Progress Toward Precision Medicine and the Challenges of Integrating Genomics into Electronic Health Records.

Rex L. Chisholm, PhD
Adam and Richard T. Lind Professor of Medical Genetics
Vice Dean for Scientific Affairs
Feinberg School of Medicine
Associate Vice President for Research
Northwestern University
Critical factors for Defining Genetic Contributions to Disease: A key to Precision Medicine

• Methods to measure genetic variation in an individual human genome
• *Large numbers of well phenotyped human genomes*
• Development of standards of care / best practices
• Methods to deploy genomics based decision support
Northwestern’s Biobank: NUgene

- Launched in fall 2002
- IRB approved
- Voluntary participation via informed consent
- Collect biological specimens under a broad consent
- One-time questionnaire at enrollment
- Longitudinal medical information captured from EMR
- Secure application and database
- Enables High-throughput phenotyping
- Recontact option for additional research
- Resource facilitating genetic research institution-wide
- Genotype data returned to NUgene
NUgene today

• About 11,000 participants are enrolled
  – Male: 42% Female: 58%
  – Median age: 51
  – Age distribution: 18 - 85+
  – Ethnic breakdown similar to census data for 6-county area
  – Over half enrolled through &/or seen at primary care clinics
  – Average participant has 31 distinct diagnoses (ICD9 non V/E codes), &
    an average of 16 distinct diagnoses assigned at least twice**
  – Average patient followed over 8 years, some patients >20 years

• Overall participation rate is 28%
  – Uptake rate is ~52% if physician mentions the study

• 92% of participants agree to be contacted for future research or
  additional health information

**assigned the same ICD-9 code on 2 or more dates
Data Sources

• Questionnaire (self-report):
  – Completed once, at time of enrollment:
    • Demographic information
    • Environmental exposures
    • Medications
    • Self-reported family and medical history for select conditions

• Electronic billing record data
  – Retrospective and prospective diagnosis (ICD9) & procedure (CPT/ICD9CM) codes

• Electronic medical record data
  – Retrospective and prospective:
    • Medical history and diagnoses
    • Lab tests and results, including pathology reports
    • Medications and therapies
    • Family and social history
    • Free text physician notes
<table>
<thead>
<tr>
<th>ICD9</th>
<th>ICD9 Description</th>
<th># Pts**</th>
<th>% Pts**</th>
</tr>
</thead>
<tbody>
<tr>
<td>401</td>
<td>Essential hypertension</td>
<td>3310</td>
<td>34%</td>
</tr>
<tr>
<td>272</td>
<td>Disorders of lipoid metabolism</td>
<td>3269</td>
<td>34%</td>
</tr>
<tr>
<td>530</td>
<td>Diseases of esophagus</td>
<td>2259</td>
<td>23%</td>
</tr>
<tr>
<td>427</td>
<td>Cardiac dysrhythmias</td>
<td>1825</td>
<td>19%</td>
</tr>
<tr>
<td>451</td>
<td>Phlebitis and thrombophlebitis</td>
<td>1688</td>
<td>17%</td>
</tr>
<tr>
<td>278</td>
<td>Overweight, obesity and other hyperalimentation</td>
<td>1566</td>
<td>16%</td>
</tr>
<tr>
<td>715</td>
<td>Osteoarthritis and allied disorders</td>
<td>1411</td>
<td>15%</td>
</tr>
<tr>
<td>477</td>
<td>Allergic rhinitis</td>
<td>1408</td>
<td>14%</td>
</tr>
<tr>
<td>216</td>
<td>Benign neoplasm of skin</td>
<td>1380</td>
<td>14%</td>
</tr>
<tr>
<td>250</td>
<td>Diabetes mellitus</td>
<td>1367</td>
<td>14%</td>
</tr>
<tr>
<td>311</td>
<td>Depressive disorder, not elsewhere classified</td>
<td>1261</td>
<td>13%</td>
</tr>
<tr>
<td>692</td>
<td>Contact dermatitis and other eczema</td>
<td>1238</td>
<td>13%</td>
</tr>
<tr>
<td>244</td>
<td>Acquired hypothyroidism</td>
<td>1102</td>
<td>11%</td>
</tr>
<tr>
<td>493</td>
<td>Asthma</td>
<td>1043</td>
<td>11%</td>
</tr>
<tr>
<td>366</td>
<td>Cataract</td>
<td>1010</td>
<td>10%</td>
</tr>
</tbody>
</table>

