Topics

- Background
- Project Overview
- Use Case
- Methodology
- Research Design
- Results to-date
- Next Steps
Background

• CENTERPHASE SOLUTIONS, INC. is an innovative, technology-driven company formed through a collaboration with Mayo Clinic. Our goal is to leverage medical data and clinical expertise from a network of academic medical centers to address a broad array of healthcare issues and opportunities.

• Our initial focus is to help our customers make better decisions for planning, evaluating and executing clinical trials.

Future areas include comparative effectiveness, outcomes research, compliance, retrospective analyses, epidemiological studies.

• Our markets:
  - Biopharmaceutical Industry
  - Government
  - Insurance
  - Academic Medical Centers
  - Foundations
  - Private equity/venture capital
Project Overview

Objectives: Compare the cost-effectiveness of phenotyping algorithms against a manual process for patient cohort identification

Approach:
1. Choose a use case (disease) that can provide valuable insights into a real world application
2. Create a phenotyping methodology (inclusion/exclusion “work flow”) to identify the patient cohort
3. Generate a random sample of patients with the given diagnosis, and conduct algorithm-driven and manual processes (study coordinator) in parallel
4. Compare the time, cost and accuracy of results from the algorithm-driven to manual processes
Initial Use Case

• Type II Diabetes Mellitus (T2DM)
• Multi-stage phenotype, representing a combined adaptation of:
  – The eMERGE Northwestern T2DM algorithm for clinical trial selection and
  – The group practice reporting options (GPRO) as defined under NCQA for population management under the Southeast Minnesota Beacon project
Phenotype Methodology

EMERGE Algorithm for T2DM

Screen 1: Age

Screen 2: Medication

Screen 3: Labs & Vitals

Patient Cohort

T2DM ICD9 Code

Identified as T2DM patient

Identified as high risk or “RED” patient

Note: All screens based on two-year measurement period 1/1/09 – 12/31/10
Phenotype Workflow (EMERGE)

Algorithm for the Identification of Subjects with Type 2 Diabetes

Patient Population

T2DM ICD9 code(s)

Treated with insulin medication

Never on T2DM medication

No T1DM ICD9 code & <2 diagnosis^ date

Type 2 Diabetes Cases

On T2DM medication in past

No T1DM ICD9 code & =2 diagnoses^ dates

Treated with T2DM medication

No DM medication but abnormal lab*

Treated with T2DM medication & have an abnormal lab*

No T2DM and T1DM ICD9 codes

This side of workflow excluded from algorithm due to extremely low prevalence of T2DM cases

* Random glucose > 200mg/dl, Fasting glucose > 125 mg/dl, or hemoglobin A1c = 6.5%.

^ Encounter or problem list diagnoses only (all other diagnoses in this chart could also include diagnoses in the medical history)

Note: All screens based on two-year measurement period 1/1/09 – 12/31/10
Phenotype Workflow (Beacon)

Patients with T2DM are categorized into care management populations based upon most recent lab values and vitals during the measurement period. Objective is to identify high risk or RED patients who would require near-term follow-up

Blood Glucose
HbA1c > or = 9

Cholesterol
LDL > 130

Blood pressure
Systolic > 160, Diastolic > 100

If ANY of the most recent values exceed allowable levels OR ANY of these elements has not been captured in the measurement period, patient is classified as high risk or RED

Note: All screens based on two-year measurement period 1/1/09 – 12/31/10
Randomly generate **ONE** sample set of patient records from database:
Based on ICD9 Codes with at least 2 diagnosis dates

**Research Design**

Study coordinator (SC) conducts manual review of patient charts, and monitors activity time

Programmer develops and runs algorithm to query records, and monitors development and run time

Compare time, cost and accuracy of results
# Data Capture Form

## SCREEN 1

<table>
<thead>
<tr>
<th>Inclusion</th>
<th>Time to exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inclusion</td>
<td>No</td>
</tr>
</tbody>
</table>

**Data**
- Patient is 75 years of age or younger as of 12/31/11: **Yes**
- Excluded, log time from opening chart to exclusion: **1.00**
- Currently prescribed only T2DM medication AND previously prescribed T2DM medication during measurement period: **Yes**
- Did patient meet T2DM medication AND previously prescribed T2DM medication criteria? If answered YES, patient is **include**.

