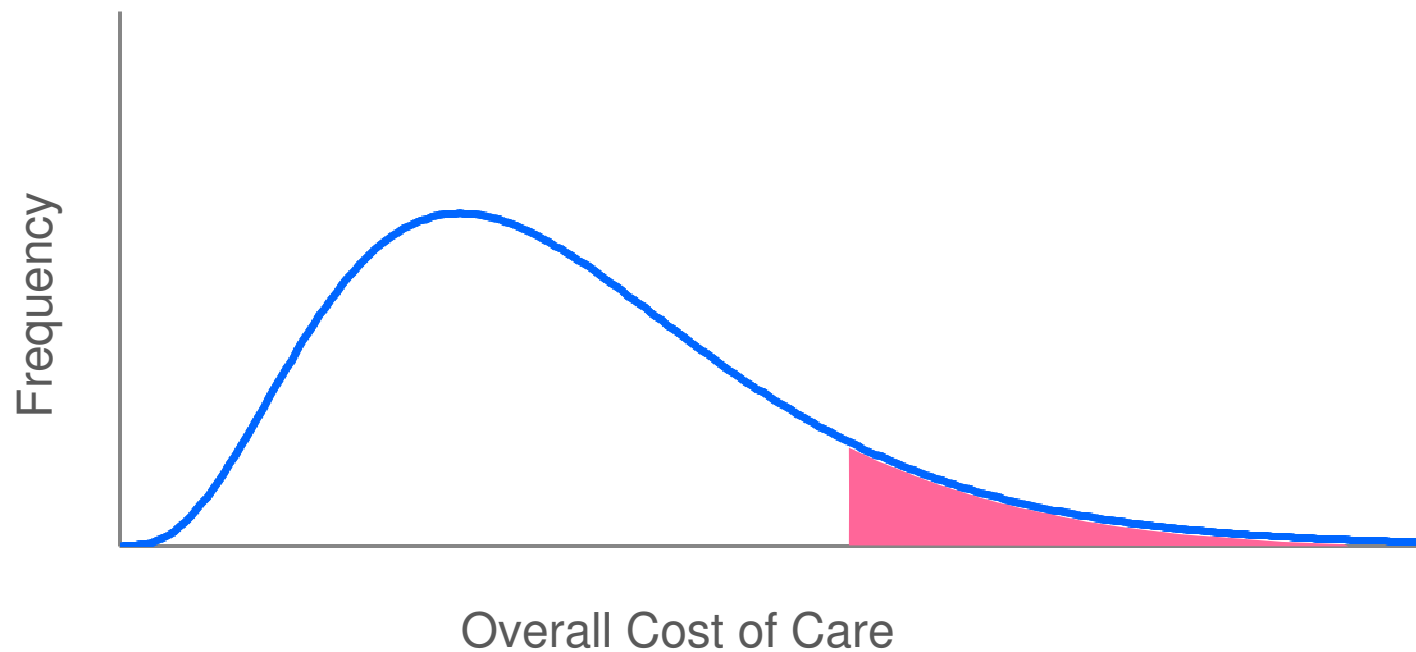
Clinical Trials

Registries

# ACO Target Improvement: Two Components

## 1. Exception management

- a. Identify potential physician or patient “outliers”
  - Clinical intuition
  - Population analytics—patterns of risk/cost
- b. Enables individualized patient attention



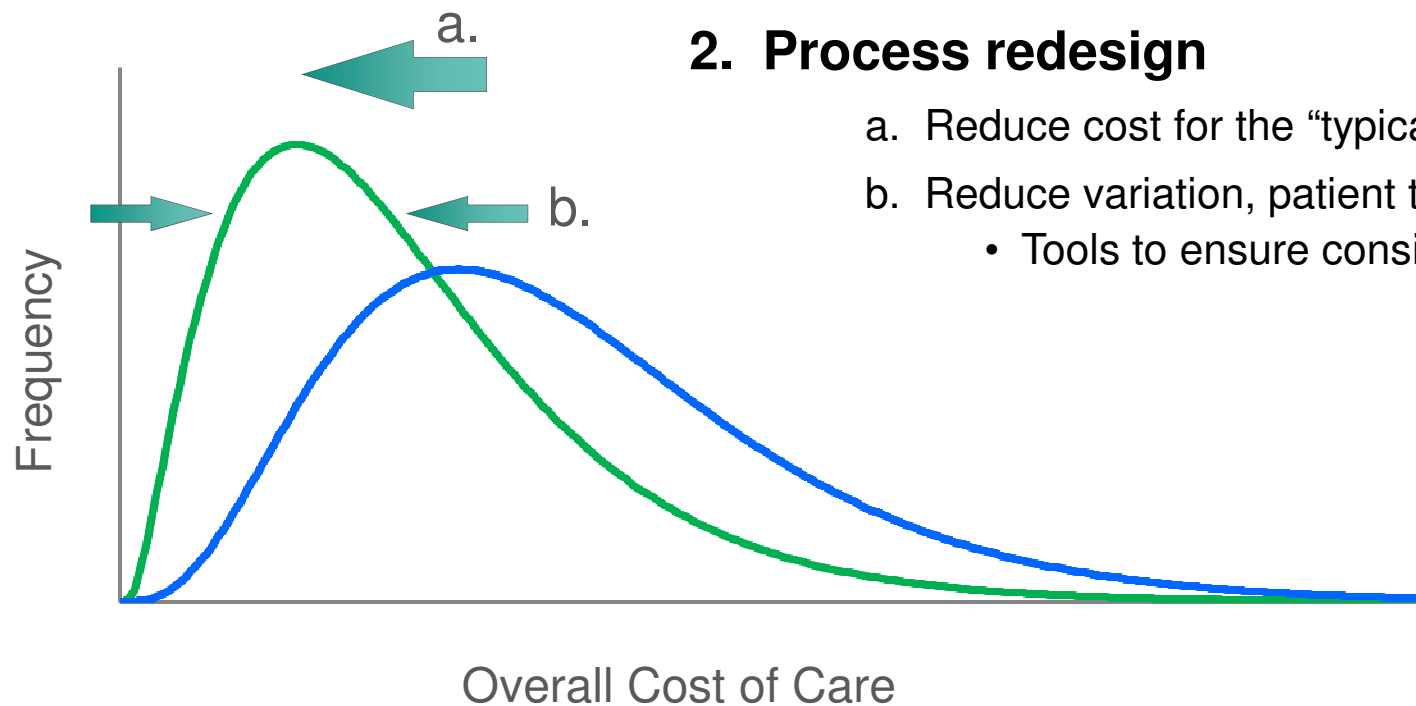
# Improvement: Two Components

## 1. Exception management

- a. Identify potential physician or patient “outliers”
  - Clinical intuition
  - Population analytics—patterns of risk/cost
- b. Enables individualized patient attention

## 2. Process redesign

- a. Reduce cost for the “typical” patient
- b. Reduce variation, patient to patient
  - Tools to ensure consistent execution