Participant Counts Per Selected Top 3-Digit ICD9s*

*From billing, encounter, problem list, med Hx.
### Top Laboratory Tests within Population

<table>
<thead>
<tr>
<th>Laboratory Tests</th>
<th># Participants*</th>
<th>% Participants*</th>
<th># of Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>8,509</td>
<td>88</td>
<td>265,963</td>
</tr>
<tr>
<td>Creatinine</td>
<td>8,422</td>
<td>87</td>
<td>124,848</td>
</tr>
<tr>
<td>Calcium</td>
<td>8,392</td>
<td>86</td>
<td>258,308</td>
</tr>
<tr>
<td>Potassium</td>
<td>8,387</td>
<td>86</td>
<td>265,022</td>
</tr>
<tr>
<td>Sodium</td>
<td>8,384</td>
<td>86</td>
<td>262,436</td>
</tr>
<tr>
<td>Chloride</td>
<td>8,384</td>
<td>86</td>
<td>261,752</td>
</tr>
<tr>
<td>Mean Corpuscular Volume (MCV)</td>
<td>8,364</td>
<td>86</td>
<td>247,740</td>
</tr>
<tr>
<td>Red Cell Distribution Width (RDW)</td>
<td>8,364</td>
<td>86</td>
<td>247,733</td>
</tr>
<tr>
<td>Mean Corp. Hemoglobin Conc. (MCHC)</td>
<td>8,364</td>
<td>86</td>
<td>247,740</td>
</tr>
<tr>
<td>Mean Corpuscular Hemoglobin (MCH)</td>
<td>8,364</td>
<td>86</td>
<td>247,740</td>
</tr>
<tr>
<td>Blood Urea Nitrogen</td>
<td>8,366</td>
<td>86</td>
<td>245,748</td>
</tr>
<tr>
<td>Bicarbonate</td>
<td>8,332</td>
<td>86</td>
<td>243,499</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>8,320</td>
<td>86</td>
<td>227,208</td>
</tr>
<tr>
<td>Hematocrit</td>
<td>8,320</td>
<td>86</td>
<td>227,201</td>
</tr>
<tr>
<td>Red Cell Count</td>
<td>8,296</td>
<td>85</td>
<td>224,380</td>
</tr>
</tbody>
</table>
How Can This EHR-linked Biobank be used to discover new gene-disease associations, improve diagnoses and measure therapeutic outcomes?
The eMERGE Network

Electronic Medical Records & Genomics

A consortium of biorepositories linked to electronic medical records data for conducting genomic studies

An NHGRI funded consortium
eMERGE I Goals

Test the ability to leverage EMRs and biobanks for genomic research

− Evaluate validity & utility of EMR phenotypes for Genomics

− Develop & validate electronic phenotyping algorithms

− Conduct association studies of genome-wide data with EMR-derived phenotypes
<table>
<thead>
<tr>
<th>Site</th>
<th>eMERGE Phase I</th>
<th>eMERGE Phase II</th>
<th>eMERGE I &amp; II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Participants</td>
<td>Genotyped</td>
<td>Participants (Still enrolling)</td>
</tr>
<tr>
<td>GHC</td>
<td>2,820</td>
<td>2,789</td>
<td>5,291</td>
</tr>
<tr>
<td>Marshfield</td>
<td>20,000</td>
<td>4,210</td>
<td>20,000</td>
</tr>
<tr>
<td>Mayo</td>
<td>3,769</td>
<td>3,755</td>
<td>6,916</td>
</tr>
<tr>
<td>NU</td>
<td>10,500</td>
<td>1,907</td>
<td>12,000</td>
</tr>
<tr>
<td>VU</td>
<td>70,000</td>
<td>6,055</td>
<td>155,000</td>
</tr>
<tr>
<td>Geisinger</td>
<td>N/A</td>
<td>N/A</td>
<td>22,000</td>
</tr>
<tr>
<td>Mt. Sinai</td>
<td>N/A</td>
<td>N/A</td>
<td>25,000</td>
</tr>
<tr>
<td>CCHMC/BCH</td>
<td>N/A</td>
<td>N/A</td>
<td>11,799</td>
</tr>
<tr>
<td>CHOP</td>
<td>N/A</td>
<td>N/A</td>
<td>60,000</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>107,089</strong></td>
<td><strong>18,716</strong></td>
<td><strong>347,090</strong></td>
</tr>
</tbody>
</table>
Approach to electronic phenotyping

Identify phenotype of interest → Case & control algorithm development and refinement → Manual review; assess precision → Deploy at site 1 → Validate at other sites → Genetic association tests; replicate

PPV < 95% → PPV ≥ 95%

Denny et al., 2011
Type II Diabetes Case Algorithm

- Random glucose > 200 mg/dl, Fasting glucose > 125 mg/dl, or hemoglobin A1c ≥6.5%.
Genomic Analysis Identifies the same genes as Purpose Built Cohorts

Imputed T2D Merged (Case/Ctrl) – 98GE SNPs, Adjusted Sex, Age, BMI, PC1, PC.

