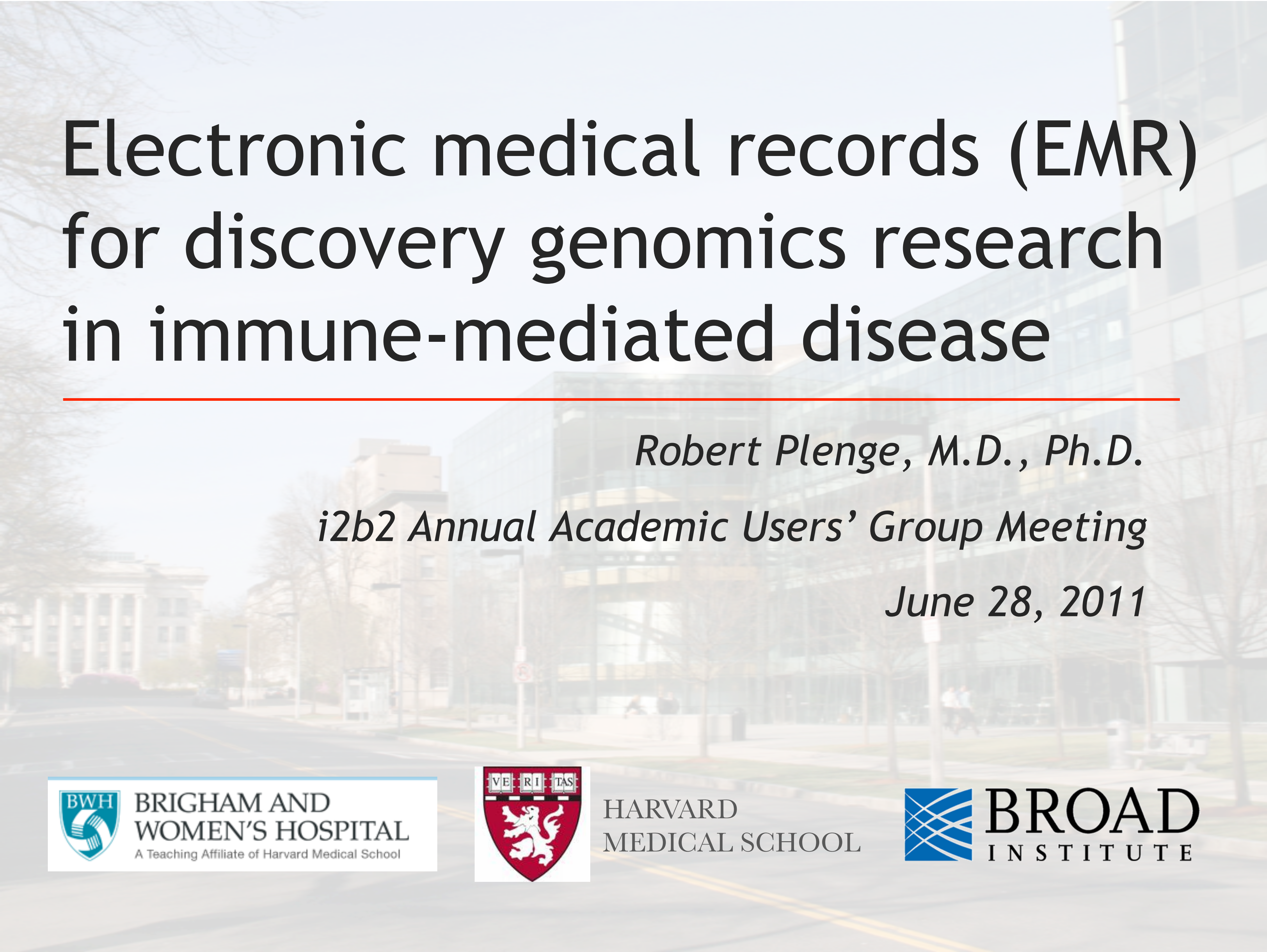


Electronic medical records (EMR) for discovery genomics research in immune-mediated disease

Robert Plenge, M.D., Ph.D.

i2b2 Annual Academic Users' Group Meeting

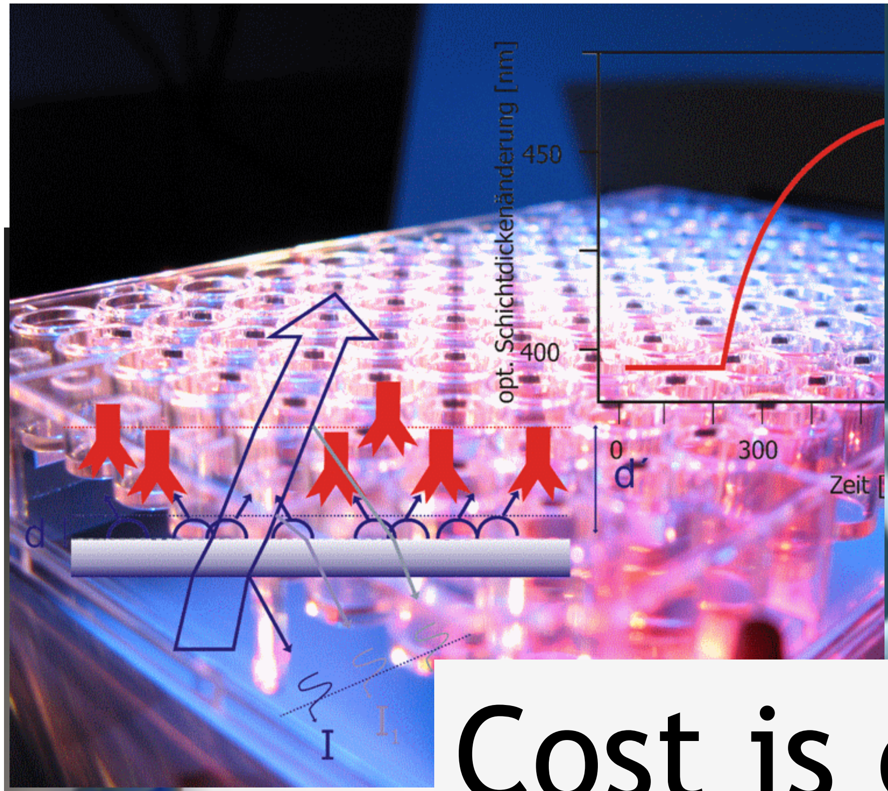
June 28, 2011

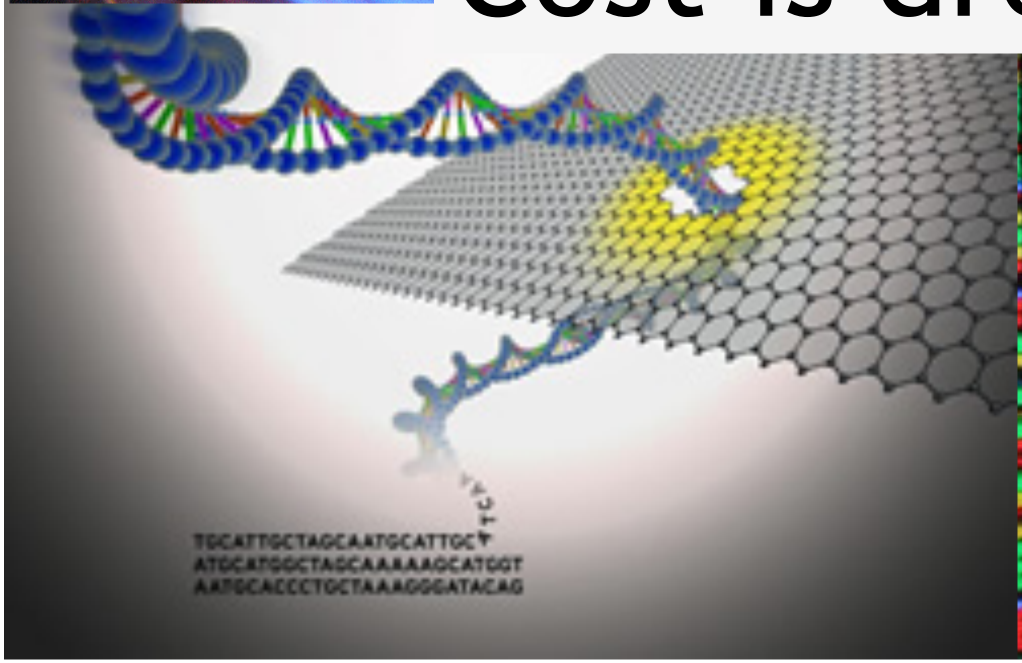
BWH BRIGHAM AND WOMEN'S HOSPITAL
A Teaching Affiliate of Harvard Medical School