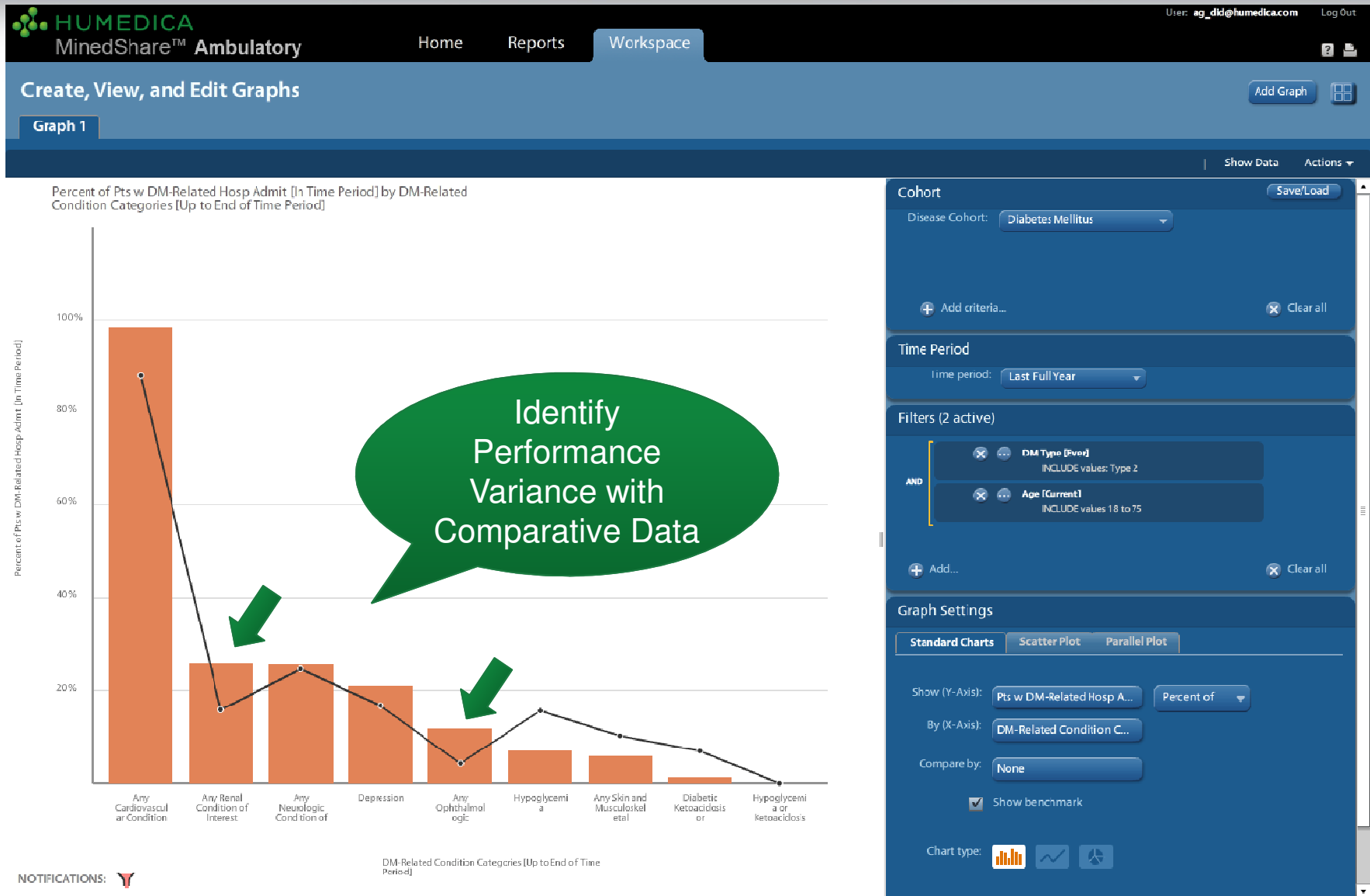


# Preventable Hospitalizations – Key Focus

- **\$30.8 Billion – Preventable admissions**
  - \$1 in \$10 hospital dollars spent could be prevented by better ambulatory care and/or patient self-care
- **Diabetes:**

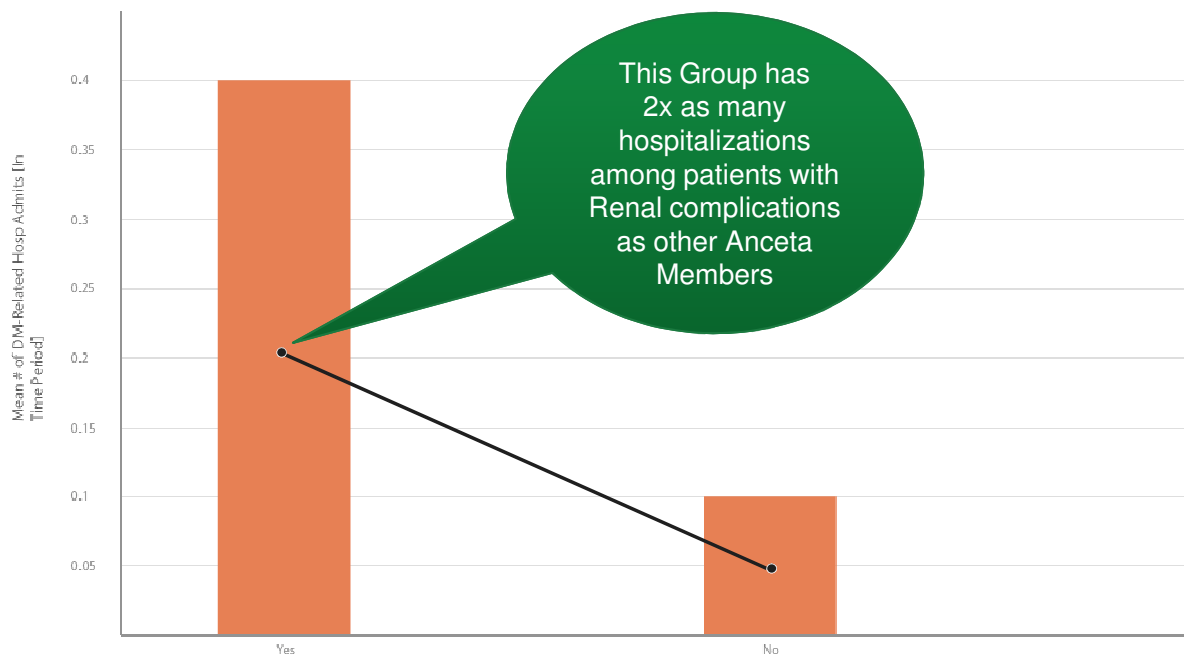
Clinical Condition (HCUP data, 2006)	Total Number of Hospital Admissions (in 000)	Number of Admissions per 100,000 Population	Total Hospital Costs (in 000,000)
Uncontrolled DM without Complications	49	22	\$227
Short Term Complications: Ketoacidosis, Hyperosmolarity, Coma	133	59	\$904
Long Term Complications: <b>Renal</b> , Ophthalmologic, Neurological, Circulatory	295	131	<b>\$2,990</b>
DM-related Lower Extremity Amputations	83	37	\$1,636

# Quickly Identify Opportunities Within Your Population



# Specific Clinical Comparisons Available Nowhere Else

Mean # of DM-Related Hosp Admits [In Time Period] by Pts w Diabetes with Renal Complications [Up to End of Time Period]



NOTIFICATIONS:

Pts w Diabetes with Renal Complications (Up to End of Time Period)

Data Patient List Graph Notes

Pts w Diabetes with Renal Complications (Up to End of Time Period)		Mean	Std.Dev	Range	Num.
Yes		0.4	1.1	0 to 17	1737
No		0.1	0.5	0 to 16	15980

Chart settings are up-to-date (9/16/2010 15:37): Refresh

**Cohort** Save/Load

All criteria are applied relative to the Humedica Diabetes cohort.

+ Add criteria... X Clear all

**Time Period**

Time period: Specific Date Range

Start: 09/01/2009 End: 09/16/2010

**Filters (1 active)**

DM Type [Ever]  
Include values: Type 2

+ Add... X Clear all

**Graph Controls**

Standard Graphs Scatter Plot Parallel Plot

View: # of DM-Related Hosp ... Mean

Displayed by: Pts w Diabetes with Re...

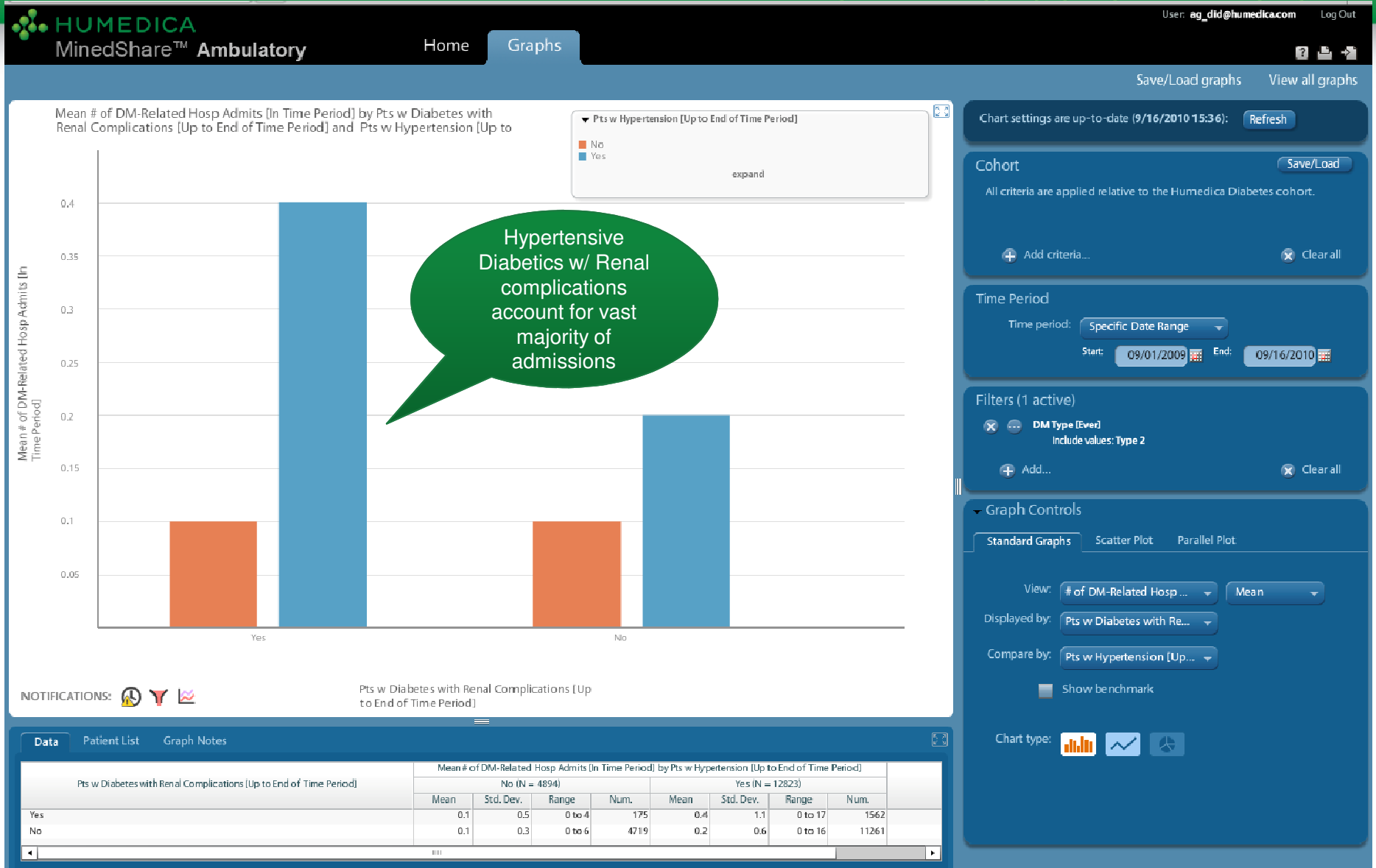
Compare by: None

Show benchmark

Chart type:



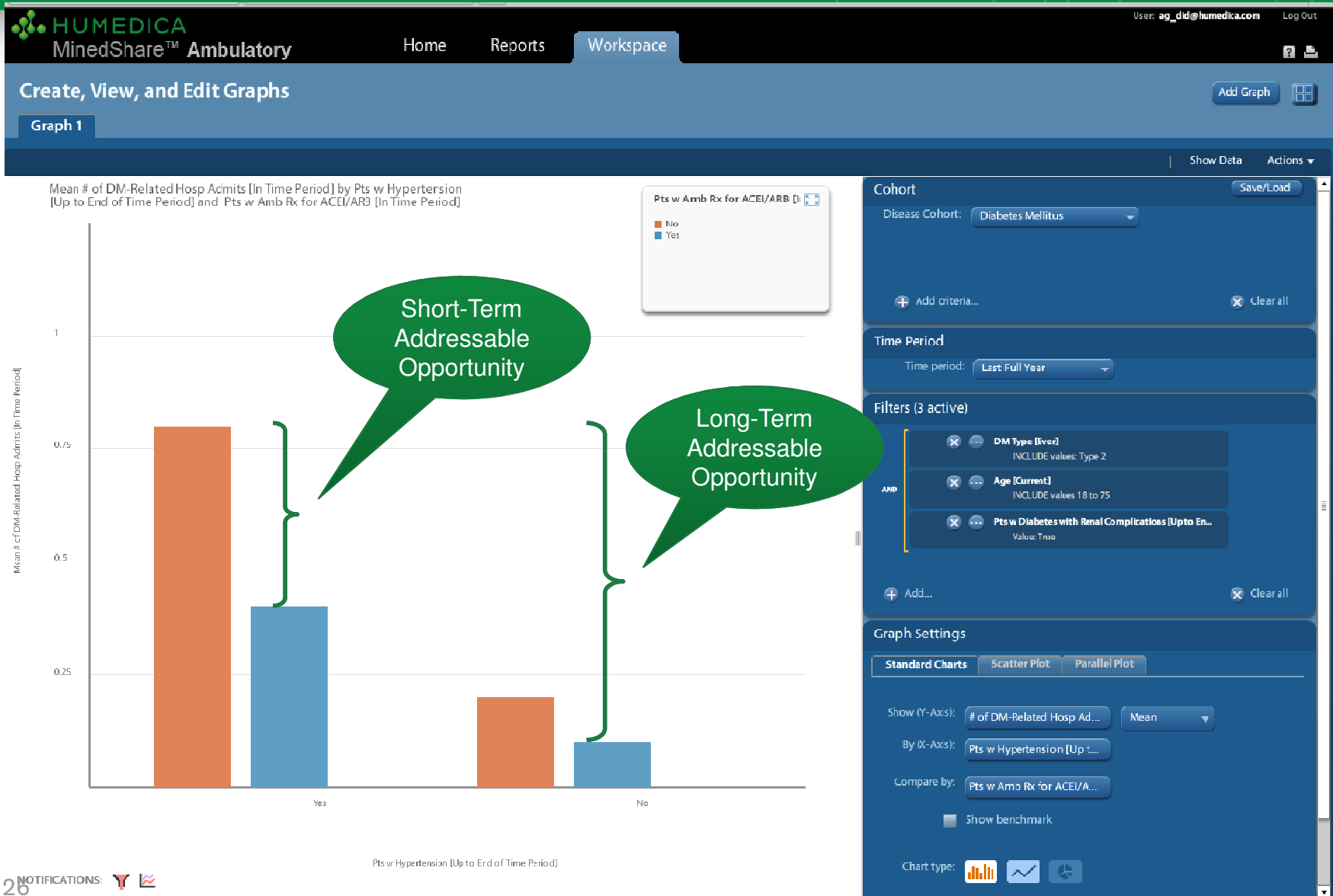
# Explore & Identify Clinical Drivers



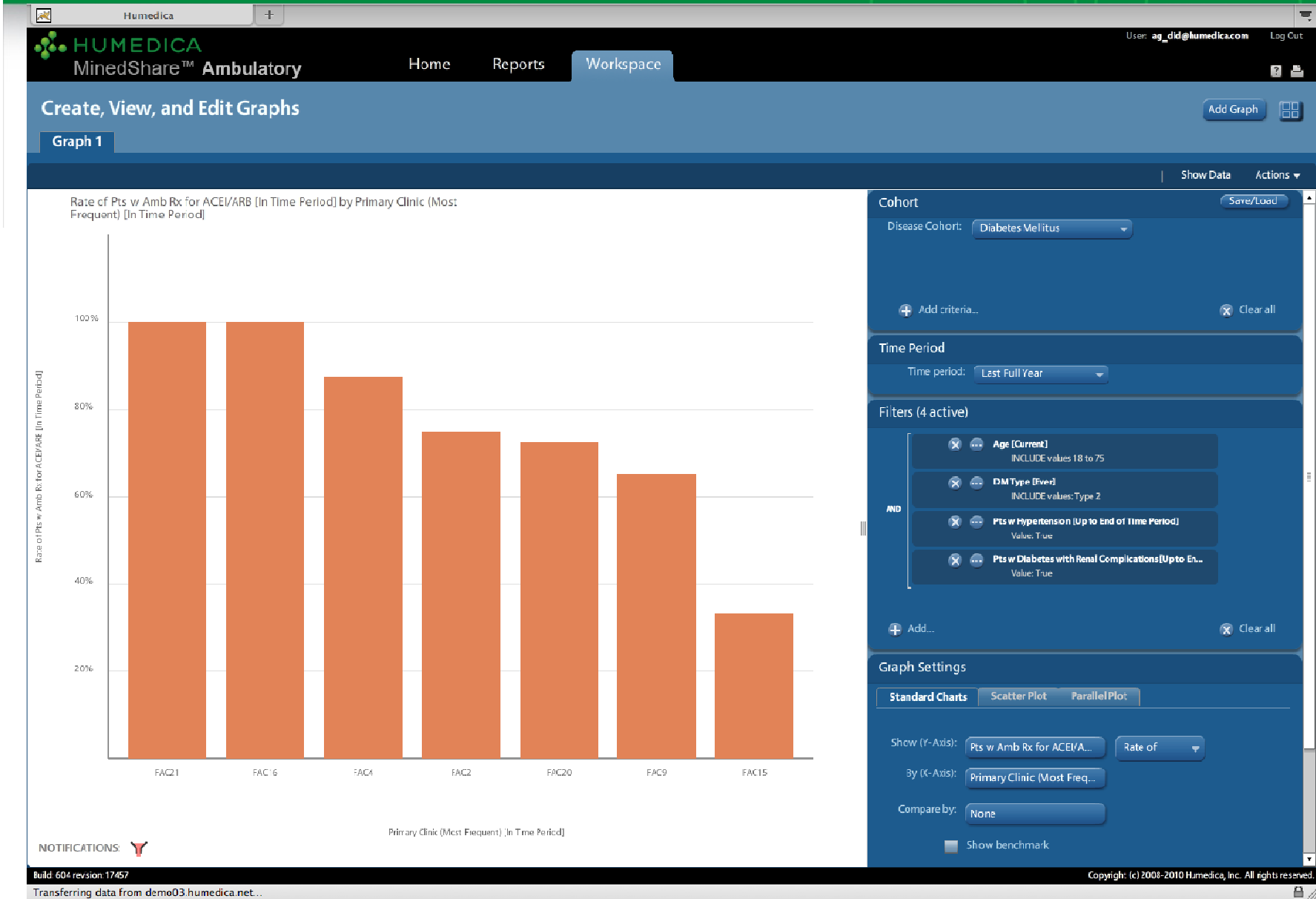
# The Evidence

- In patients with diabetes and renal disease, lowering blood pressure and the levels of urinary albumin are effective in reducing the risk of end-stage renal disease as well as that of myocardial infarction, heart failure, and stroke.
- ACE inhibitors and ARBs appear to be the most effective anti-hypertensive agents.
  - *New England Journal of Medicine* 351;19 (November 4, 2004)

