Data and Analytics: Preparing for the Future

Gregory Warren, FSA, MAAA, FCA
Vice President, Actuarial Consulting
Optum Consulting
Experience ‘in the room” where decisions are made

We offer unique knowledge of the actuarial culture, roles, language, methods, and value drivers within various health care, risk-bearing entities.

Our offerings are ultimately designed to enhance the relevance and impact of strategies and tactics designed for risk-bearing stakeholders.
Actuarial Science and Health Economics & Outcomes Research

Building Bridges: Common Foundations, Divergent Approaches and Applications

Common Foundations
- Calculus-Based Statistical Theory
- Measuring Results
- Health Economic Impact
- Dealing with Uncertainty
- Model-Building Experts

Divergent Approaches and Applications
- Law of Large Numbers (minimize statistical variation)
- Estimate Confounding Factors
- Financial Outcomes
- Book of Business Focus
- Identify Correlations
- Short/Intermediate-Term Horizons
- Characteristic-Matched Studies (minimize confounding factors)
- Eliminate Confounding Factors
- Clinical & Economic Outcomes
- Disease Focus
- Identify Causations
- Intermediate/Long-Term Horizons
Optum actuarial solutions are integrated to an interdisciplinary approach that aligns strategy, value and execution across the product lifecycle.
The Optum integrated strategic approach produces tangible *actuarial* deliverables for use in *price, value, and access* optimization.

**MARKET ACCESS STRATEGY**

**Define value landscape**
- Competition
  - Competitive analysis
  - Patient sub-population analysis
  - Provider/patient preference analysis
  - Market research
  - Evidence map

**Reimbursement**
- Global payer review
- HTA assessment
- Treatment pathways
- Site of care analysis
- Benchmark reimbursable file

**Develop value strategy**
- Target value profile
  - Value hypotheses
  - *Stakeholder value*
  - Target population
  - Active comparator
  - Target endpoints

**Pricing analysis**
- Pricing value concept
- Early research
- Horizon scanning
- *Market access schemes*
- Value based pricing
- Initial projections

**Quantify economic value**
- Actuarial models
  - Payer pipeline cost forecasts (HTP)
  - Payer formulary design models
  - Payer addressable burden analysis
  - Client ROI models

**Evidence plans**
- Study designs
- Integrated clinical/evidence plans
- HEOR and PRO plans
- post-launch plans

**Value demonstration**

**Support product launch**
- Launch plans
  - Final value proposition
  - *Reimbursement strategy*
  - Global and local value dossiers
  - Global price corridor and launch sequence

**Post-launch plans**
- Label expansion
- Target populations
- New indications
- Services strategy

**Stakeholder engagement**

**Engage key stakeholders**
- Engagement
  - Stakeholder targeting
  - Hot-button analysis
  - Message development
  - Regulatory review
  - Channel messaging
  - Training
  - Resource planning
  - Integrated value

**Risk-share design**
- Design evaluation
- Financial risk-share design modeling
- Risk-share pilot

**Measure impact and manage**
- Value management
  - Market analytics
  - Value monitoring and communication
  - Stakeholder value map
  - Stakeholder value realization plans
  - Market dynamics and scenario planning
  - Patient/provider services development, piloting and roll-out
  - Risk-share management
Optum actuarial solutions are integrated to align with evidence and outcomes research execution across the product lifecycle.

**EVIDENCE STRATEGY**
- Define evidence landscape
  - Payer addressable burden (risk stack) opportunity value?
  - What is the burden of illness?
  - How does the disease progress and what is its history?
- Develop evidence strategy
  - How do payers think about the financial impact of the disease and how to manage it?
  - Drivers of cost-effectiveness and required efficacy, safety, and price thresholds?
  - What does the disease cost?
- Quantify economic evidence
  - How are payers budgeting costs?
  - How will payers model formulary decisions?
  - Post-launch evidence needs?
  - What is the value and affordability to payers, both local and globally?
  - What are the clinical and economic trade-offs?
- Support product launch
  - How to quantify value in payer language?
  - How to apply evidence to financial value propositions?
  - What evidence to present and how best to present it?
  - What is the impact on payer pharmacy budgets?
- Stakeholder engagement
  - Stakeholder issues/priorities?
  - How to engage payer actuarial/financial stakeholders with evidence?
  - What evidence can best support decisions about risk-share agreements?