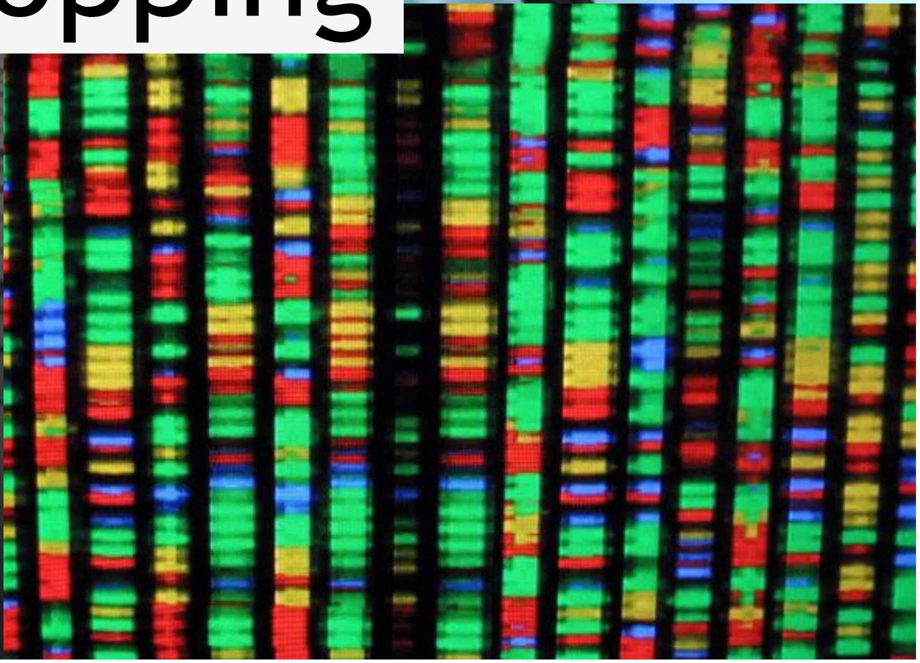
VE RI TAS HARVARD MEDICAL SCHOOL

BROAD INSTITUTE

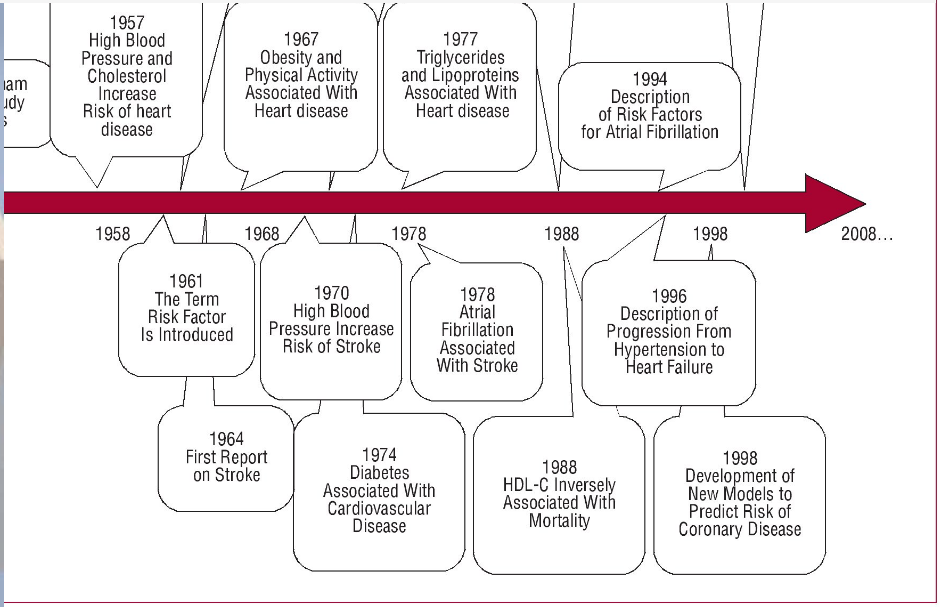
Cost is dropping




TGCATTGCTAGCAATGCATTGCTTCAA
ATGCATGGCTAGCAAAAAGCATGGT
AATGCACCCCTGCTAAAGGGATACAG



Phenotyping remains expensive



am
udy
s

1957
High Blood Pressure and Cholesterol Increase Risk of heart disease

1967
Obesity and Physical Activity Associated With Heart disease

1977
Triglycerides and Lipoproteins Associated With Heart disease

1994
Description of Risk Factors for Atrial Fibrillation

1958

1968

1978

1988

1998

2008…

1961
The Term Risk Factor Is Introduced

1970
High Blood Pressure Increase Risk of Stroke

1978
Atrial Fibrillation Associated With Stroke

1996
Description of Progression From Hypertension to Heart Failure

1964
First Report on Stroke

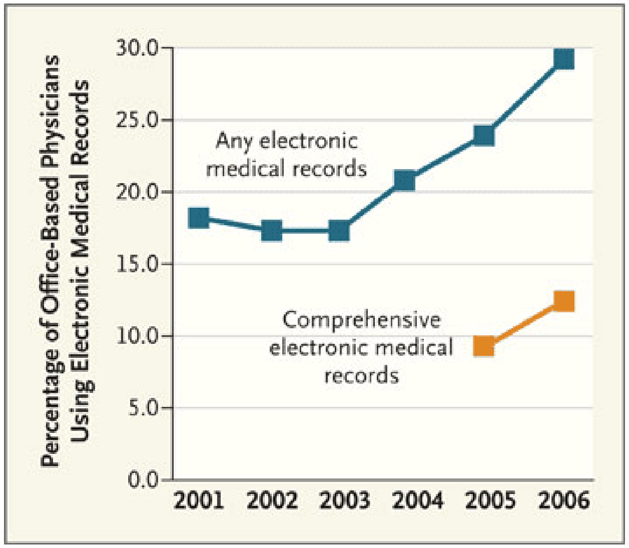
1974
Diabetes Associated With Cardiovascular Disease

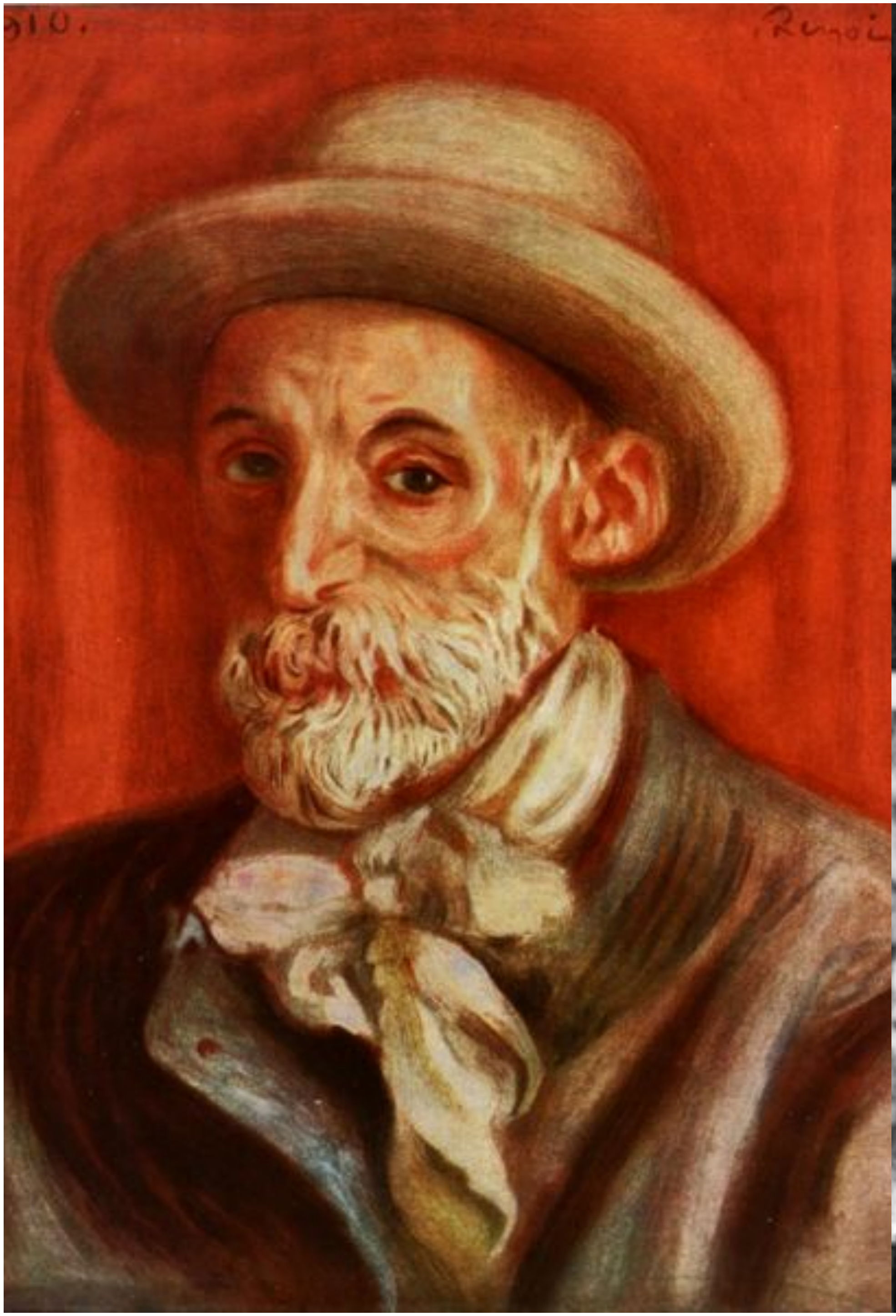
1988
HDL-C Inversely Associated With Mortality

1998
Development of New Models to Predict Risk of Coronary Disease

How will we realize the ultimate potential of genomics if phenotyping is rate-limiting?

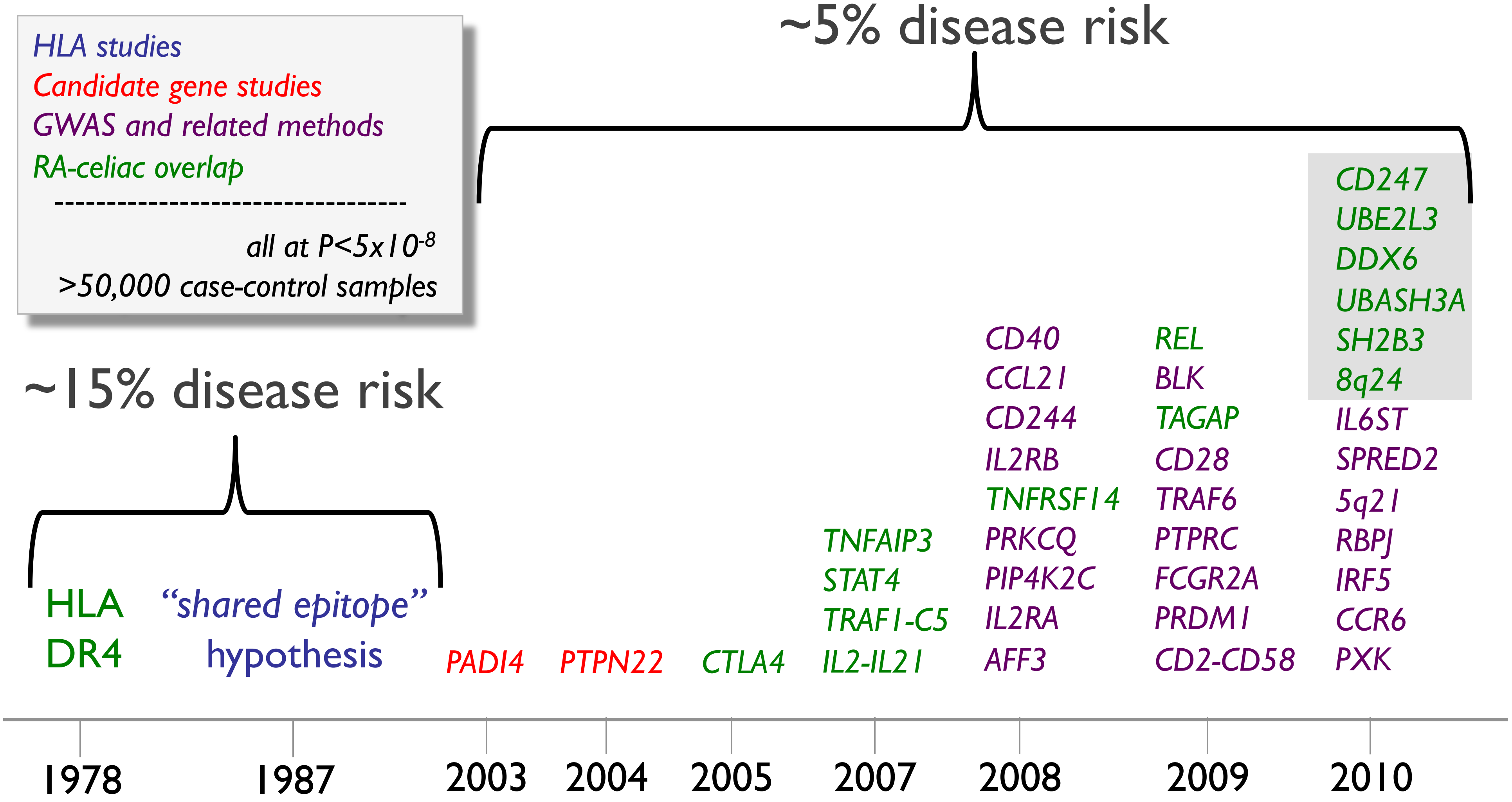
Can electronic medical records help?