# Quantify the Addressable Clinical Opportunities



# Identify Variance by Clinic



# ...Or by Physician

Humedica MinedShare™ Ambulatory Home Reports Workspace

Create, View, and Edit Graphs

Graph 1

Rate of Pts w Amb Rx for ACEI/ARB [In Time Period] by PCP [Current]

NOTIFICATIONS:

Cohort: Diabetes Mellitus

Time Period: Last Full Year

Filters (5 active):

- Age [Current] (INCLUDE values 18 to 75)
- DM Type [Ever] (INCLUDE values: Type 2)
- Pts w Hypertension [Up to End of Time Period] (Value: True)
- Pts w Diabetes with Renal Complications [Up to End of Time Period] (Value: True)
- # of DM Pts Assigned to Pt's PCP [Current] [In Time Period] (INCLUDE values 300 to 1111)

Graph Settings:

Standard Charts | Scatter Plot | Parallel Plot

Show (Y-Axis): Pts w Amb Rx for ACEI/ARB... Rate of

By (X-Axis): PCP [Current]

Compare by: None

PCP [Current]	Graph ID	Rate (%)	Num.	Denom.
MARSHALL, WYATT MD	A3	86.6	29	36
WEBSTER, MALCOLM MD	A19	72.2	13	18
REASLEY, FRANCES MD	A23	50.0	9	18
SPENCE, CAMILLE MD	A8	85.7	12	14
FALMER, HERMAN MD	A17	71.4	10	14
VELAZQUEZ, NORBERT MD	A4	92.3	12	13
PETERS, LEMUEL MD	A13	70.0	7	10
BOYLE, CLAUDIO MD	A19	60.0	6	10
WYGANSKI, JUSTON MD	A11	55.6	5	9

Build: 604 revision: 17457

Copyright (c) 2008-2010 Humedica, Inc. All rights reserved.

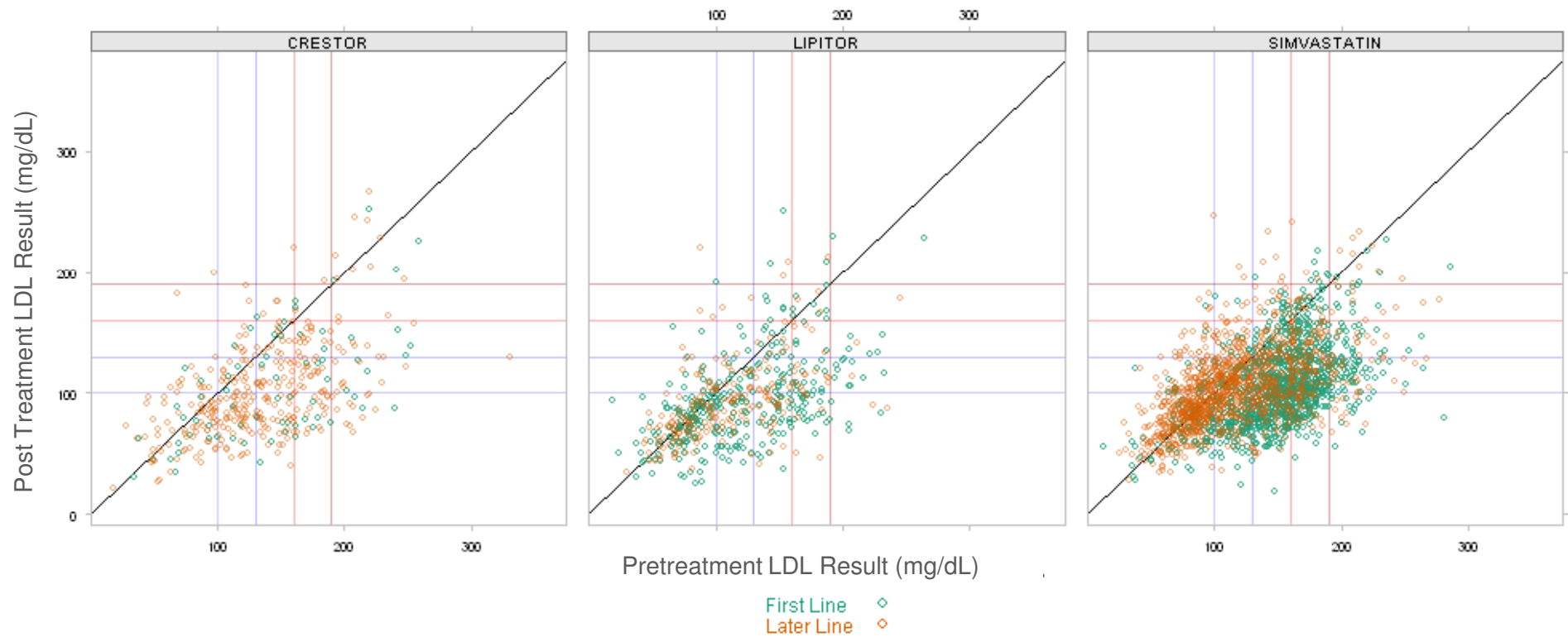
Transferring data from demo03.humedica.net...

# Pharma Marketing Questions Answerable with EHR Data

Group	Outstanding Questions	EHR Data-Based Solution
<b>Patients</b>	How is my brand being used (clinically) relative to competitive products?	Brand-level comparison of <b>patient's clinical profile at treatment initiation and follow-up</b>
<b>Physicians</b>	What are the drivers of physician prescribing behavior?	<b>Natural language processing to derive insight from physician notes</b> , including physician treatment rationale
<b>Hospitals</b>	How are hospital stays benefiting or disrupting brand use?	Analyze patient flow (i.e., in and out of hospital) and corresponding treatment patterns with <b>longitudinal clinical data</b>
<b>Payers</b>	How is the hospital formulary and/ or PBM formulary influencing the <i>clinical</i> adoption of brands?	Link patients' clinical profiles with insurance coverage to determine <b>formulary's influence on clinical practice</b> (i.e., treatment initiation) and leverage data to <b>inform negotiations with key payers</b>

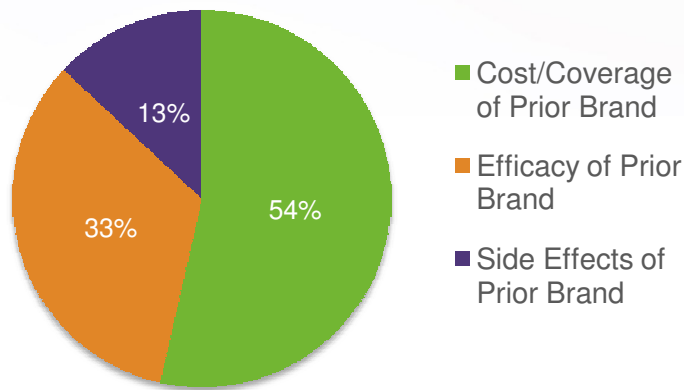
# Lab, Radiology and Pathology Data Allows You to Visualize Patient Outcomes

Pre vs. Post LDL Results by Brand and Line of Therapy



# Understand Rationale Behind Treatment Changes

## Distribution of Switching Rationale: Dyslipidemia



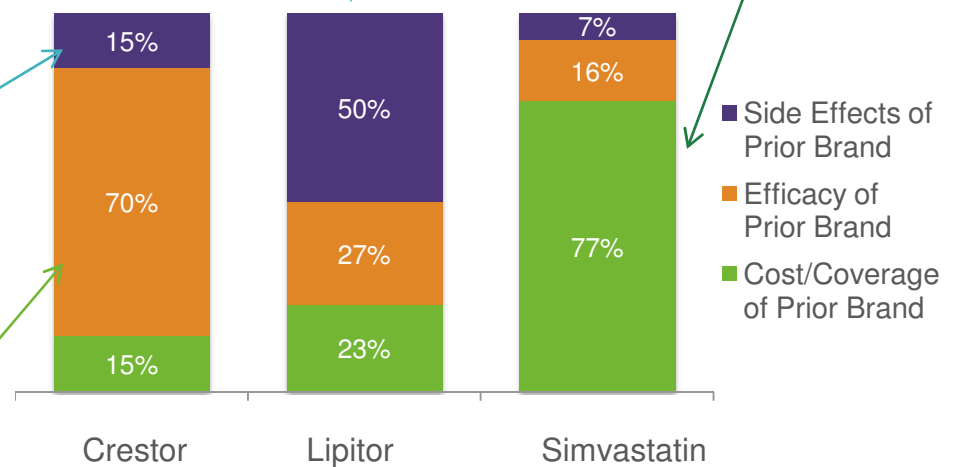
“The family would like to switch him from Lipitor to Crestor because reports indicate that the Crestor/grapefruit interaction is much less than with Lipitor and he very much likes to eat grapefruit.”