TCF7L2
## Phase I Phenotypes

<table>
<thead>
<tr>
<th></th>
<th>GHC/UW</th>
<th>Marshfield</th>
<th>Mayo</th>
<th>Northwestern</th>
<th>Vanderbilt</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dementia</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Cataract</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAD</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Type 2 Diabetes</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>QRS Duration</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Secondary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WBC</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Diabetic Retinopathy</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBC</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Lipids</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Height</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PheWAS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>HDL</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td><strong>Network</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypothyroidism</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Resistant HTN</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
Approach to electronic phenotyping

1. Identify phenotype of interest
2. Case & control algorithm development and refinement
3. Manual review; assess precision
   - PPV < 95%
4. Deploy at site 1
   - PPV ≥ 95%
5. Validate at other sites
6. Genetic association tests; replicate

Denny et al., 2011
Merged Genotype Dataset

• 17,046 eMERGE samples with GWAS data

• Majority of samples genotyped using 660W

• Samples collected for various studies
  GH – Dementia
  Marshfield – Cataracts and HDL-C
  Mayo – PAD
  NW – T2D
  VU – Normal ECGs

Can we use existing dataset for another experiment?
Algorithm for Hypothyroidism

- **No thyroid-altering medications (e.g., Phenytoin, Lithium)**
  - ICD-9s for Hypothyroidism
  - Abnormal TSH/FT4
    - Thyroid replacement medication
    - **No secondary causes** (e.g., pregnancy, ablation)
    - **Case**
  - **Control**

- **2+ non-acute visits**
  - **No ICD-9s for Hypothyroidism**
  - **Normal TSH**
    - **No thyroid replace. meds**
    - **No hx of myasthenia gravis**
    - **Control**
Genomic Analysis

[Graph showing genomic analysis results with chromosome numbers on the x-axis and some form of density or frequency on the y-axis.]
FoxE1 is associated with Hypothyroidism

**GWAS:** Target phenotype

**The phenome-wide association study (PheWAS)**

**PheWAS:** Target genotype

**PheWAS requirement:** A large cohort of patients with genotype data and many diagnoses
PheWAS for rs965513 (FOXE1)

Analysis of 866 phenotypes in 13,617 European Americans
Adjusted for age and sex
eMERGE Phenotyping: Sharable, High-Throughput

• Methods
  - 44 phenotypes (complete or in development)
  - Sharable formats – KNIME, QDM
  - Machine Learning algorithms
  - Portable Methods

• Tools
  - PheKB with data standardization & validation tools
  - eMERGE RC
  - PheWAS
  - Downloadable NLP Tools – cTAKES, MedEx
## Phase II Phenotyping