Percentage of Office-Based Physicians Using Electronic Medical Records
30.0
25.0
20.0
15.0
10.0
5.0
0.0
2001
2002
2003
2004
2005
2006
Any electronic medical records
Comprehensive electronic medical records



Many risk loci remain “hidden”

HLA studies
Candidate gene studies
GWAS and related methods
RA-celiac overlap
----
all at $P<5x10^{-8}$
>50,000 case-control samples



~15% disease risk

HLA DR4

“shared epitope” hypothesis

~5% disease risk

PADI4

PTPN22

CTLA4

TNFAIP3
STAT4
TRAF1-C5
IL2-IL21

CD40
CCL21
CD244
IL2RB
TNFRSF14
PRKCQ
PIP4K2C
IL2RA
AFF3

REL
BLK
TAGAP
CD28
TRAF6
PTPRC
FCGR2A
PRDM1
CD2-CD58

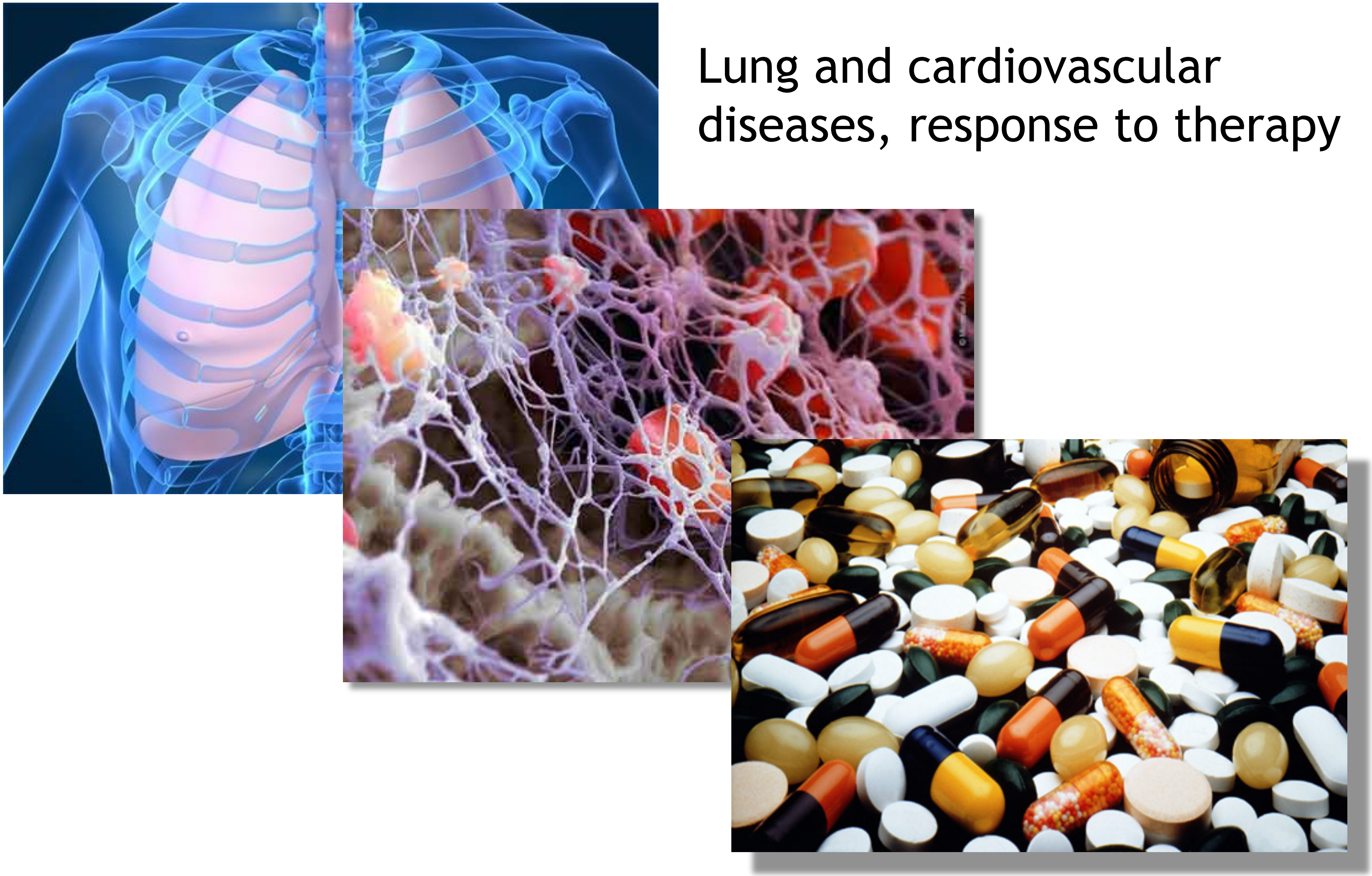
CD247
UBE2L3
DDX6
UBASH3A
SH2B3
8q24
IL6ST
SPRED2
5q21
RBPJ
IRF5
CCR6
PXK

1978 1987 2003 2004 2005 2007 2008 2009 2010

Zhernakova et al *PLoS Genetics* 2011

Clinically relevant subsets of RA

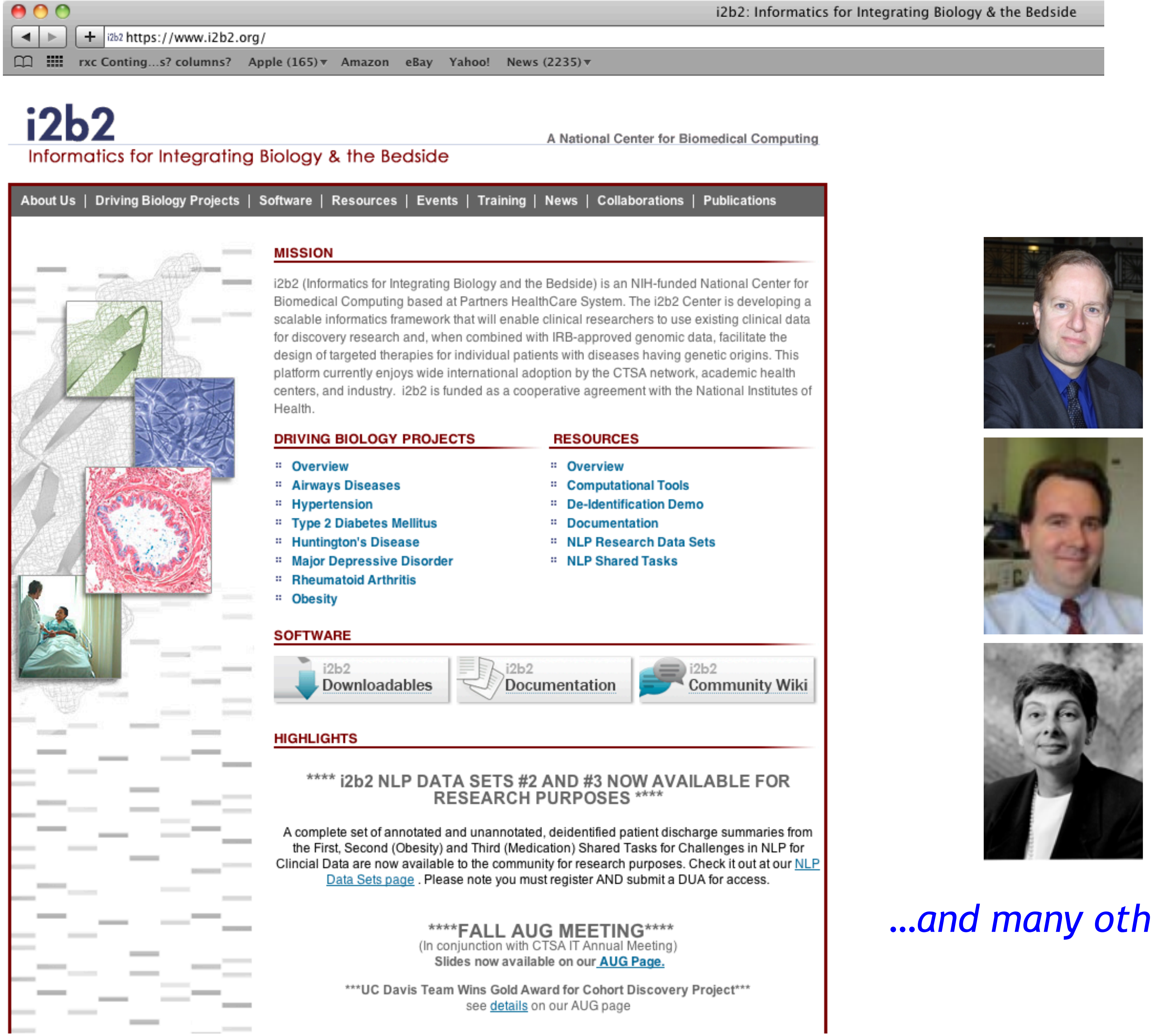
Lung and cardiovascular diseases, response to therapy



What are the options for collecting clinical data *and* DNA for genetic studies?

Options for clinical + DNA

design	Clinical data	DNA	Sample size	cost
clinical trial	+++	+++	+	$$$
registry	++	+++	++	$$
claims data	+	n/a	+++	$
EMR	++	+++	+++	$


i2b2: Informatics for Integrating Biology & the Bedside
https://www.i2b2.org/
rxc Conting...s? columns?
Apple (165)
Amazon
eBay
Yahoo!
News (2235)

i2b2
Informatics for Integrating Biology & the Bedside

A National Center for Biomedical Computing

About Us | Driving Biology Projects | Software | Resources | Events | Training | News | Collaborations | Publications

MISSION

i2b2 (Informatics for Integrating Biology and the Bedside) is an NIH-funded National Center for Biomedical Computing based at Partners HealthCare System. The i2b2 Center is developing a scalable informatics framework that will enable clinical researchers to use existing clinical data for discovery research and, when combined with IRB-approved genomic data, facilitate the design of targeted therapies for individual patients with diseases having genetic origins. This platform currently enjoys wide international adoption by the CTSA network, academic health centers, and industry. i2b2 is funded as a cooperative agreement with the National Institutes of Health.

DRIVING BIOLOGY PROJECTS

- Overview
- Airways Diseases
- Hypertension
- Type 2 Diabetes Mellitus
- Huntington's Disease
- Major Depressive Disorder
- Rheumatoid Arthritis
- Obesity

RESOURCES

- Overview
- Computational Tools
- De-Identification Demo
- Documentation
- NLP Research Data Sets
- NLP Shared Tasks

SOFTWARE

HIGHLIGHTS

**** i2b2 NLP DATA SETS #2 AND #3 NOW AVAILABLE FOR RESEARCH PURPOSES ****

A complete set of annotated and unannotated, deidentified patient discharge summaries from the First, Second (Obesity) and Third (Medication) Shared Tasks for Challenges in NLP for Clincial Data are now available to the community for research purposes. Check it out at our NLP Data Sets page . Please note you must register AND submit a DUA for access.

****FALL AUG MEETING****
(In conjunction with CTSA IT Annual Meeting)
Slides now available on our AUG Page.

UC Davis Team Wins Gold Award for Cohort Discovery Project
see details on our AUG page

Kohane

Murphy

Churchill

...and many others!