“Given the need for an LDL cholesterol solidly below 100 mg/dL, and preferably around 70 mg/dL, she will change from simvastatin to Crestor 20 mg daily.”

“He finds Lipitor too expensive and is interested in switching to generic.”

“He has had resolution of his headache since we switched from simvastatin to Lipitor.”

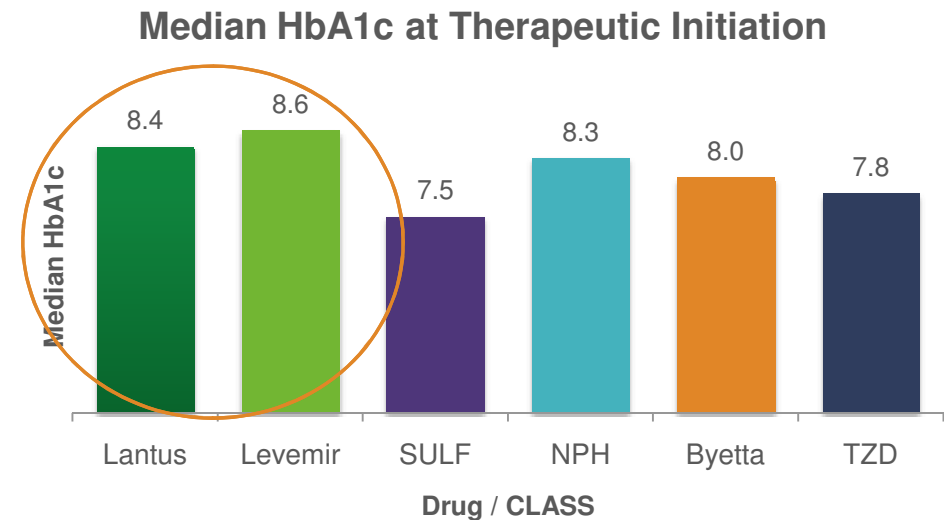
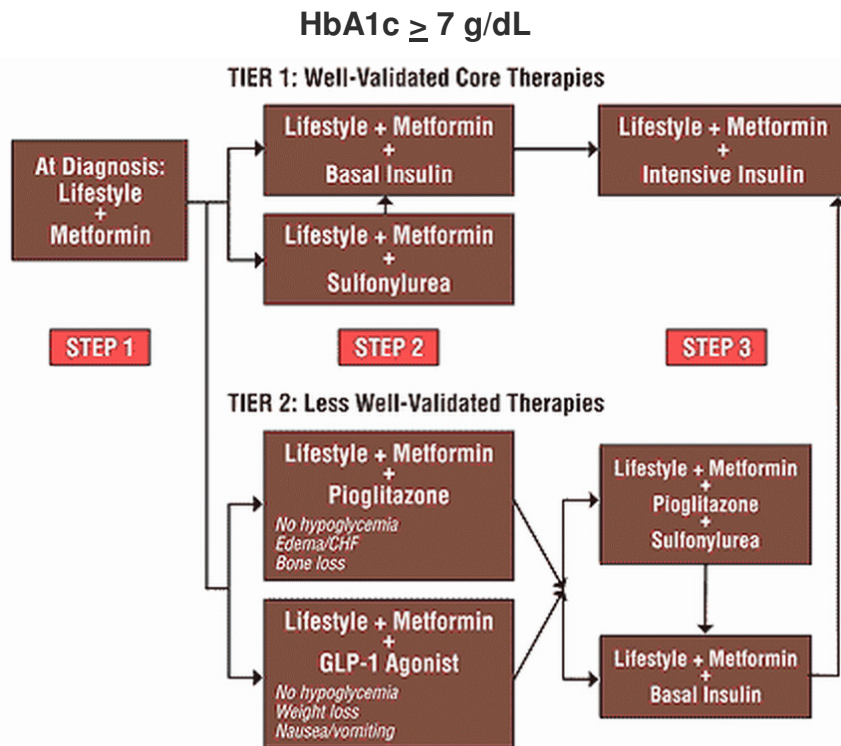
## Physician Rationale for Switching to



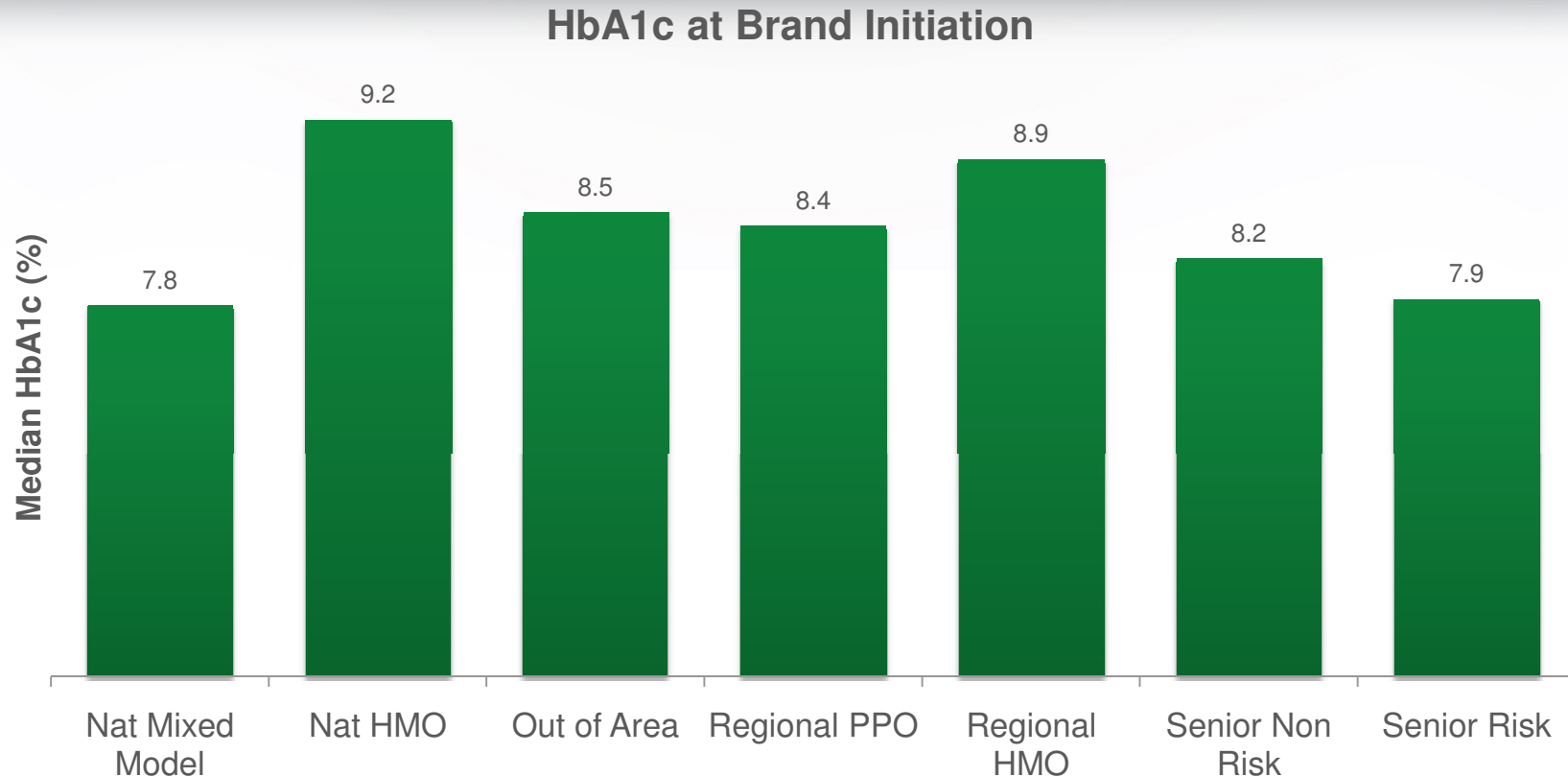


# Compare “Real-world” Clinical Practice to Established Treatment Protocols

Although the ADA recommends Basal Insulin as a first-line add-on to metformin when HbA1c is at or above 7 g/dL, Lantus and Levemir are prescribed at higher median HbA1c levels when compared to other medications, including other injectables.