**EVIDENCE DEMONSTRATION**
- Define evidence landscape
- Develop evidence strategy
- Quantify economic evidence
- Support product launch
- Stakeholder engagement

**STAKEHOLDER ENGAGEMENT**
- Measure real world impact
  - Actual outcomes and cost of care?
  - Evidence of financial impact of services to stakeholders?
  - How to assess/manage financial risk?
  - How do payers incorporate evidence in their financial models?
  - What is the value and affordability in a real-world setting?

Actuarial modeling and payer financial perspectives incorporated into traditional economic, clinical and pricing approaches.
The Optum integrated strategic approach produces tangible **actuarial deliverables** for use in evidence development and demonstration.

<table>
<thead>
<tr>
<th>Deliverables</th>
<th>Discovery/preclinical</th>
<th>Phase I</th>
<th>Phase II</th>
<th>Phase III</th>
<th>Post launch/phase IV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EVIDENCE STRATEGY</strong></td>
<td>Define evidence landscape</td>
<td>Develop evidence strategy</td>
<td>Quantify economic evidence</td>
<td>Support product launch</td>
<td>Stakeholder engagement</td>
</tr>
</tbody>
</table>
| Competition | • Competitive analysis  
  • Evidence map  
  • Early phase economic models  
  • Burden-of-illness models  
  • Natural history of disease studies | Early phase model and report | Actuarial models  
  • Payer pipeline cost forecasts (HTP)  
  • Payer formulary design models  
  • Client ROI models | Launch plans | Risk-share design  
  • Design evaluation  
  • Financial risk-share design modeling  
  • Risk-share pilot |
| Cost-of-illness model and report | • For publication and value dossier development | Evidence plans | Evidence update strategy | Post-Launch plans | Updated Economic Models |
| Actuarial models | • Payer addressable burden models defining potential for total cost of care savings | Economic Models | • For AMCP dossiers, Global HTA submissions  
  • Cost-effectiveness models  
  • Budget impact models  
  • Risk-benefit models | | Economic models and reports based on real-world data  
  • Submission to drug plans and publications |
| Evidence plans | | | | | |

**Competition**
- Competitive analysis
- Evidence map
- Early phase economic models
- Burden-of-illness models
- Natural history of disease studies

**Early phase model and report**
- Cost-effectiveness with threshold analyses
- For publication and value dossier development

**Actuarial models**
- Payer addressable burden models defining potential for total cost of care savings

**Evidence plans**
- For AMCP dossiers, Global HTA submissions
- Cost-effectiveness models
- Budget impact models
- Risk-benefit models

**Economic Models**
- For AMCP dossiers, Global HTA submissions
- Cost-effectiveness models
- Budget impact models
- Risk-benefit models

**Launch plans**
- Evidence communication strategy

**Post-Launch plans**
- Evidence update strategy

**Risk-share design**
- Design evaluation
- Financial risk-share design modeling
- Risk-share pilot

**Evidence management**
- Updated evidence supporting risk-share management

**Updated Economic Models**
- Submission to drug plans and publications
Actuarial Review of Life Sciences Models

- Formal actuarial review of models to provide:
  - Insight to how actuaries may use the models
  - Detailed observations of model strengths and perceived gaps from an actuarial perspective
  - Recommendations on the potential to modify the existing models to gain more actuarial acceptance

- May be applied to:
  - Budget impact models
  - Cost-effectiveness models
  - Long-range planning financial models