### Phenotyping Algorithm Development

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Lead Site</th>
<th>Secondary Site</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clostridium difficile</td>
<td>GroupHealth</td>
<td>Vanderbilt</td>
<td>Complete</td>
</tr>
<tr>
<td>Abdominal aortic aneurysm</td>
<td>Geisinger</td>
<td>Mayo</td>
<td>Complete</td>
</tr>
<tr>
<td>Venous Thromboembolism</td>
<td>Mayo</td>
<td>Vanderbilt</td>
<td>Complete</td>
</tr>
<tr>
<td>Ocular Hypertension</td>
<td>MC/EIRH/PSU</td>
<td>Geisinger</td>
<td>Complete</td>
</tr>
<tr>
<td>Diverticulosis</td>
<td>Northwestern</td>
<td>Vanderbilt</td>
<td>Complete</td>
</tr>
<tr>
<td>Glaucoma</td>
<td>MC/EIRH/PSU</td>
<td>Geisinger</td>
<td>Complete</td>
</tr>
<tr>
<td>Herpes Zoster</td>
<td>GroupHealth</td>
<td>Vanderbilt</td>
<td>Complete</td>
</tr>
<tr>
<td>ACE-Inhibitor Induced Cough</td>
<td>Vanderbilt</td>
<td>Northwestern</td>
<td>Complete</td>
</tr>
<tr>
<td>Cardio Respiratory Fitness</td>
<td>Mayo</td>
<td>Geisinger</td>
<td>Complete</td>
</tr>
<tr>
<td>Extreme Obesity</td>
<td>Geisinger</td>
<td>MC/EIRH/PSU</td>
<td>Complete</td>
</tr>
<tr>
<td>Asthma</td>
<td>CHOP</td>
<td>MC/EIRH/PSU</td>
<td>Complete</td>
</tr>
<tr>
<td>Child Obesity</td>
<td>CCHMC/BCH</td>
<td>CHOP</td>
<td>Complete</td>
</tr>
<tr>
<td>Heart Failure</td>
<td>Mayo</td>
<td>GroupHealth</td>
<td>In process (exp. Jan. 2014)</td>
</tr>
<tr>
<td>Colon Polyps</td>
<td>Northwestern</td>
<td>MC/EIRH/PSU</td>
<td>In process (exp. Jan. 2014)</td>
</tr>
<tr>
<td>Autism</td>
<td>CCHMC</td>
<td>BCH</td>
<td>In process (exp. Jan. 2014)</td>
</tr>
<tr>
<td>Statins for MACE</td>
<td>Vanderbilt</td>
<td>MC/EIRH/PSU</td>
<td>In process (exp. Jan. 2014)</td>
</tr>
<tr>
<td>Age-related Macular Degeneration</td>
<td>MC/EIRH/PSU</td>
<td>Northwestern</td>
<td>In process (exp. Jan. 2014)</td>
</tr>
<tr>
<td>Atopic Dermatitis</td>
<td>CHOP</td>
<td>MC/EIRH/PSU</td>
<td>In process (exp. Jan. 2014)</td>
</tr>
<tr>
<td>Remission of Diabetes after ROUX-EN-Y</td>
<td>CHOP</td>
<td>MC/EIRH/PSU</td>
<td>In process (exp. Feb. 2014)</td>
</tr>
<tr>
<td>CAAD as Quantitative Measure</td>
<td>Geisinger</td>
<td>Northwestern</td>
<td>In process (exp. Feb. 2014)</td>
</tr>
</tbody>
</table>

**Upcoming Phenotypes:** Upper GI/PUD, GERD, Appendicitis, Epilepsy, Lipids, Pulmonary HTN, Diabetic Hypertensive CKD, Rapid Renal Decline in Diabetic HTN Nephropathy, caMRSA, ADHD
• View existing algorithms
• Enter or create new algorithms
• Collaborate to create or review algorithms
• View implementation details for existing algorithms
eMERGE II phenotyping: Lower GI Phenotypes

Colon Polyps

Diverticulosis

© 2004 MedicineNet, Inc.
Colon Polyp NLP Algorithm

1. ±Colon Polyps: checks EHR for colonoscopy with linked path report w. polyp mention

2. Type + Location: NLP on path reports to extract type and location

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity</th>
<th>PPV</th>
<th>Specificity</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>±Colon Polyps</td>
<td>98%</td>
<td>94%</td>
<td>94%</td>
<td>98%</td>
</tr>
<tr>
<td>Type + Location</td>
<td>96%</td>
<td>98%</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>
Colon Polyp Findings By Location

All Colon Polyps: Count per Location

- Hepatic Flexure: 1257
- Transverse: 3522
- Splenic Flexure: 574
- Cecum: 4629
- Ascending: 2374
- Ileum: 259
- Descending: 2178
- Sigmoid: 4906
- Rectum: 4454

Adenoma Count Per Location

- Hepatic Flexure: 904
- Transverse: 2590
- Splenic Flexure: 403
- Cecum: 3423
- Ascending: 1844
- Ileum: 181
- Descending: 1490
- Sigmoid: 2449
- Rectum: 1511

Northwestern Medicine
Returning Genomic Results to the EHR

• How to store...PDF inadequate
• What to store
  – Only variants of medical significance
  – Actionable
  – Evidence changes over time
• Clinical Decision support
  – Sharable Decision Logic
  – Patient facing supporting material
Data from CLIA Labs

<table>
<thead>
<tr>
<th>MRN</th>
<th>CYP2C9</th>
<th>SLC01B1</th>
<th>VKORC1</th>
<th>CYP2C19</th>
</tr>
</thead>
<tbody>
<tr>
<td>12345</td>
<td>*1/*2</td>
<td>T/C</td>
<td>A/A</td>
<td>*1/*1</td>
</tr>
<tr>
<td>12346</td>
<td>*2/*3</td>
<td>T/T</td>
<td>A/G</td>
<td>*1/*17</td>
</tr>
</tbody>
</table>

... 