# Understand How Payers Influence Clinical Practice



- Understand and track how formularies influence clinical practice
- Clear understanding of your product's initiation level *by payer category*
  - Ability to link patient's clinical profile (i.e. lab values, prescriptions, etc) with insurance coverage information

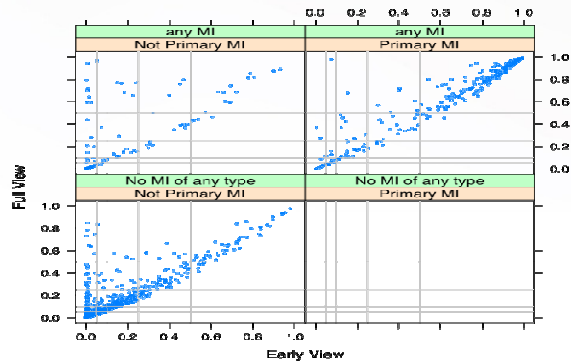
# Directions: Data Visualization



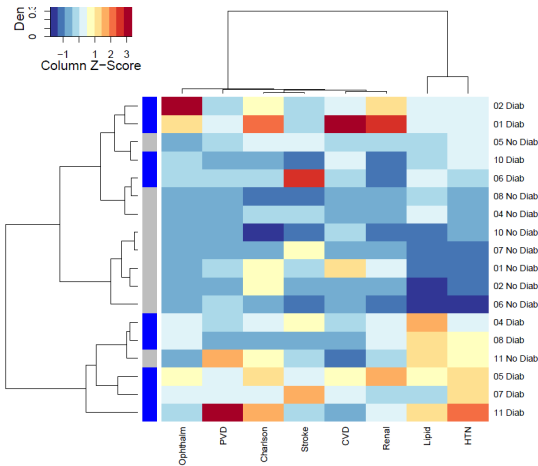
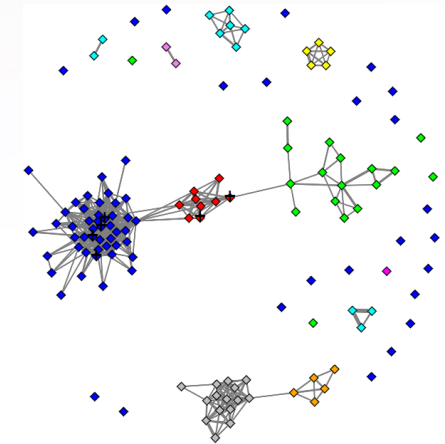
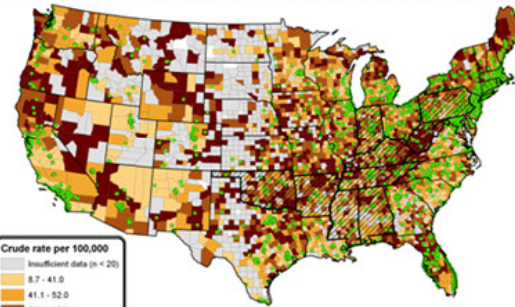
738942  
24322344323

243345098543534353  
59333

# Data Mining: Creating Value from Healthcare Data

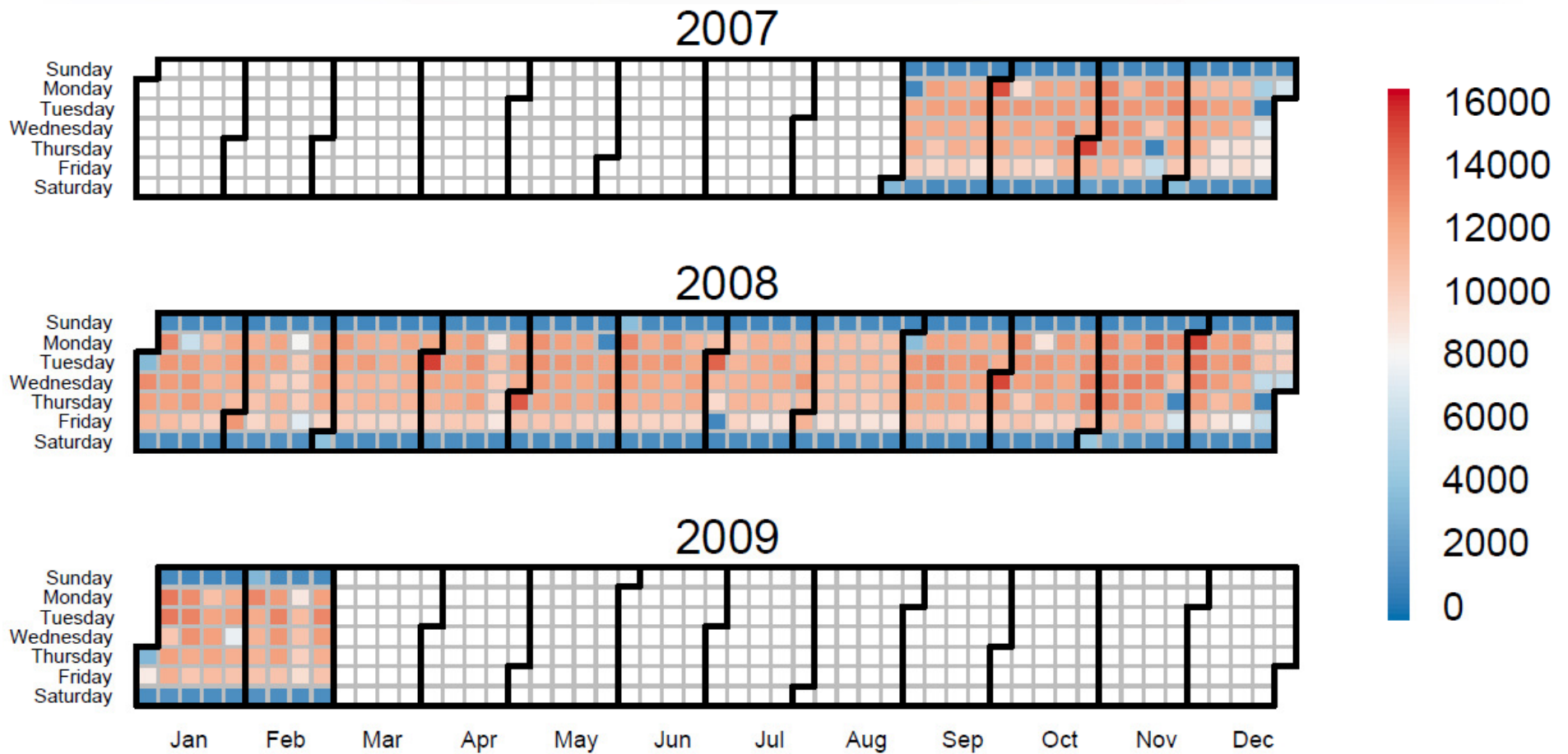


Crude COPD mortality rates per 100,000 population by county (1999-2002) with pulmonologist practice locations. Ten highest state-level current smoking rates in cross-hatching