<table>
<thead>
<tr>
<th>Feedback</th>
<th>How to Make Better</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) The BIM assume 12 months per member - actuaries would not accept this assumption</td>
<td>Use actual member months instead of members</td>
</tr>
<tr>
<td>2) Actuaries would find it hard to accept “billed” or “acquisition” costs</td>
<td>Use allowed or paid costs</td>
</tr>
<tr>
<td>3) Health insurance companies consider costs and are structured based on books of business</td>
<td>Show the range results for Medicare, Medicaid, and Commercial populations</td>
</tr>
<tr>
<td>4) The ability to adjust the population are important for actuarial analyses</td>
<td>Add a table showing prevalence, utilization, and costs across enrollee demographics</td>
</tr>
<tr>
<td>5) Trend factors are very important for actuarial analyses</td>
<td>Explicitly include trend factors for prevalence, utilization, and costs</td>
</tr>
</tbody>
</table>
Our actuarial modeling engagements have begun as stand-alone “translational” projects but can now be seen as important sequential phases in assessing, building and delivering financial value propositions to actuarial and financial analytics decision-makers in health plans, PBMAs and ACOs.
Payer Addressable Burden Analyses

Episodes of care define the opportunity for products to “bend the cost curve” based on the total costs of care and the pharmacy versus medical cost mix.

- Costs of Primary Condition
- Total Patient Costs
- Drivers of Cost Changes
- Severity and Costs
Formulary Design Modeling

Actuarial models similar to what is used by payers when determining formulary placement, medical drug benefits, and developing clinical program and utilization management policies. These modeling engagements provide insights to key levers that payers may use to exploit or mitigate financial risks.

### Table: Future State 1

<table>
<thead>
<tr>
<th>Year</th>
<th>Company Direct</th>
<th>Indirect</th>
<th>Overall Market</th>
<th>Company Direct</th>
<th>Indirect</th>
<th>Total</th>
<th>Scripts/1000</th>
<th>Revenue PMPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017</td>
<td>4%</td>
<td>22%</td>
<td>74%</td>
<td>$0.24</td>
<td>$0.24</td>
<td>$0.48</td>
<td>129</td>
<td>$3.77</td>
</tr>
<tr>
<td>2018</td>
<td>5%</td>
<td>24%</td>
<td>71%</td>
<td>$0.42</td>
<td>$4.50</td>
<td>$4.92</td>
<td>144</td>
<td>$3.75</td>
</tr>
<tr>
<td>2019</td>
<td>5%</td>
<td>26%</td>
<td>69%</td>
<td>$0.67</td>
<td>$6.36</td>
<td>$7.03</td>
<td>169</td>
<td>$4.94</td>
</tr>
</tbody>
</table>

### Model Assumptions

1. Eligible population will remain stable throughout years 2016 - 2018
2. Formulary status will remain consistent between 2017 and 2018
3. Four new drugs will enter in 2016 as specialty drugs
4. All drugs require prior authorization
5. New products will be excluded from formularies if the net costs are above market costs
6. Current drugs manufacturers will decrease rebates as exclusive contracts end after 2016
7. Pricing for new products is based on Product 1 AWP for all scenarios
8. The prevalence of Comorbidity X is 15%
9. Annual Utilization Trend – Expected to remain flat
10. Annual Unit Cost Trend – Remains flat due to current high costs of HCV therapies
11. Assumed average treatment of 12 weeks
12. Ultimate share is achieved in 2018
Client Return-On-Investment (ROI) Modeling

- Payer ROI modeling estimates the incremental financial “investment” and “return” of specific products relative to their competitors.
- Models for various client types utilizing “proxy client data” from Optum databases, with sensitivity testing of various population and assumption scenarios

**Risk Share Agreement Analysis**

**Payer ROI Model**

**Inputs**
- Performance Measure
- Distribution
- Gap Method
- Confidence Interval
- Pharmacy Claim Cost
- Member Cost Share
- Formulary Rebate %
- Number of Members

**Calculations**

<table>
<thead>
<tr>
<th>Performance Measure Option</th>
<th>Input</th>
<th>Distribution</th>
<th>Gap Method</th>
<th>Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Option 1</td>
<td>Lognormal</td>
<td>15%</td>
<td>95%</td>
</tr>
</tbody>
</table>

**Outputs**
- Annual Payer ROI with Product: 14% 11% 7%

**Contract guarantee**

**Payer “return”**

**Payer “investment”**

**Payer “ROI”**
Risk Share Strategies and Design

Structures, balancing risks/opportunities, selling, contracting, reconciliation and risk mitigation

1. Actuarially project future episode costs
2. Apply study outcomes to episode costs to calculate guaranteed savings
3. If actual savings are less, reimburse payer
4. If actual savings are more, share in the excess
Evolving Uses of “Real World” Data