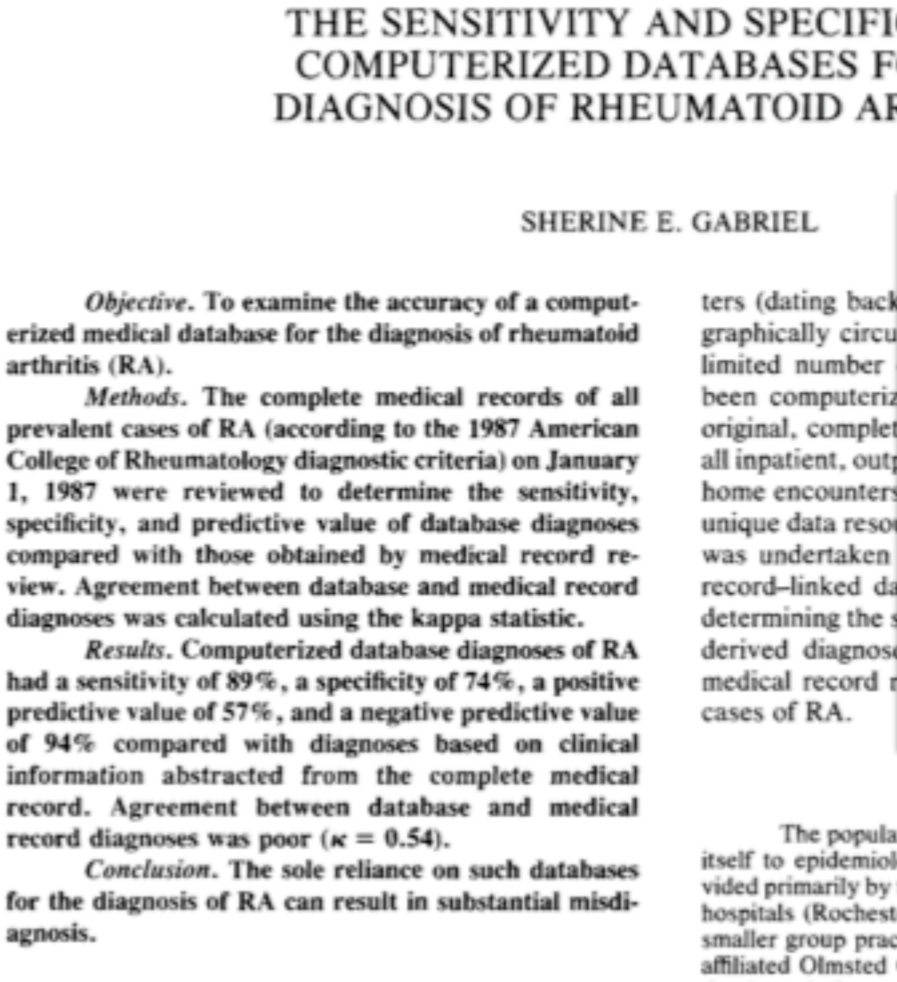
Outline of talk today

- Demonstration: developing an algorithm to define an RA cohort, proof-of-concept genomic studies

- Portability: implementing the EMR classification algorithm at other institutions

- Application: defining subsets of patients with clinically-relevant outcomes – and cardiovascular disease in particular

This is not a new idea…



THE SENSITIVITY AND SPECIFICITY OF COMPUTERIZED DATABASES FOR THE DIAGNOSIS OF RHEUMATOID ARTHRITIS

SHERINE E. GABRIEL

Objective. To examine the accuracy of a computerized medical database for the diagnosis of rheumatoid arthritis (RA).

Methods. The complete medical records of all prevalent cases of RA (according to the 1987 American College of Rheumatology diagnostic criteria) on January 1, 1987 were reviewed to determine the sensitivity, specificity, and predictive value of database diagnoses compared with those obtained by medical record review. Agreement between database and medical record diagnoses was calculated using the kappa statistic.

Results. Computerized database diagnoses of RA had a sensitivity of 89%, a specificity of 74%, a positive predictive value of 57%, and a negative predictive value of 94% compared with diagnoses based on clinical information abstracted from the complete medical record. Agreement between database and medical record diagnoses was poor (κ = 0.54).

Conclusion. The sole reliance on such databases for the diagnosis of RA can result in substantial misdiagnosis.

ters (dating back
graphically circu
limited number
been computeriz
original, complet
all inpatient, outp
home encounters
unique data reso
was undertaken
record–linked da
determining the s
derived diagnos
medical record r
cases of RA.

The population
itself to epidemiologic
vided primarily by the M
hospitals (Rochester M
smaller group practice
affiliated Olmsted Com

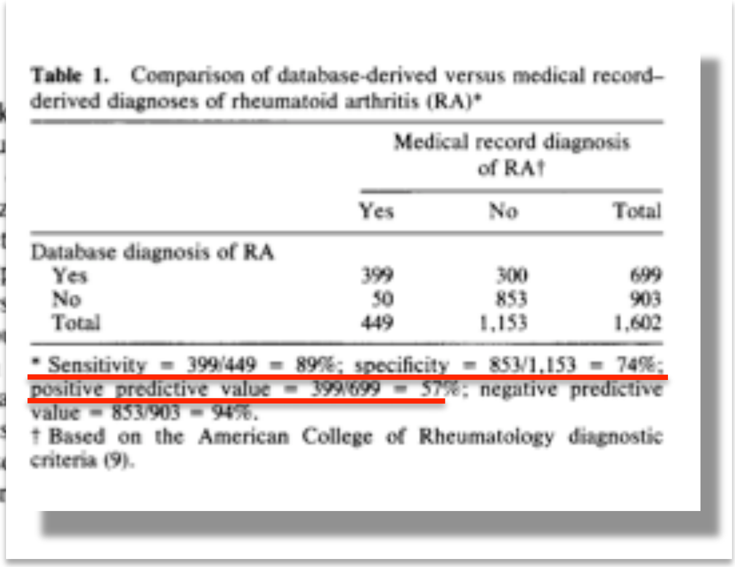


Table 1. Comparison of database-derived versus medical record–derived diagnoses of rheumatoid arthritis (RA)*

	Medical record diagnosis of RA†		
	Yes	No	Total
Database diagnosis of RA			
Yes	399	300	699
No	50	853	903
Total	449	1,153	1,602

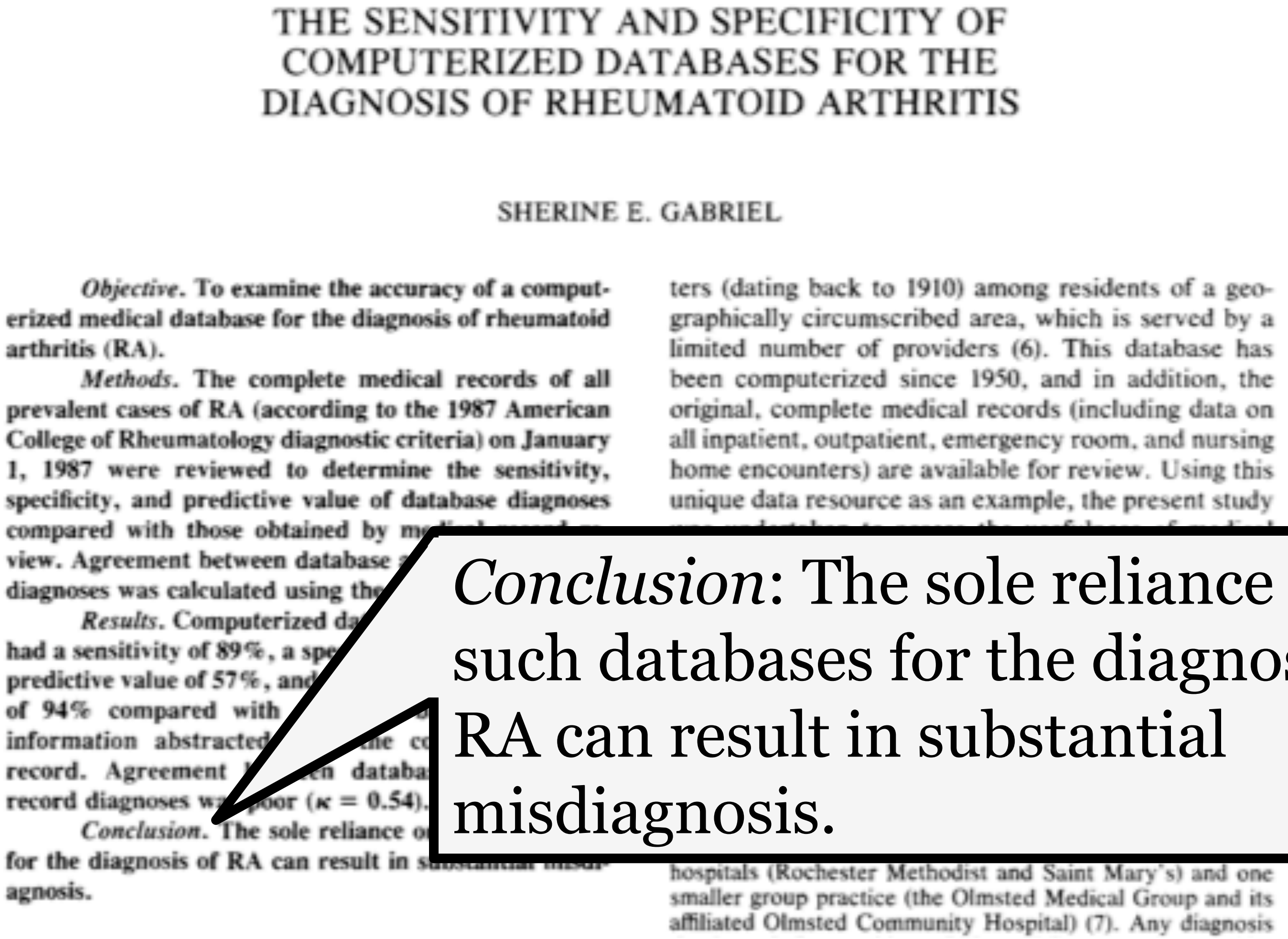
* Sensitivity = 399/449 = 89%; specificity = 853/1,153 = 74%; positive predictive value = 399/699 = 57%; negative predictive value = 853/903 = 94%.

† Based on the American College of Rheumatology diagnostic criteria (9).

Sens:	89%
PPV:	57%

Gabriel (1994) *Arthritis and Rheumatism*

…but EMR data are “*dirty*”



THE SENSITIVITY AND SPECIFICITY OF COMPUTERIZED DATABASES FOR THE DIAGNOSIS OF RHEUMATOID ARTHRITIS

SHERINE E. GABRIEL

Objective. To examine the accuracy of a computerized medical database for the diagnosis of rheumatoid arthritis (RA).

Methods. The complete medical records of all prevalent cases of RA (according to the 1987 American College of Rheumatology diagnostic criteria) on January 1, 1987 were reviewed to determine the sensitivity, specificity, and predictive value of database diagnoses compared with those obtained by me
view. Agreement between database
diagnoses was calculated using the

Results. Computerized da
had a sensitivity of 89%, a spe
predictive value of 57%, and
of 94% compared with
information abstracted
record. Agreement
record diagnoses w
poor (κ = 0.54).

Conclusion. The sole reliance o
for the diagnosis of RA can result in s
agnosis.

ters (dating back to 1910) among residents of a geographically circumscribed area, which is served by a limited number of providers (6). This database has been computerized since 1950, and in addition, the original, complete medical records (including data on all inpatient, outpatient, emergency room, and nursing home encounters) are available for review. Using this unique data resource as an example, the present study

hospitals (Rochester Methodist and Saint Mary's) and one smaller group practice (the Olmsted Medical Group and its affiliated Olmsted Community Hospital) (7). Any diagnosis

Conclusion: The sole reliance on such databases for the diagnosis of RA can result in substantial misdiagnosis.

Gabriel (1994) *Arthritis and Rheumatism*

Partners HealthCare: *4 million patients*

BRIGHAM AND WOMEN'S HOSPITAL *

MASSACHUSETTS GENERAL HOSPITAL *

FAULKNER HOSPITAL

NEWTON-WELLESLEY HOSPITAL

McLEAN HOSPITAL*

PARTNERS CONTINUING CARE

MGH INSTITUTE OF HEALTH PROFESSIONS

PARTNERS COMMUNITY HEALTHCARE, INC.