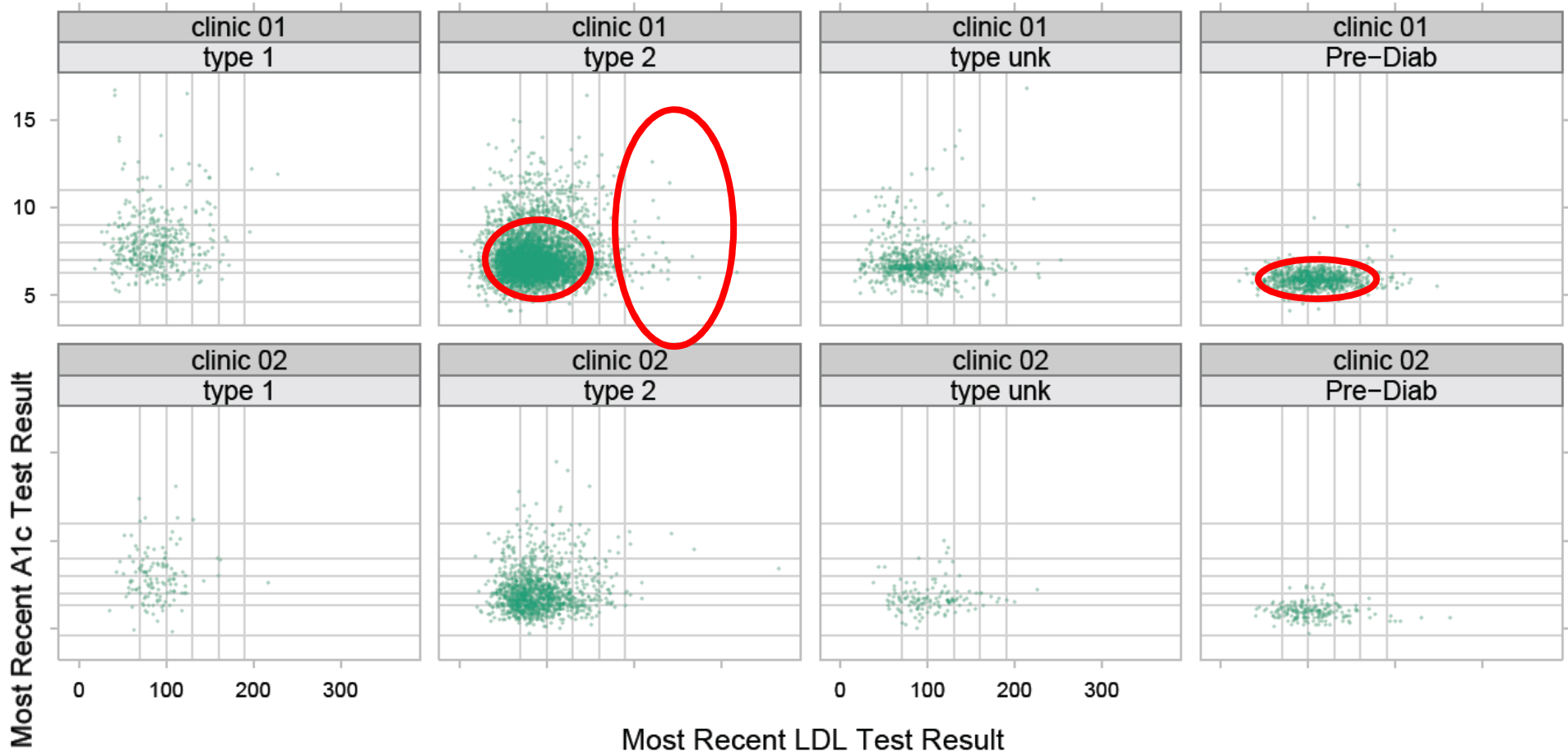


- Data Visualization
- Predictive Analytics
- Network and Clustering Analysis
- Geographical Analysis (GIS)

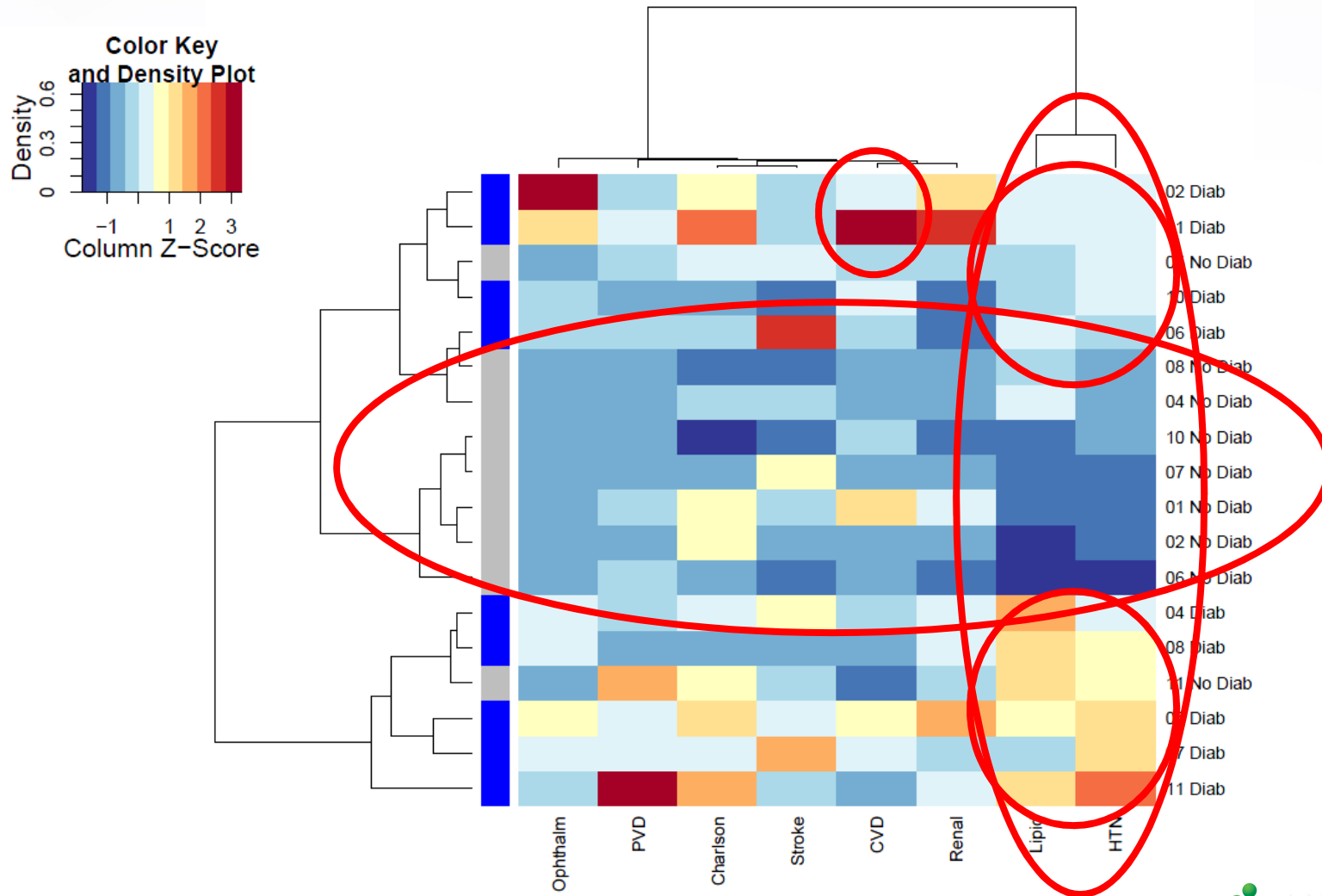
# Daily Encounter Volume



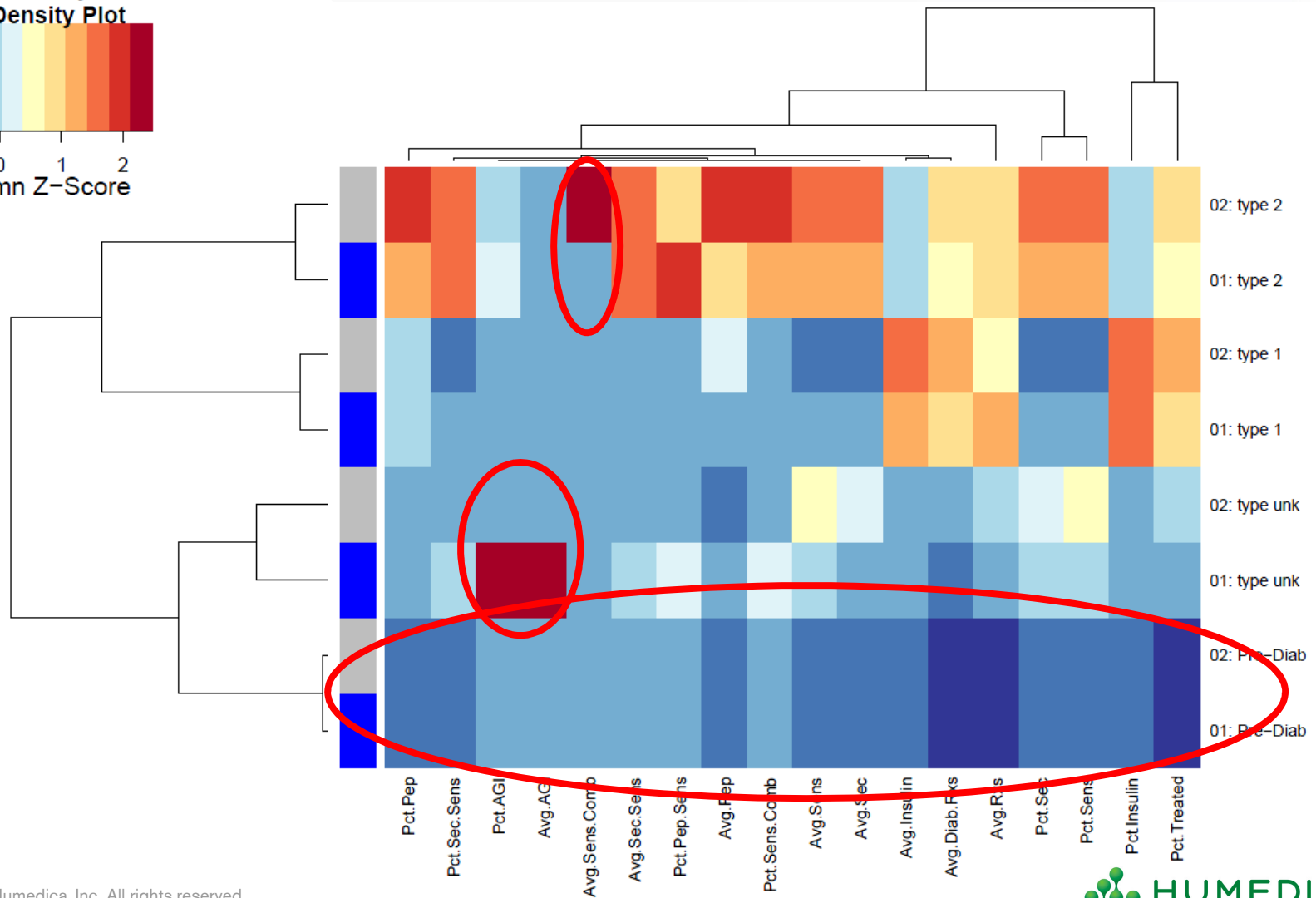
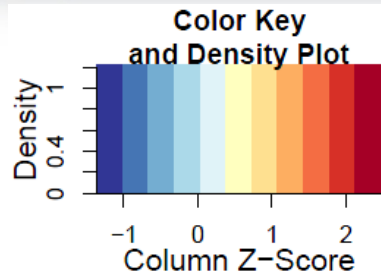
# HbA1c vs. LDL Scatterplot: Clinic 02 vs. 01