David Dore, PharmD, PhD

Vice President, Epidemiology – Optum Life Sciences
Adjunct Assistant Professor – Brown University School of Public Health
Humedica Business Model

Staging Area

- Medications
- Lab Results
- Vital Signs
- Physician Notes
- Diagnoses
- Procedures
- Demographics
- Hospitalizations
- Outpatient Visits

Normalization & De-Identification Processes

Humedica Comprehensive Dataset
TARGETED CONTEXTUALIZATION OF SAFETY SIGNALS
Imagine this is your trial…

A Randomized Trial of Intensive versus Standard Blood-Pressure Control

The SPRINT Research Group*

N ENGL J MED 373;22 NEJM.ORG NOVEMBER 26, 2015
Table 1. Baseline Characteristics of the Study Participants.*

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Intensive Treatment (N = 4678)</th>
<th>Standard Treatment (N = 4683)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion for increased cardiovascular risk — no. (%)†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age ≥75 yr</td>
<td>1317 (28.2)</td>
<td>1319 (28.2)</td>
</tr>
<tr>
<td>Chronic kidney disease‡</td>
<td>1330 (28.4)</td>
<td>1316 (28.1)</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinical</td>
<td>940 (20.1)</td>
<td>937 (20.0)</td>
</tr>
<tr>
<td>Subclinical</td>
<td>779 (16.7)</td>
<td>783 (16.7)</td>
</tr>
<tr>
<td>Framingham 10-yr cardiovascular disease risk score ≥15%</td>
<td>247 (5.3)</td>
<td>246 (5.3)</td>
</tr>
<tr>
<td>Female sex — no. (%)</td>
<td>1684 (36.0)</td>
<td>1648 (35.2)</td>
</tr>
<tr>
<td>Age — yr</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>67.9±9.4</td>
<td>67.9±9.5</td>
</tr>
<tr>
<td>Among those ≥75 yr of age</td>
<td>79.8±3.9</td>
<td>79.9±4.1</td>
</tr>
<tr>
<td>Race or ethnic group — no. (%)‡</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic black</td>
<td>1379 (29.5)</td>
<td>1423 (30.4)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>503 (10.8)</td>
<td>481 (10.3)</td>
</tr>
<tr>
<td>Non-Hispanic white</td>
<td>2698 (57.7)</td>
<td>2701 (57.7)</td>
</tr>
<tr>
<td>Other</td>
<td>98 (2.1)</td>
<td>78 (1.7)</td>
</tr>
</tbody>
</table>
Mimicking the Trial Population

Table 1. SPRINT inclusion criteria. To be eligible, a participant must meet 1, 2, and 3.

General inclusion criteria
1. $\geq$50 years old
2. Systolic blood pressure (SBP)
   - SBP: 130–180 mm Hg on 0 or 1 medication
   - SBP: 130–170 mm Hg on up to 2 medications
   - SBP: 130–160 mm Hg on up to 3 medications
   - SBP: 130–150 mm Hg on up to 4 medications
3. At risk (one or more of the following):
   (a) Presence of clinical or subclinical cardiovascular disease (CVD) other than stroke
      (i) Clinical CVD (other than stroke)
         a. Previous myocardial infarction (MI), percutaneous coronary intervention (PCI), coronary artery bypass grafting (CABG), carotid endarterectomy (CE), carotid stenting
         b. Peripheral artery disease (PAD) with revascularization
         c. Acute coronary syndrome with or without resting ECG change, ECG changes on a graded exercise test (GXT), or positive cardiac imaging study
         d. At least a 50% diameter stenosis of a coronary, carotid, or lower extremity artery
         e. Abdominal aortic aneurysm (AAA) $\geq$5 cm with or without repair
      (ii) Subclinical CVD
         (a) Coronary artery calcium score $\geq$ 400 Agatston units within the past 2 years
         (b) Ankle brachial index (ABI) $\leq$ 0.90 within the past 2 years
         (c) Left ventricular hypertrophy (LVH) by ECG (based on computer reading), echocardiogram report, or other cardiac imaging procedure report within the past 2 years.