NORTH SHORE MEDICAL CENTER

MARTHA'S VINEYARD HOSPITAL

NANTUCKET COTTAGE HOSPITAL

Partners HealthCare: *linked by EMR*

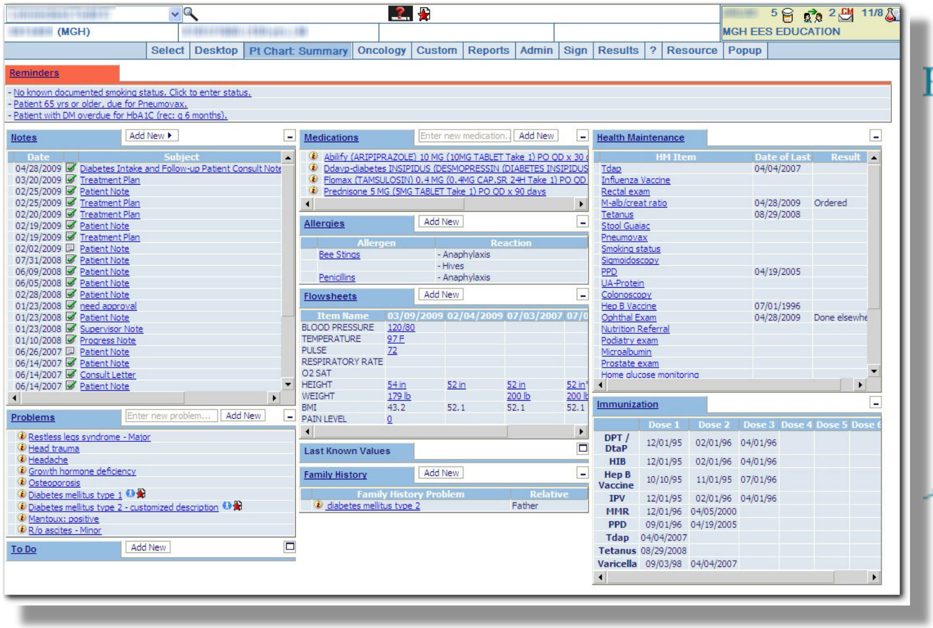
BRIGHAM AND WOMEN'S HOSPITAL *

MASSACHUSETTS GENERAL HOSPITAL *

FAULKNER HOSPITAL

NEWTON-WELLESLEY HOSPITAL

McLEAN HOSPITAL*



Reminders

- No known documented smoking status. Click to enter status.
- Patient 65 yrs or older, due for Pneumovax.
- Patient with DM overdue for HbA1C (rec: q 6 months).

Notes

Date		Subject
04/28/2009	✓	Diabetes Intake and Follow-up Patient Consult Note
03/20/2009	✓	Treatment Plan
02/25/2009	✓	Patient Note
02/25/2009	✓	Treatment Plan
02/20/2009	✓	Treatment Plan
02/19/2009	✓	Patient Note
02/19/2009	✓	Treatment Plan
02/02/2009	▤	Patient Note
07/31/2008	✓	Patient Note
06/09/2008	✓	Patient Note
06/05/2008	✓	Patient Note
02/28/2008	✓	Patient Note
01/23/2008	✓	need approval
01/23/2008	✓	Patient Note
01/23/2008	✓	Supervisor Note
01/10/2008	✓	Progress Note
06/26/2007	▤	Patient Note
06/14/2007	✓	Patient Note
06/14/2007	✓	Consult Letter
06/14/2007	✓	Patient Note

Problems

- Restless legs syndrome - Major
- Head trauma
- Headache
- Growth hormone deficiency
- Osteoporosis
- Diabetes mellitus type 1
- Diabetes mellitus type 2 - customized description
- Mantoux: positive
- R/o ascites - Minor

To Do

Medications

- Abilify (ARIPIPRAZOLE) 10 MG (10MG TABLET Take 1) PO QD x 30
- Ddavp-diabetes INSIPIDUS (DESMOPRESSIN (DIABETES INSIPIDUS
- Flomax (TAMSULOSIN) 0.4 MG (0.4MG CAP.SR 24H Take 1) PO QD
- Prednisone 5 MG (5MG TABLET Take 1) PO QD x 90 days

Allergies

Allergen	Reaction
Bee Stings	- Anaphylaxis
	- Hives
Penicillins	- Anaphylaxis

Flowsheets

Item Name	03/09/2009	02/04/2009	07/03/2007	07/0
BLOOD PRESSURE	120/80			
TEMPERATURE	97 F			
PULSE	72			
RESPIRATORY RATE				
O2 SAT				
HEIGHT	54 in	52 in	52 in	52 in*
WEIGHT	179 lb		200 lb	200 lb
BMI	43.2	52.1	52.1	52.1
PAIN LEVEL	0			

Last Known Values

Family History

Family History Problem	Relative
diabetes mellitus type 2	Father

Health Maintenance

HM Item	Date of Last	Result
Tdap	04/04/2007	
Influenza Vaccine		
Rectal exam		
M-alb/creat ratio	04/28/2009	Ordered
Tetanus	08/29/2008	
Stool Guaiac		
Pneumovax		
Smoking status		
Sigmoidoscopy		
PPD	04/19/2005	
UA-Protein		
Colonoscopy		
Hep B Vaccine	07/01/1996	
Ophthal Exam	04/28/2009	Done elsewhe
Nutrition Referral		
Podiatry exam		
Microalbumin		
Prostate exam		
Home glucose monitoring		

Immunization

	Dose 1	Dose 2	Dose 3	Dose 4	Dose 5	Dose 6
DPT / DtaP	12/01/95	02/01/96	04/01/96			
HIB	12/01/95	02/01/96	04/01/96			
Hep B Vaccine	10/10/95	11/01/95	07/01/96			
IPV	12/01/95	02/01/96	04/01/96			
MMR	12/01/96	04/05/2000				
PPD	09/01/96	04/19/2005				
Tdap	04/04/2007					
Tetanus	08/29/2008					
Varicella	09/03/98	04/04/2007				

PARTNERS CONTINUING CARE

MGH INSTITUTE OF HEALTH PROFESSIONS

PARTNERS COMMUNITY HEALTHCARE, INC.

NORTH SHORE MEDICAL CENTER

MARTHA'S VINEYARD HOSPITAL

NANTUCKET COTTAGE HOSPITAL

Partners HealthCare: *organized by i2b2*

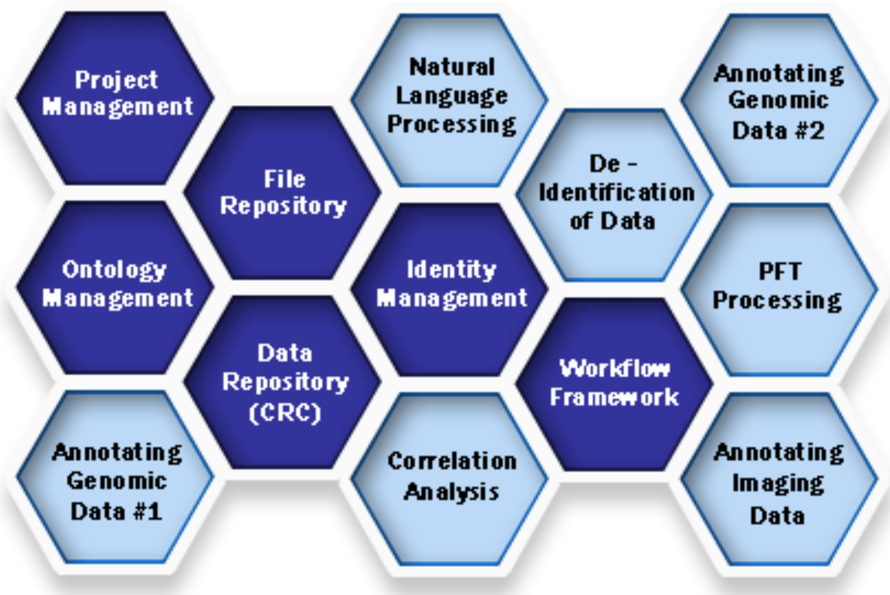
BRIGHAM AND WOMEN'S HOSPITAL *

MASSACHUSETTS GENERAL HOSPITAL *

FAULKNER HOSPITAL

NEWTON-WELLESLEY HOSPITAL

McLEAN HOSPITAL*



PARTNERS CONTINUING CARE

MGH INSTITUTE OF HEALTH PROFESSIONS

PARTNERS COMMUNITY HEALTHCARE, INC.

NORTH SHORE MEDICAL CENTER

MARTHA'S VINEYARD HOSPITAL

NANTUCKET COTTAGE HOSPITAL

Our library of RA phenotypes

- Natural language processing (NLP)
  - *disease terms (e.g., RA, lupus)*
  - *medications (e.g., methotrexate)*
  - *autoantibodies (e.g., CCP, RF)*
  - *radiographic erosions*

Qing Zeng

Guergana Savova

Concept/term	Accuracy of concept
presence of erosion	88%
seropositive	96%
CCP positive	98.7%
RF positive	99.3%
etanercept	100%
methotrexate	100%

Our library of RA phenotypes

- Natural language processing (NLP)
  - *disease terms (e.g., RA, lupus)*
  - *medications (e.g., methotrexate)*
  - *autoantibodies (e.g., CCP, RF)*
  - *radiographic erosions*

- Codified data
  - *ICD9 disease codes*
  - *prescription medications*
  - *laboratory autoantibodies*

Shawn Murphy

4 million patients

ICD9 RA and/or CCP checked
(goal = high sensitivity)

31,171 patients

Classification algorithm
(goal = high PPV)

3,585
RA patients

Liao et al (2010) *Arth Res Therapy*

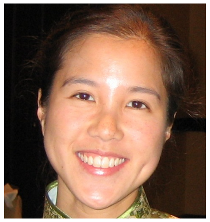
High PPV with adequate sensitivity

Model	PPV (SE)	Sensitivity (SE)
Codified + NLP	*0.93 (0.02)* ✪	*0.63 (0.06)*
NLP only	*0.89 (0.02)*	*0.56 (0.05)*
Codified only	*0.88 (0.02)*	*0.51 (0.05)*

✪392 out of 400 (98%) had definite or possible RA!

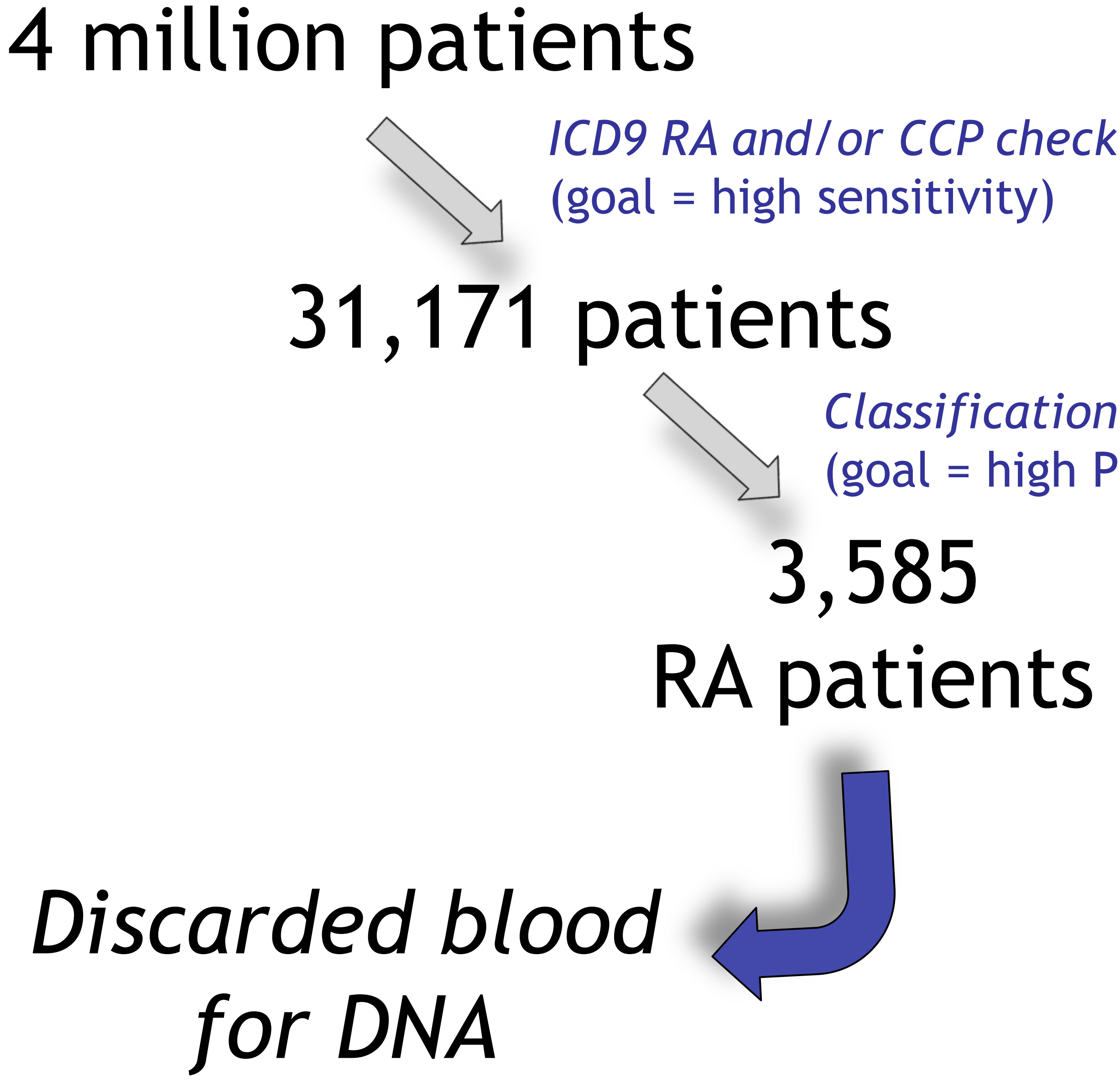
Liao et al (2010) *Arth Res Therapy*

Clinical features of patients



Characteristics	*i2b2* RA	CORRONA
total number	3,585	7,971
Mean age (SD)	57.5 (17.5)	58.9 (13.4)
Female (%)	79.9	74.5
Anti-CCP(%)	63	N/A
RF (%)	74.4	72.1
Erosions (%)	59.2	59.7
MTX (%)	59.5	52.8
Anti-TNF (%)	32.6	22.6