- Want to preserve the original data
- Want to let physicians make sense of it
- Want to let patients make sense of it
Ancillary Genomic System

- Entity-Attribute-Value (EAV) schema

<table>
<thead>
<tr>
<th>ID</th>
<th>PatientID</th>
<th>ParentID</th>
<th>AttrID</th>
<th>AttrName</th>
<th>Value1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>H0150629</td>
<td>NULL</td>
<td>5</td>
<td>CYP2C9 Diploype</td>
<td>*1/*1</td>
</tr>
<tr>
<td>2</td>
<td>H0150629</td>
<td>NULL</td>
<td>13</td>
<td>CYP2C19 Diploype</td>
<td>*1/*2</td>
</tr>
<tr>
<td>3</td>
<td>H0150629</td>
<td>NULL</td>
<td>9</td>
<td>VKORC1 Genotype</td>
<td>A/A</td>
</tr>
<tr>
<td>4</td>
<td>H0150629</td>
<td>NULL</td>
<td>14</td>
<td>rs12248560</td>
<td>C</td>
</tr>
<tr>
<td>5</td>
<td>H0150629</td>
<td>NULL</td>
<td>15</td>
<td>rs28399504</td>
<td>AG</td>
</tr>
<tr>
<td>6</td>
<td>H0150629</td>
<td>NULL</td>
<td>16</td>
<td>rs41291556</td>
<td>T</td>
</tr>
<tr>
<td>7</td>
<td>H0150629</td>
<td>1</td>
<td>18</td>
<td>Value Observation</td>
<td>Inferred</td>
</tr>
<tr>
<td>8</td>
<td>H0150629</td>
<td>2</td>
<td>18</td>
<td>Value Observation</td>
<td>Direct</td>
</tr>
</tbody>
</table>

- Create computed observations

**Clopidogrel Metabolism:** Poor Metabolizer (Predicted)

**Simvastatin Metabolism:** Normal Metabolizer (Predicted)
<table>
<thead>
<tr>
<th>Creation Date</th>
<th>Patient Name</th>
<th>DOB</th>
<th>MRN</th>
<th>Test Name</th>
<th>Results</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>12/31/2013 10:44:20 AM</td>
<td>Zztest, Hamish</td>
<td>12/26/1961</td>
<td>100000084</td>
<td>Simvastatin Metabolism</td>
<td>Normal Activity (Predicted)</td>
<td>Details</td>
</tr>
<tr>
<td>1/26/2014 10:01:56 PM</td>
<td>Zztest, Hamish</td>
<td>12/26/1961</td>
<td>100000084</td>
<td>Warfarin Dosing</td>
<td></td>
<td>Details</td>
</tr>
</tbody>
</table>

**AGS Interpretation - Report Preview**

**Warfarin Dosing**

- **Value:**
  - Created: 01/26/2014 22:01:56
- **Report:**
  - RESULT
  - CYP2C9 *1/*2
  - VKORC1 (rs9923231) G/G

Use specific information below to determine the appropriate starting dose at http://www.warfarindosing.org.

- VKORC1-1639/3673: GG (warfarin insensitive)
- CYP2C9*2: CT (heterozygous)
- CYP2C9*3: AA (wildtype)
- CYP2C9*5: CC (wildtype)
- CYP2C9*6: AA (wildtype)

**INTERPRETATION**

Patient carries one active and one reduced activity CYP2C9 allele and, therefore, is expected to be able to metabolize medications via CYP2C9 less effectively. Intermediate metabolizers may require non-conventional doses of medications whose major metabolic pathway is CYP2C9 or use of another drug that is not processed by CYP2C9.
Clinical Decision Support

**Patient on clopidogrel, but genetic results indicate patient may be a poor metabolizer**

Medication may be ineffective – consider alternative.

- Acknowledge Reason:
- Discussed results with patient

- Open SmartSet: Patient education (After Visit Summary)
- Click to review medications
- View clinical references related to results
- View patient materials related to results

---

**MyResults.org**

**Clopidogrel (Plavix)**

- Simvastatin
- Tegretol (Carbamazepine)
- Warfarin (Coumadin)

**Overview**

- What is Clopidogrel?
  - Clopidogrel is also called Plavix. It is a drug used by doctors to treat or prevent strokes and reducing the risk of death by preventing the blood from clotting so that it flows easier through the body. Clopidogrel is sometimes used with aspirin to lower the chance of heart attack, stroke, or blood clots.