Clinical Trials
2014, Vol. 11(5) 532–546
# Mimicking the Trial Population

Table 5. Baseline characteristics of the 9361 SPRINT participants, presented as N (%) or mean (SD).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>SPRINT (N = 9361)</th>
<th>Intensive (N = 4678)</th>
<th>Standard (N = 4683)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (female)</td>
<td>3331 (35.6)</td>
<td>1684 (36.0)</td>
<td>1647 (35.2)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>67.9 (9.4)</td>
<td>67.9 (9.4)</td>
<td>67.9 (9.5)</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>2802 (29.9)</td>
<td>1379 (29.5)</td>
<td>1423 (30.4)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>984 (10.5)</td>
<td>503 (10.8)</td>
<td>481 (10.3)</td>
</tr>
<tr>
<td>White</td>
<td>5399 (57.7)</td>
<td>2698 (57.7)</td>
<td>2701 (57.7)</td>
</tr>
<tr>
<td>Other</td>
<td>176 (1.9)</td>
<td>98 (2.1)</td>
<td>78 (1.7)</td>
</tr>
<tr>
<td>Age ≥75 years</td>
<td>2636 (28.2)</td>
<td>1317 (28.2)</td>
<td>1319 (28.2)</td>
</tr>
<tr>
<td>Age among ≥75 (years)</td>
<td>79.8 (4.0)</td>
<td>79.8 (3.9)</td>
<td>79.9 (4.1)</td>
</tr>
<tr>
<td>Baseline blood pressure (mm Hg)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systolic</td>
<td>139.7 (15.6)</td>
<td>139.7 (15.8)</td>
<td>139.7 (15.4)</td>
</tr>
<tr>
<td>Diastolic</td>
<td>78.1 (11.9)</td>
<td>78.2 (11.9)</td>
<td>78.0 (12.0)</td>
</tr>
<tr>
<td>Baseline chronic kidney disease (estimated glomerular filtration rate &lt;60)</td>
<td>2648 (28.3)</td>
<td>1331 (28.5)</td>
<td>1317 (28.1)</td>
</tr>
<tr>
<td>Serum creatinine (mg/dL)</td>
<td>1.07 (0.34)</td>
<td>1.07 (0.34)</td>
<td>1.08 (0.34)</td>
</tr>
<tr>
<td>Estimated glomerular filtration rate (mL/min/1.73 m²)</td>
<td>71.8 (20.6)</td>
<td>71.8 (20.7)</td>
<td>71.7 (20.6)</td>
</tr>
<tr>
<td>Estimated glomerular filtration rate &lt;45 mL/min/1.73 m²</td>
<td>890 (9.5)</td>
<td>446 (9.5)</td>
<td>444 (9.5)</td>
</tr>
</tbody>
</table>

*Clinical Trials*
2014, Vol. 11(5) 532–546
Baseline characteristics of participants in the Trial and patients in the Trial-like Cohort within the Humedica Research Database (01 January 2008 - 31 December 2013)

<table>
<thead>
<tr>
<th></th>
<th>Trial Participants n ~ 10,000</th>
<th>Trial-like Cohort n ~ 208,672</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed</td>
<td>Standardized</td>
</tr>
<tr>
<td>Age, Mean (SD)</td>
<td>64.3 (7.2)</td>
<td>64.9 (9.7)</td>
</tr>
<tr>
<td>Gender (male), %</td>
<td>64.3</td>
<td>49.7</td>
</tr>
<tr>
<td>Ethnicity, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>12.2</td>
<td>6.3</td>
</tr>
<tr>
<td>Race, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>77.5</td>
<td>76.0</td>
</tr>
<tr>
<td>Asian</td>
<td>9.9</td>
<td>1.8</td>
</tr>
<tr>
<td>Black</td>
<td>8.3</td>
<td>11.1</td>
</tr>
<tr>
<td>Other</td>
<td>4.3</td>
<td>11.1</td>
</tr>
<tr>
<td>Weight, kg, Mean (SD)*</td>
<td>81.8 (21.0)</td>
<td>73.9 (23.5)</td>
</tr>
<tr>
<td>Body mass index, Mean (SD)</td>
<td>32.5 (6.3)</td>
<td>32.9 (7.4)</td>
</tr>
<tr>
<td>Hyperlipidemia, %</td>
<td>77.0</td>
<td>54.0</td>
</tr>
<tr>
<td>Coronary artery disease, %</td>
<td>56.8</td>
<td>19.2</td>
</tr>
<tr>
<td>eGFR, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;30</td>
<td>1.9</td>
<td>1.4</td>
</tr>
<tr>
<td>30-60</td>
<td>19.9</td>
<td>6.0</td>
</tr>
<tr>
<td>60-90</td>
<td>41.3</td>
<td>28.6</td>
</tr>
<tr>
<td>&gt;90</td>
<td>36.9</td>
<td>64.0</td>
</tr>
</tbody>
</table>