CCP has an OR = 1.5 for predicting erosions



4 million patients

ICD9 RA and/or CCP checked
(goal = high sensitivity)

31,171 patients

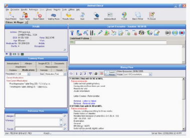
Classification algorithm
(goal = high PPV)

3,585
RA patients

Discarded blood
for DNA

“*On demand*” biorepository

NLP data

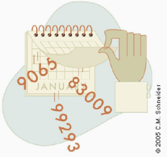


Narrative electronic medical record

Natural Language Processing (NLP)

Codified data

codified data (e.g., billing codes)



i2b2 informatics infrastructure

Algorithm to define patients with RA

i2b2 RA-DataMart

30,655 patients

NLP queries

autoantibody status

medication history

codified data

billing codes

laboratory values

Within 1 year (at $30/sample):

1,800 RA cases

2,400 matched controls

RA patients

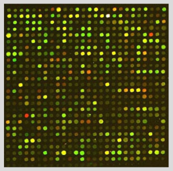
IDs	IDs
13100	87443
65773	61103
23001	49011
12543	12543

Match!

FIREWALL

anonymous clinical data

discarded blood sample for DNA

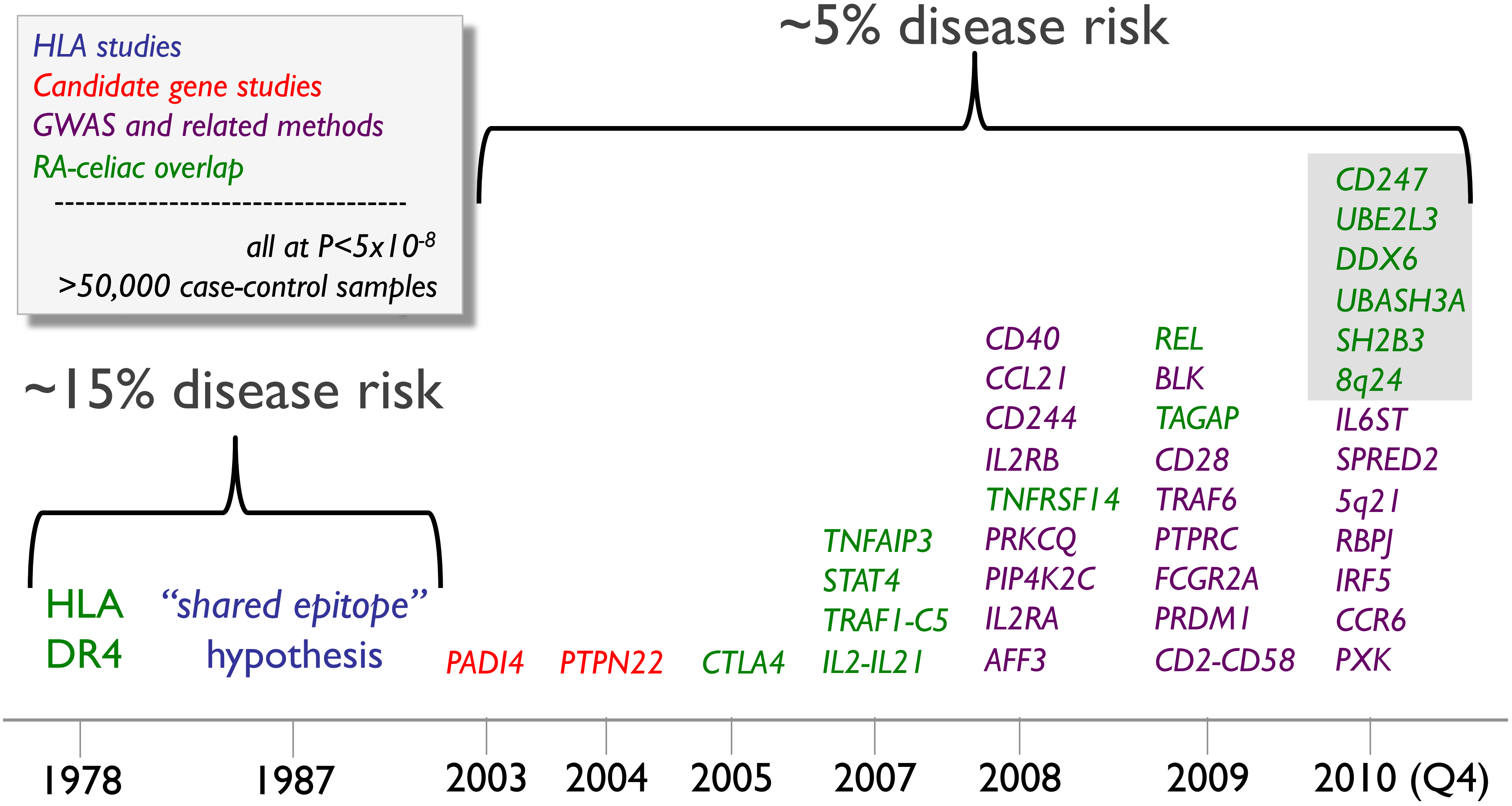


FINAL ANONYMOUS PATIENT CLINICAL DATA WITH DNA FOR GENETIC STUDIES

June 2011: >*35 RA risk loci*

HLA studies
Candidate gene studies
GWAS and related methods
RA-celiac overlap

all at $P<5x10^{-8}$
>50,000 case-control samples



~15% disease risk

~5% disease risk

1978
HLA DR4

1987
"shared epitope" hypothesis

2003
PADI4

2004
PTPN22

2005
CTLA4

2007
TNFAIP3
STAT4
TRAF1-C5
IL2-IL21

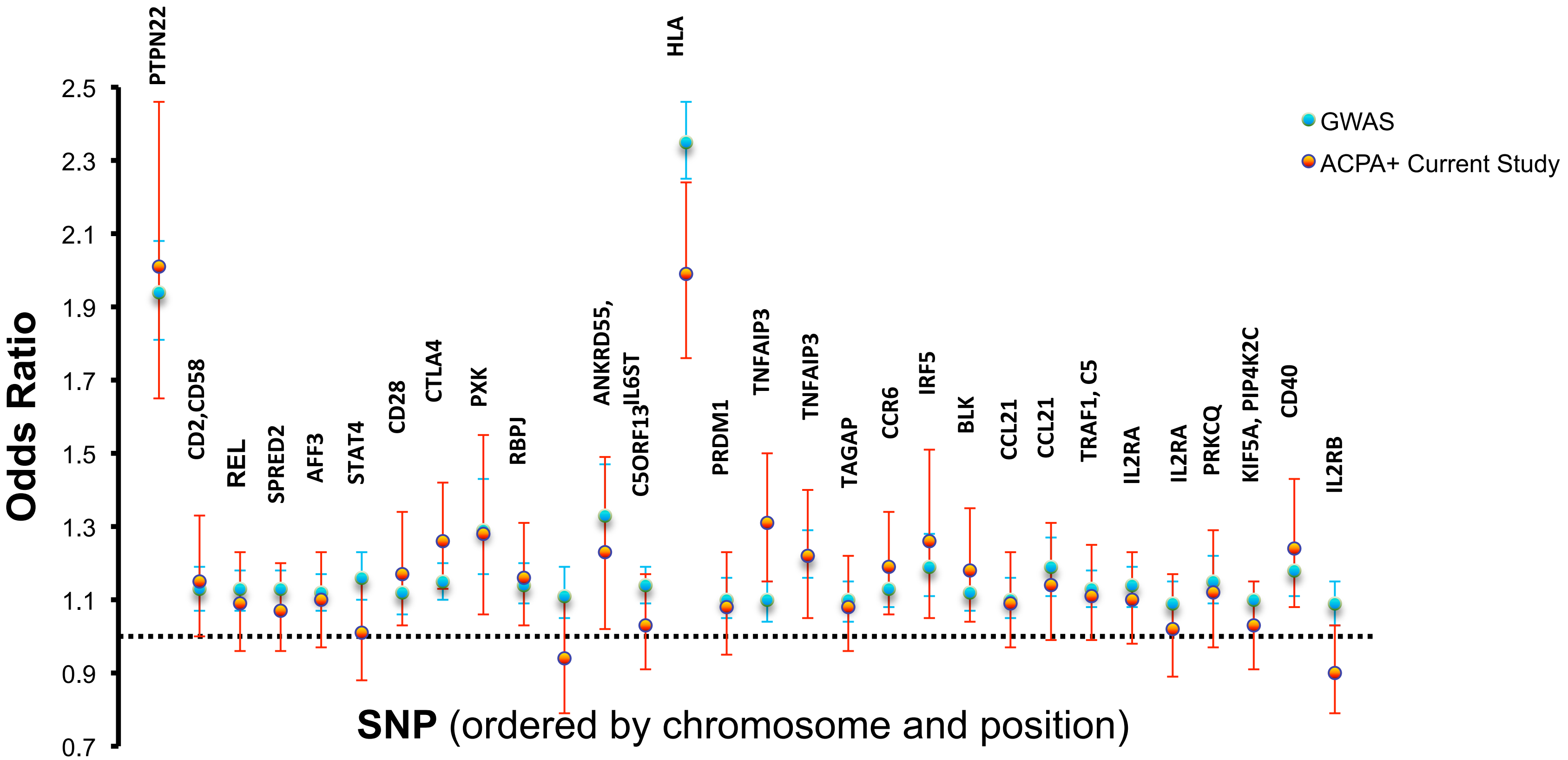
2008
CD40
CCL21
CD244
IL2RB
TNFRSF14
PRKCQ
PIP4K2C
IL2RA
AFF3

2009
REL
BLK
TAGAP
CD28
TRAF6
PTPRC
FCGR2A
PRDM1
CD2-CD58

2010 (Q4)
CD247
UBE2L3
DDX6
UBASH3A
SH2B3
8q24
IL6ST
SPRED2
5q21
RBPJ
IRF5
CCR6
PXK

Zhernakova et al *PLoS Genetics* 2011 (in press)