# Comorbidities by Site and Diabetes Status

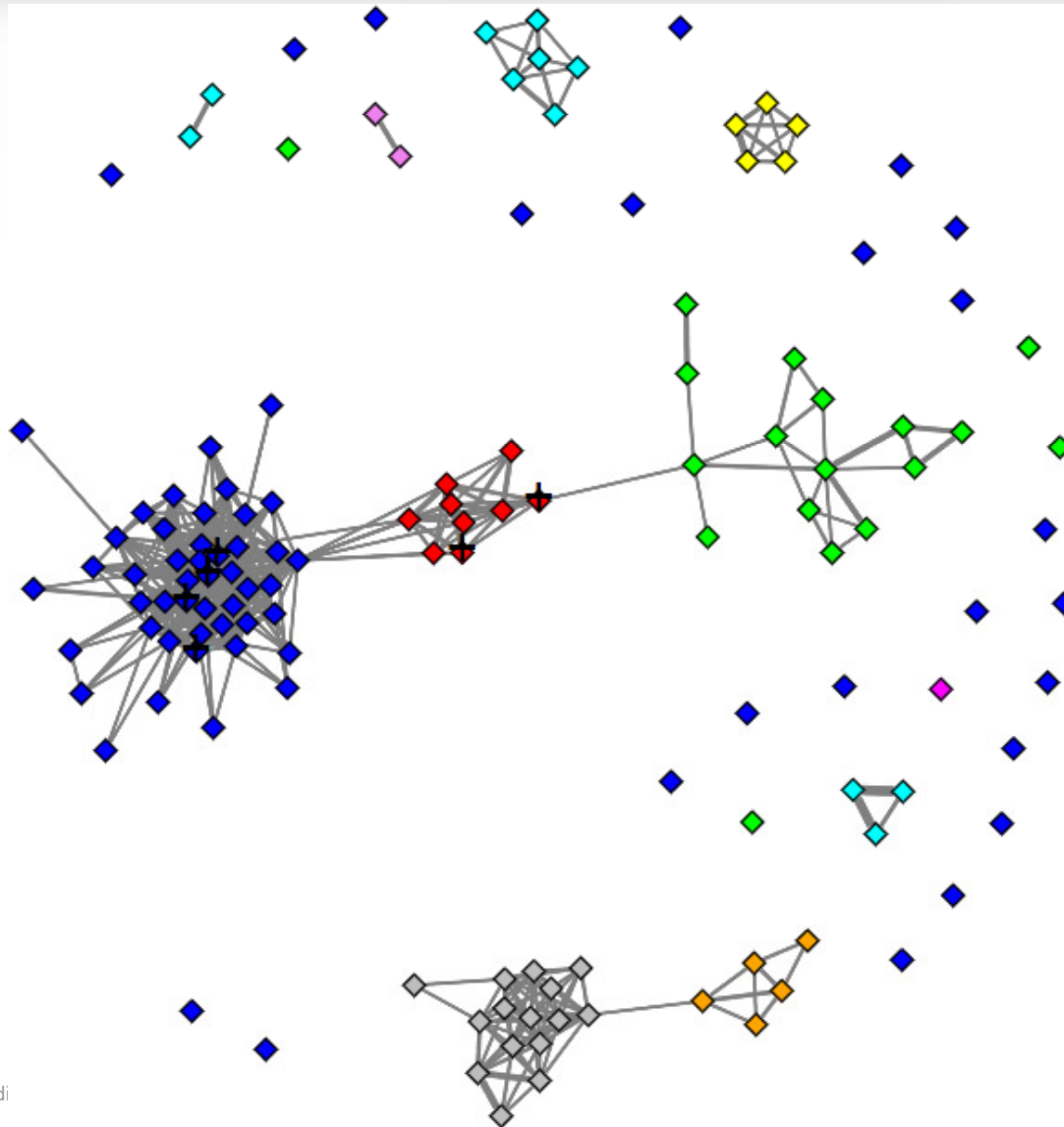


# Diabetes Prescribing Patterns





# Primary Care Physicians for Diabetes Patients

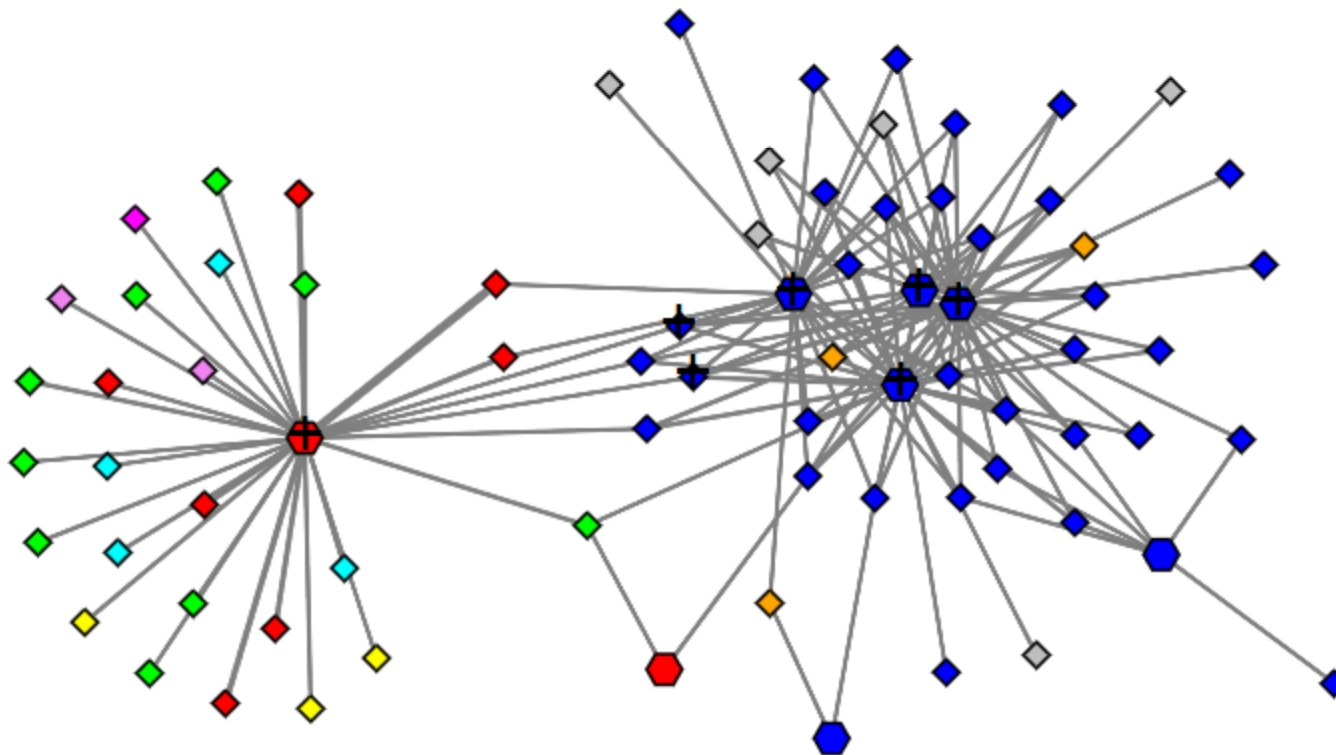


**Clinic 01 - blue**  
**Clinic 02 - red**  
Clinic 03 - cyan  
Clinic 04 - green  
Clinic 05 - grey  
Clinic 07 - yellow  
Clinic 08 - orange  
Clinic 10 - violet  
Clinic 11 - magenta

# Type 2 Diabetes Patients

Diamond – 1° care MD  
Hexagon – Endocrinologist

**Clinic 01 - blue**  
**Clinic 02 - red**  
Clinic 03 - cyan  
Clinic 04 - green  
Clinic 05 - grey  
Clinic 07 - yellow  
Clinic 08 - orange  
Clinic 10 - violet  
Clinic 11 – magenta



# Conclusions

- Secondary use of data has significant value for many stakeholders
- EHR adoption and improvements make this data available today
- NLP will remain an important source of data for the foreseeable future
- Secondary data must be managed carefully to prevent disclosure of private information
- Data visualization provides powerful insights, but requires a mindset and skills that are often not (yet) available in end users