*Values edited to protect confidentiality
Does Hypoglycemia Increase Risk of Acute Cardiovascular Events?

CHRONIC DISEASE EPIDEMIOLOGY
First improve capture of hypoglycemia

Prevalence of hypoglycemia diagnoses or mentions, by year and identification method

<table>
<thead>
<tr>
<th>HYPOGLYCEMIA SEVERITY</th>
<th>Any CVD</th>
<th>Acute MI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Full Cohort (n=82,321)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Events, N</td>
<td>39,209</td>
<td>1,649</td>
</tr>
<tr>
<td>Unadjusted RR (95%CI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>2.02 (1.91-2.13)</td>
<td>1.92 (1.52-2.41)</td>
</tr>
<tr>
<td>Mild-moderate</td>
<td>1.55 (1.08-2.21)</td>
<td>1.05 (0.17-6.31)</td>
</tr>
<tr>
<td>Unknown</td>
<td>1.57 (1.54-1.61)</td>
<td>1.35 (1.19-1.52)</td>
</tr>
<tr>
<td>Adjusted* RR (95%CI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>1.69 (1.59-1.79)</td>
<td>1.66 (1.31-6.30)</td>
</tr>
<tr>
<td>Mild-moderate</td>
<td>1.43 (1.00-2.05)</td>
<td>0.96 (0.15-1.35)</td>
</tr>
<tr>
<td>Unknown</td>
<td>1.30 (1.26-1.33)</td>
<td>1.19 (1.06-4.12)</td>
</tr>
<tr>
<td><strong>Restricted Cohort, No baseline CVD (n=57,177)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Events, N</td>
<td>17,179</td>
<td>626</td>
</tr>
<tr>
<td>Adjusted** RR (95%CI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>1.77 (1.61-1.95)</td>
<td>1.18 (0.66-2.12)</td>
</tr>
<tr>
<td>Mild-moderate</td>
<td>1.47 (0.90-2.39)</td>
<td>1.79 (0.48-6.72)</td>
</tr>
<tr>
<td>Unknown</td>
<td>1.48 (1.42-1.54)</td>
<td>1.49 (1.24-1.81)</td>
</tr>
</tbody>
</table>

* Adjusted for age, gender, and prior CVD (any)
** Adjusted for age and gender
Can we identify a cohort of patients with binge eating disorder?
Natural Language Processing for Binge Eating Disorder

- Mapping of observed free text to unique medical concepts

**Observed Text Counts**
- Binge eating: 2,919
- Binge eating disorder: 813
- Binge eat: 226
- Binge-eating: 104
- Binge eats: 80
- Binge eater: 62
- Bing eating: 9
- Binge-eat: 5

**Mapped Text Counts**
- Binge eating: 3,389
- Binge eating disorder*: 829

Nunes et al. ICPE 2015 Abstract #941
## Natural Language Processing for Binge Eating Disorder

### Age at First Observed Diagnosis (years)

<table>
<thead>
<tr>
<th>Age</th>
<th>Full BED Cohort (N=7,115)</th>
<th>Exact &quot;Binge Eating Disorder&quot; Text Cohort (N=497)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>18-29</td>
<td>1,243</td>
<td>17.5</td>
</tr>
<tr>
<td>30-39</td>
<td>1,396</td>
<td>19.6</td>
</tr>
<tr>
<td>40-49</td>
<td>1,505</td>
<td>21.2</td>
</tr>
<tr>
<td>50-59</td>
<td>1,677</td>
<td>23.6</td>
</tr>
<tr>
<td>60+</td>
<td>1,294</td>
<td>18.2</td>
</tr>
</tbody>
</table>