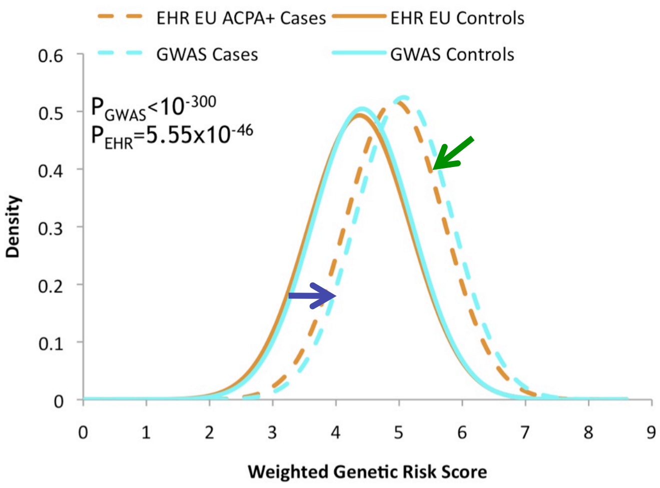
OR similar in EMR cohort



~1,500 multi-ethnic RA cases and 1,500 matched controls

Kurreeman et al *(*2011) *AJHG*

Genetic risk score similar…



Low GRS High GRS

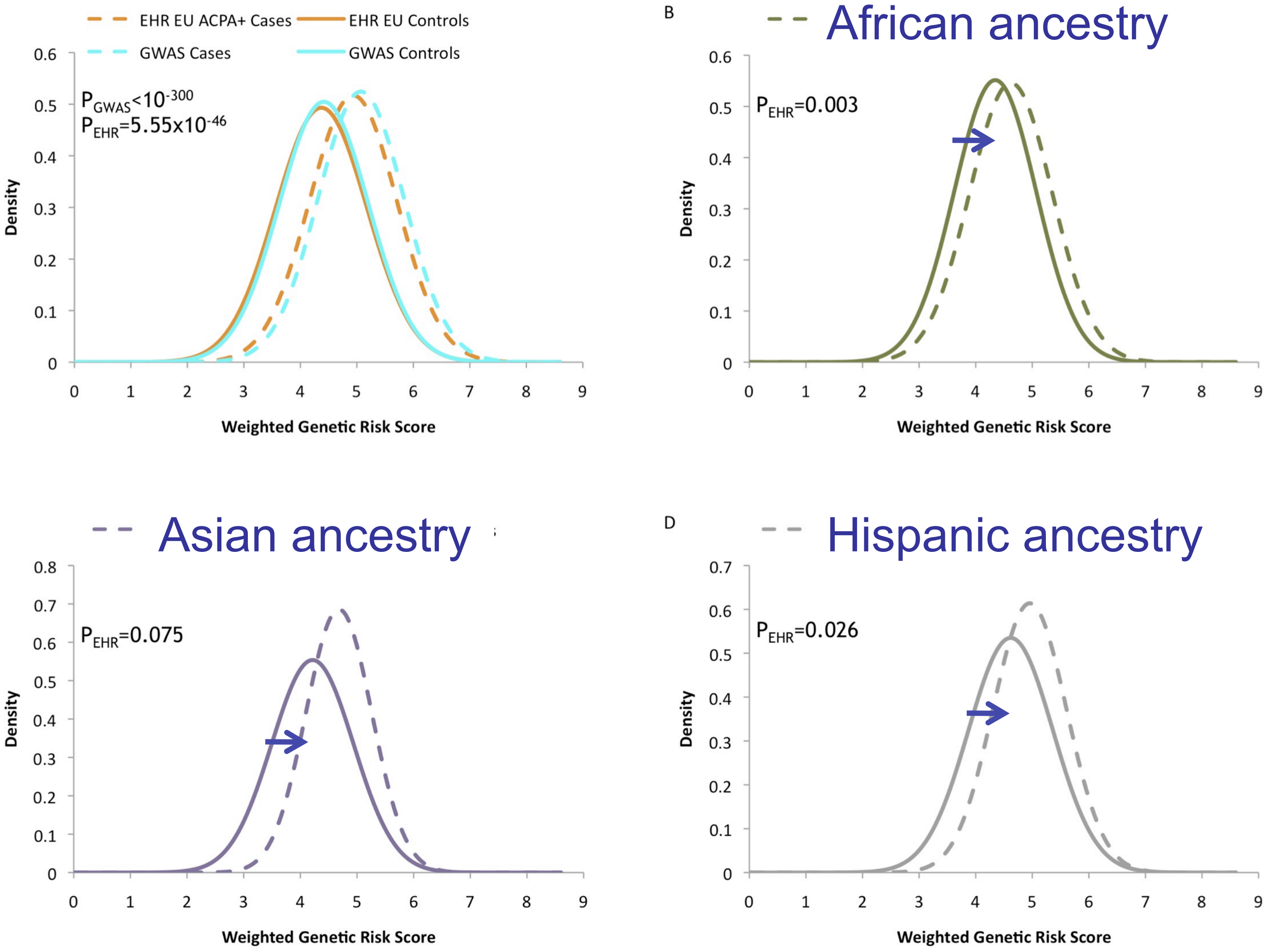
European ancestry

→ RA case vs control

→ GWAS vs EMR (no difference!)

1. Assign each risk allele a weight based on OR
2. Sum weights across all risk alleles per person (= “genetic risk score”)
3. Compare distribution of weighted GRS in cases vs controls
4. Compare GWAS GRS *vs* EMR GRS

… across all ethnic groups



Outline of talk today

- Demonstration: developing an algorithm to define an RA cohort, proof-of-concept genomic studies

- Portability: implementing the EMR classification algorithm at other institutions

- Application: defining subsets of patients with clinically-relevant outcomes – and cardiovascular disease in particular

Portability to other institutions

BRIGHAM AND WOMEN'S HOSPITAL *
MASSACHUSETTS GENERAL HOSPITAL *
FAULKNER HOSPITAL
NEWTON-WELLESLEY HOSPITAL
McLEAN HOSPITAL *
PARTNERS CONTINUING CARE
MGH INSTITUTE OF HEALTH PROFESSIONS
PARTNERS COMMUNITY HEALTHCARE, INC.
NORTH SHORE MEDICAL CENTER
MARTHA'S VINEYARD HOSPITAL
NANTUCKET COTTAGE HOSPITAL

VANDERBILT

PGRN Pharmacogenomics Research Network

Good portability to other institutions

Institution	PPV (SE)	Sensitivity (SE)
Partners	*0.93 (0.02)*	*0.63 (0.06)*
Northwestern	*0.80 (0.02)*	*0.50 (0.05)*
Vanderbilt	*0.92 (0.02)*	*0.54 (0.05)*

Note: it took us 2 years to develop the algorithm at Partners, but ~2 months to apply it at Northwestern/Vanderbilt. *Still, this needs to be faster (e.g., 2 minutes!)*

Outline of talk today

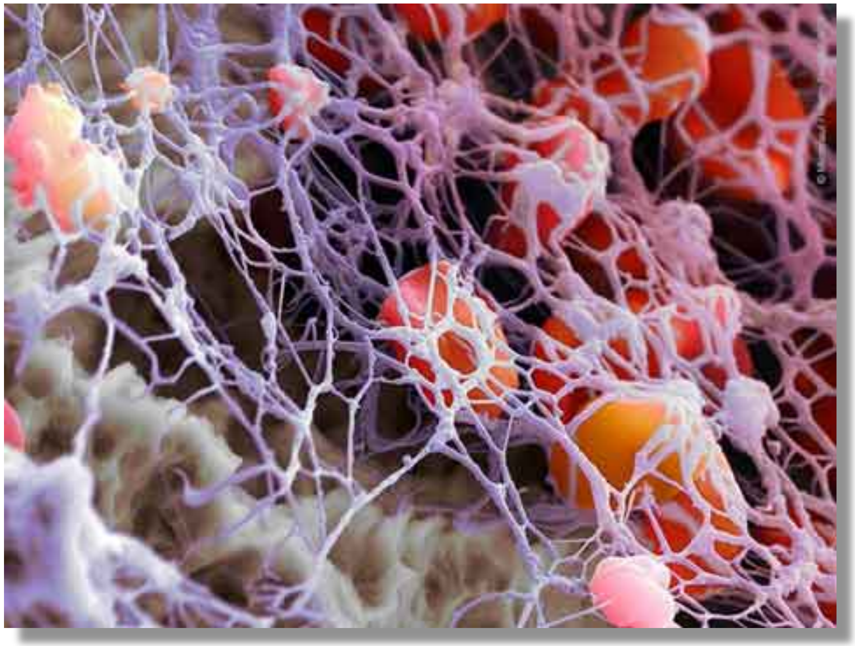
- Demonstration: developing an algorithm to define an RA cohort, proof-of-concept genomic studies

- Portability: implementing the EMR classification algorithm at other institutions

- Application: defining subsets of patients with clinically-relevant outcomes – and cardiovascular disease in particular

Clinically relevant subsets of RA

cardiovascular disease



i2b2

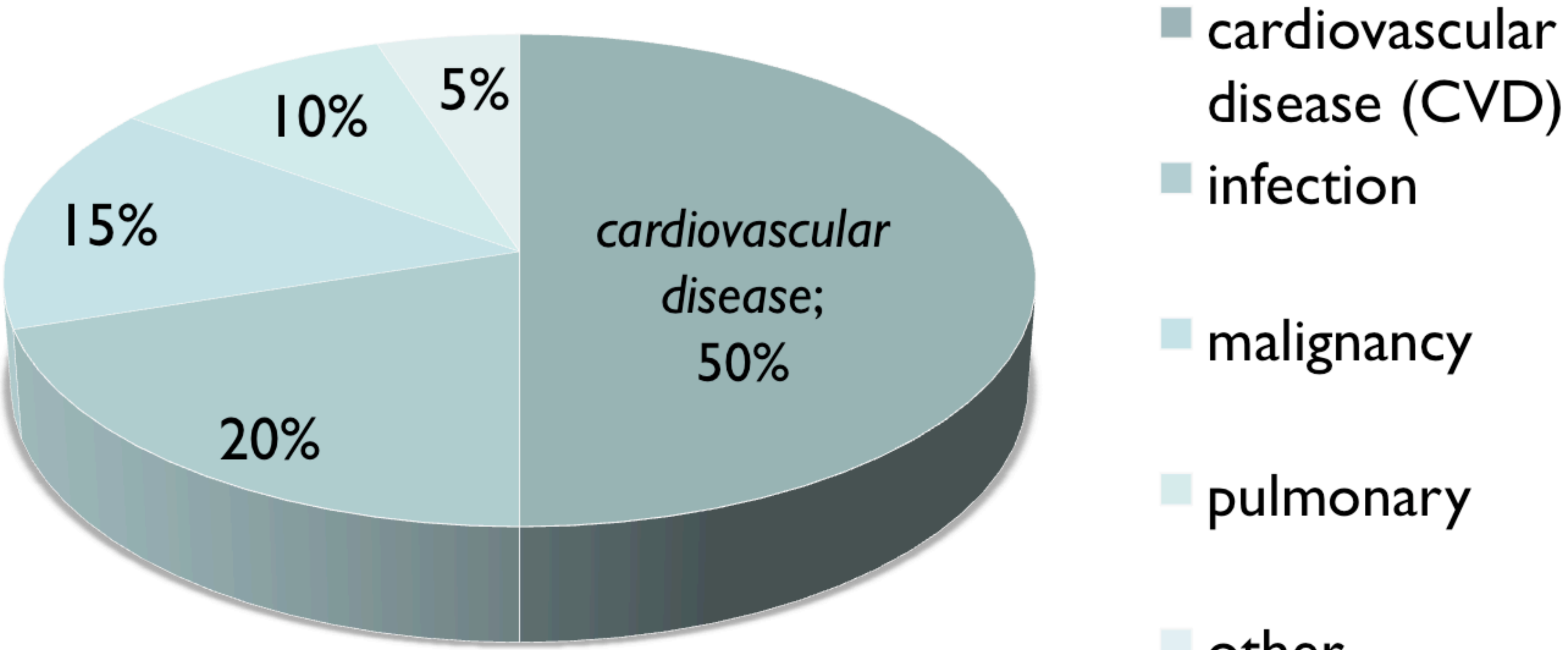
Informatics for Integrating Biology & the Bedside