- What is being tested?
  - People react differently to medicine and some of those different reactions can be related to the effects of their genes. People who are poor metabolizers may not respond to particular medications as other people. The gene called, CYP2C19. This test will look for some of the genetic differences in the CYP2C19 gene.

- How will this affect my health care?
  - People who are poor metabolizers might be less responsive to clopidogrel. Depending on results, doctors might need to adjust your dose of clopidogrel.
About These Results

Many things can explain why a person has a medical condition, or why different people respond to the same medication in different ways. Genetic testing, which looks for changes (also called polymorphisms, or mutations) in your DNA, can help your doctor determine if you are at risk for getting a condition, or which medication you might need.

It’s important to know that being at risk for a condition doesn’t mean you will necessarily get it. The results of these tests should be used with other pieces of evidence collected by your doctor to make any type of medical decision.

Hemochromatosis

Results
C282Y Homozygote
Tested on 3/15/2012

What does this mean?
You may be at risk for developing a condition known as hemochromatosis. Your doctor can perform additional tests to see if you currently have this condition.

Clopidogrel (Plavix) Metabolism

Results
CYP2C19*2 Homozygote
Tested on 5/17/2011

What does this mean?
Your body may not get any benefit from the drug clopidogrel (also known as Plavix). If you are currently on this drug, you should contact your doctor to discuss if an alternative drug may be right for you. If you are not currently taking Plavix, it is important to know about this in case your doctor needs to prescribe it in the future.
eMERGE OnLine Resources

eMERGE is a national network organized and funded by the National Human Genome Research Institute (NHGRI) that combines DNA biorepositories with electronic medical record (EMR) systems for large scale, high-throughput genetic research in support of implementing genomic medicine.

POPULAR TOOLS (CLICK ON A BUTTON BELOW)

- **PheKB**: A knowledgebase for discovering phenotypes from electronic medical records.
- **MyResults.org**: An informational tool for educating patients about genetic test results.
- **SPHINX**: A data exploration tool for genetics-related drug response hypothesis generation.
- **Infobutton Project template**
- **emerge Model Consent Language**
- **PheWAS catalog**
Summary

• Biobanks and EHRs are increasingly playing a critical role in identifying associations between genetic variation, disease risk, drug efficacy and clinical outcomes

• Longitudinal mining of electronic medical records can be used to provide the most up to date phenotype associated with human biospecimens

• Research use can be an important driver of EHR quality

• Networks of EHR-linked biobanks that share samples and data have the potential to increase statistical power to detect genetic associations, population diversity in these studies, and overall research efficiency

• Methods to store genomic variation in EHR will enable personalized medicine
How Can EHR-linked Biobanks be used to discover new gene-disease associations and Improve Disease Models?

How well does this work?
• Pretty well, but could be better with more standardization

What are the issues with this approach?
• Heavy reliance on NLP
• Poor consistency between different clinical sites, even for coded data
How Can EHR-linked Biobanks be used to discover new gene-disease associations and improve Disease Models?

- **What can be done to improve the quality of phenotypes?**
  - Broader adoption of standards
  - Include research uses in meaningful use criteria

- **How can this inform disease models?**
  - Phenotype algorithms should also help inform disease models—integrate lessons learned from the “interactive” phenotyping process.
Acknowledgements

• NUgene Governance Committee & Community Advisory Committee

• NUgene team:
  – Rex Chisholm, PhD (PI)
  – Warren Kibbe, PhD (Co-founder)
  – William Lowe, MD (Medical Dir.)
  – Wendy Wolf, PhD
  – Maureen Smith, MS, CGC
  – Jennifer A. Pacheco
  – Tony Miqueli
  – Sharon Aufox, MS, CGC
  – Oana Popescu
  – Nicole Sheehan
  – Noah Goss
  – Maribeth Miceli

• More information about NUgene: [http://www.nugene.org](http://www.nugene.org)

• NU eMERGE team:
  – Rex Chisholm
  – Phil Greenland
  – Abel Kho
  – Bill Lowe
  – Wendy Wolf
  – Maureen Smith
  – Geoff Hayes
  – Loren Armstrong
  – Doug Scheftner
  – Laura Rasmussen-Torvik
  – Pedro Avila
  – Amy Lemke
  – May Law
  – Jen A. Pacheco
  – Joel Humowiecki
  – Will Thompson
  – Arun Muthalagu
  – Krishna Dornadula

• Funding for eMERGE studies is provided by NHGRI (U01HG004609)