### Sex

<table>
<thead>
<tr>
<th></th>
<th>Full BED Cohort</th>
<th>Exact &quot;Binge Eating Disorder&quot; Text Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Female</td>
<td>5,475</td>
<td>77.0</td>
</tr>
<tr>
<td>Male</td>
<td>1,640</td>
<td>23.0</td>
</tr>
</tbody>
</table>

### Anthropomorphic Measures and Vital Status

<table>
<thead>
<tr>
<th>Measure</th>
<th>Full BED Cohort</th>
<th>Exact &quot;Binge Eating Disorder&quot; Text Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Median</td>
</tr>
<tr>
<td>Body Mass Index</td>
<td>5,739</td>
<td>35.6</td>
</tr>
<tr>
<td>Systolic Blood Pressure</td>
<td>6,760</td>
<td>122.0</td>
</tr>
<tr>
<td>Diastolic Blood Pressure</td>
<td>6,758</td>
<td>77.0</td>
</tr>
<tr>
<td>Heart Rate</td>
<td>1,163</td>
<td>78.0</td>
</tr>
</tbody>
</table>
De-identified Virtual Medical Records

REMEMBERING OUR ROOTS
<table>
<thead>
<tr>
<th>Days from Index Date/Timestamp</th>
<th>Visit/Encounter ID (NLP Sequence)</th>
<th>Observation Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>=CONCATENATE(&quot;Test Name (local, mapped):&quot;, &quot;,&quot;, DAT!AI2, &quot;,&quot;, &quot;,&quot;, DAT!AQ2, CHAR(10), &quot;Specimen Type (local):&quot;, &quot;,&quot;, DAT!AC2, CHAR(10), &quot;Result (local, normalized):&quot;, &quot;,&quot;, DAT!AA2, &quot;,&quot;, &quot;,&quot;, DAT!AL2, CHAR(10), &quot;Units (local, normalized):&quot;, &quot;,&quot;, DAT!AE2, &quot;,&quot;, &quot;,&quot;, DAT!AJ2, CHAR(10), &quot;Normal Range:&quot; &quot;,&quot;, &quot;,&quot;, DAT!AM2, &quot;,&quot;, DAT!AK2)</td>
</tr>
</tbody>
</table>

**Date Specimen Collected:**
labres_collected_date

**Date Result Available:**
labres_available_date

**Laboratory Result (local, normalized):**
localresult_inferred

**Units (local, normalized):**
localunits

**Normalized Value:**
normalizedvalue

**Normal Range:**
(normalrange)

**Qualitative Test Result:**
mapped_qual_code

**Test Name (local, mapped):**
name_m

**Specimen Type (local):**
localspecimentype

**Standard Code CUI:**
standardcode_cui
| Date Specimen Collected: 20AUG2013 | Laboratory Result | Test Name: Calcium.total Specimen Type (local): Serum Result (local, normalized): 9.40, mg/dL Normal Range: (8.9 - 10.3) Qualitative Test Result: Normal |
| Date Result Available: 20AUG2013 | |
| Date Updated: 14JAN2015 | Allergy | Type of Allergy: Drug Specific Allergen: Amoxicillin |
| Date of Onset: 06SEP2014 | |
| Date of Diagnosis: 17AUG2011 | Diagnosis | Diagnosis Code: 272.4 Diagnosis Description: Other or Unspecified Lipidemia Present on Admission?: YES Diagnosis Active?: YES |
| Date Updated: 13JAN2015 | |
| Update Date: 11JAN2015 | PAIN (Observation) | Value: 6 Unit: out of 10 |
| Observation Date: 29MAY2014 | |
| Update Date: 10JAN2015 | Pulmonary Function Test | FEV1: 3.27 FVC: 4.23 FEV1 - FVC Ratio: 77 |
| Result Date: 08MAY2014 | |
Thank you

David Dore
781.472.8458
david.dore@optum.com