response to therapy

PGRN

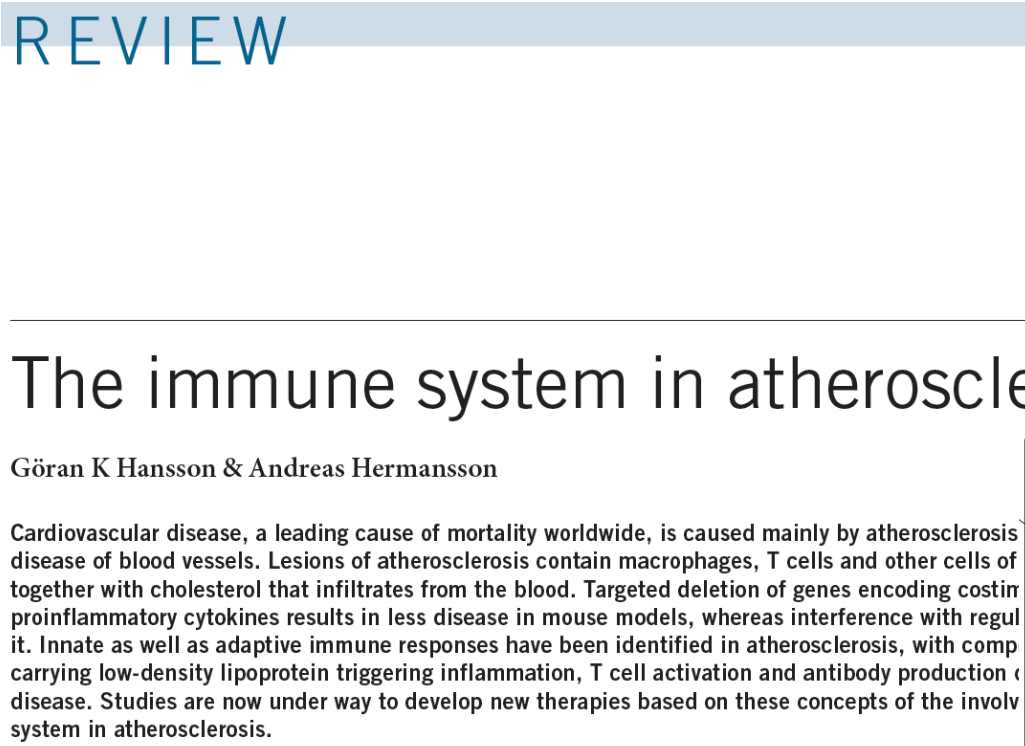
Pharmacogenomics
Research Network

Subset patients in clinically meaningful ways: *causes of mortality*



There is a 2-fold increased risk of CVD in RA patients…is this due to inflammation?

Link between CVD and inflammation



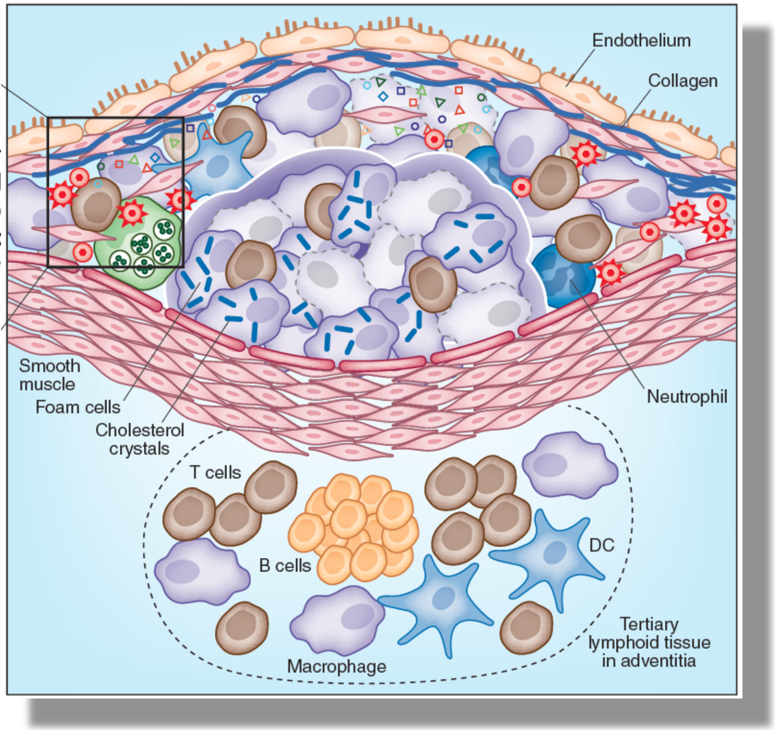
REVIEW

nature immunology

The immune system in atherosclerosis

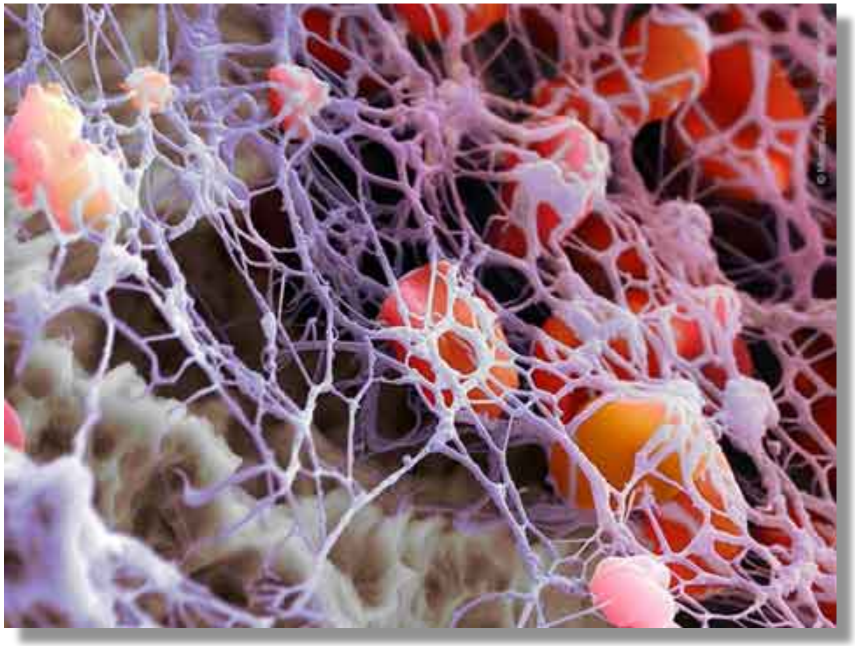
Göran K Hansson & Andreas Hermansson

Cardiovascular disease, a leading cause of mortality worldwide, is caused mainly by atherosclerosis disease of blood vessels. Lesions of atherosclerosis contain macrophages, T cells and other cells of together with cholesterol that infiltrates from the blood. Targeted deletion of genes encoding costim proinflammatory cytokines results in less disease in mouse models, whereas interference with regul it. Innate as well as adaptive immune responses have been identified in atherosclerosis, with comp carrying low-density lipoprotein triggering inflammation, T cell activation and antibody production c disease. Studies are now under way to develop new therapies based on these concepts of the involv system in atherosclerosis.



Work in progress: *model of CVD in RA*

cardiovascular disease



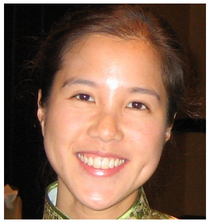
genetics + autoAbs

↓ ???

CVD

i2b2

Informatics for Integrating Biology & the Bedside



Clinical characteristics of CVD in our EMR RA cohort

Characteristic	CAD yes, n=335 (7.5%)	CAD no, n=4118 (92.5%)	P-value
Age, years, mean (SD)	72.9 (10.2)	60.0 (14.7)	<0.0001
Female gender, n (%)	207 (5.9)	3316 (94.1)	<0.0001
Male gender, n (%)	128 (13.8)	802 (68.2)	
Race- white, n (%)	265 (79.0)	2714 (91.1)	
Seropositive, n (%)	87 (67.4)	1099 (60.2)	0.10
MTX, n (%)	158 (47.2)	1851 (45.0)	0.45
TNFi, n (%)	96 (28.7)	1189 (28.8)	1.0
Plaquenil, n (%)	101 (30.2)	1200 (29.1)	0.71
CRP mean, median (mg/L)	10.2, 4.2	7.9, 2.0	<0.0001
ESR_mean (mm/hr)	36.5	26.2	<0.0001
Erosions, n (%)	206 (61.5)	2168 (52.6)	0.0021
HTN, n (%)	252 (75.2)	1160 (28.2)	<0.0001
DM, n (%)	108 (32.2)	375 (9.1)	<0.0001
Hyperlipidemia, n (%)	214 (63.9)	817 (19.8)	<0.0001

Clinical characteristics of CVD in our EMR RA cohort

Characteristics	OR (95% CI)
Age	1.06 (1.05, 1.08)
Female gender	0.35 (0.27, 0.46)
HTN	2.64 (1.88, 3.72)
DM	1.64 (1.20, 2.23)
Hyperlipidemia	2.86 (2.10, 3.90)
Ever smoker	2.30 (1.73, 3.04)

Conclusions

EMRs for discovery research

design	Clinical data	DNA	Sample size	cost
clinical trial	+++	+++	+	$$$
registry	++	+++	++	$$
claims data	+	n/a	+++	$
EMR	++	+++	+++	$

Conclusion: Informatics methods can yield accurate clinical data.

EMRs for discovery research

design	Clinical data	DNA	Sample size	cost
clinical trial	+++	+++	+	$$$
registry	++	+++	++	$$
claims data	[illegible]	[illegible]	[illegible]	$
EMR	++	+++	+++	$

Conclusion: EMR-based biorepositories for genetic studies yield effect sizes similar to traditional cohorts.

EMRs for discovery research

design	Clinical data	DNA	Sample size	cost
clinical trial	+++	+++	+	$$$
registry	++	+++	++	$$
claims data				$
EMR	++	+++	+++	$

Conclusion: It should be possible to extend this same framework to a multitude of other phenotyes across multiple institutions, **but**...

Of course, this is not the only way

- This approach will be good for some applications, and not good for others.
- This may serve as an effective way to generate hypotheses.
- There will always be a role for traditional registries.

i2b2 and PGRN acknowledgments

Zak Kohane
Susanne Churchill
Vivian Gainer
Kat Liao
Tianxi Cai
Shawn Murphy
Beth Karlson
Raul Guzman-Perez
Qing Zing
Pete Szolovits
Lee-Jen Wei
Lynn Bry (Crimson)
Ashwin Ananthakrishnan
Barbara Mawn
Zongqi Xia
Phil De Jager
& many others !

Josh Denny
Abel Kho
Will Thompson
Richard Pope
Anne Eyler
Chad Boomershine
Eric Ruderman
Art Mandelin
Tom Thomas
Robert Carroll

& others !

Funding

Within Our Reach

Finding a Cure for Rheumatoid Arthritis

AMERICAN COLLEGE OF RHEUMATOLOGY
RESEARCH AND EDUCATION FOUNDATION

NIAMS

PGRN

NIGMS

BURROUGHS
WELLCOME
FUND

i2b2

Informatics for Integrating Biology & the Bedside

rplenge @ partners.org