## SCREEN 2

<table>
<thead>
<tr>
<th>Inclusion</th>
<th>Inclusion</th>
<th>Inclusion</th>
<th>Inclusion</th>
<th>Time to exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inclusion</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

**Data**
- Currently prescribed only T2DM medication AND previously prescribed T2DM medication during measurement period: **Yes**
- Did patient meet T2DM medication AND previously prescribed T2DM medication criteria? If answered YES, patient is **include**.

## SCREEN 3

<table>
<thead>
<tr>
<th>Inclusion</th>
<th>Inclusion</th>
<th>Inclusion</th>
<th>Inclusion</th>
<th>Time to exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inclusion</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

**Data**
- Currently prescribed only T2DM medication OR previously prescribed T2DM medication during measurement period: **Yes**
- Did patient meet T2DM medication OR previously prescribed T2DM medication criteria? If answered YES, patient is **include**.

## Log time from opening chart to completion of chart review

- Chart 1: 11
- Chart 2: 10
- Chart 3: 10
- Chart 4: 10
- Chart 5: 10
- Chart 6: 10
- Chart 7: 8
- Chart 8: 8
- Chart 9: 12
- Chart 10: 12
- Chart 11: 12
- Chart 12: 10
- Chart 13: 10
- Chart 14: 10
- Chart 15: 10
- Chart 16: 6
- Chart 17: 8
- Chart 18: 8
- Chart 19: 0.30
- Chart 20: 0.30

**Total Include**: 15
**Total Excluded**: 5
**Total Screen 1**: 20

**Total Include**: 11
**Total Excluded**: 4
**Total Screen 2**: 15

**Total Include**: 9
**Total Excluded**: 8
**Total Screen 3**: 11
Validation Process

Stage 1
- 20 Charts
  - Review each chart for manual and algorithm processes to identify any screening errors
  - Confirm approaches are consistent
  - Refine procedures as appropriate

Stage 2
- 50 Charts
  - Compare patient result set for manual and algorithm process
  - Identify root causes of differences
  - Refine as necessary
  - Validation complete

Stage 3
- 500 Charts
  - “Live run”
  - Conduct data queries
  - Collect time, cost and patient result sets
  - Analyze results and evaluate / compare performance of methods
Results to-Date

Stage 1

20 Charts

Manual Process

• Initially identified 4 REDs
  Charts 1, 3, 8, 11
• After detailed review, determined SC had incorrectly included chart 11 based on lab results
  Final REDs: 1, 3, 8
• Time:
  Initial: 2.3 hours
  Validation: 0.5 hours
  Total: 2.8 hours

Algorithm-Driven Process

• Initially identified 4 REDs
  Charts 1, 3, 8, 20
• After detailed review, determined algorithm had incorrectly included chart 20 based on measurement period
  Final REDs: 1, 3, 8
• Time:
  Development: 5.0 hours*
  Run time: 8 seconds
  Validation: 0.5 hours
  Total: 5.5 hours

* Will need to ensure time includes NLP development
Cost per Patient Analysis

<table>
<thead>
<tr>
<th>Manual Chart Review</th>
<th>Algorithm-Driven Process</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Hours</strong></td>
<td>2.8</td>
</tr>
<tr>
<td><strong>Total Cost</strong></td>
<td>$242</td>
</tr>
<tr>
<td><strong>Number of RED Patients</strong></td>
<td>3.0</td>
</tr>
<tr>
<td><strong>Cost Per Patient</strong></td>
<td>$81</td>
</tr>
<tr>
<td><strong>Total Hours</strong></td>
<td>5.5</td>
</tr>
<tr>
<td><strong>Total Cost</strong></td>
<td>$916</td>
</tr>
<tr>
<td><strong>Number of RED Patients</strong></td>
<td>3.0</td>
</tr>
<tr>
<td><strong>Cost Per Patient</strong></td>
<td>$305</td>
</tr>
</tbody>
</table>

- "Manual Hours Applied"
- "Algorithm Hours Applied"

- "Manual Cost Per Patient"
- "Algorithm Cost Per Patient"
Preliminary Findings

- High level of consistency in first ‘experiment’
- Able to identify errors and refine both processes
- For 20 charts, algorithm-development time exceeds manual review time and cost
- Due to query run time, extrapolation of results to 500 charts likely to show significant time and cost efficiencies
Next Steps / Preliminary Timelines

- Begin Stage 2 this week: 50 charts
  - Complete analysis and review by late June/early July
  - Make refinements as necessary
- Initiate Stage 3 in July/August: 500 charts
  - Complete analysis and review in August/September
  - Prepare white paper/manuscript by September/October