Using CPCSSN Data in Primary Care Research
The Art of the Possible

K Keshavjee, K Martin, T Williamson, R Birtwhistle
InfoClin Inc, CPCSSN, Queen’s University

The Opportunity

What types of research does CPCSSN enable?

CPCSSN Facts

- 7 Provinces
- 10 PBRNS
- 380 Physicians
- 450,000 patients
- 10 EMR systems

- Developed a method to extract anonymized data from 10 different EMRs
- Standardizes data for Surveillance, QI and Research
- Extractions done quarterly

CPCSSN Data Holdings

<table>
<thead>
<tr>
<th>Data Item</th>
<th>Total</th>
<th>% Before Cleaning</th>
<th>% After Cleaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sites</td>
<td>380</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patients</td>
<td>440,994</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Encounter</td>
<td>6,494,718</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health Condition</td>
<td>2,005,940</td>
<td>61%</td>
<td>73%</td>
</tr>
<tr>
<td>Encounter Diagnosis</td>
<td>8,799,009</td>
<td>63%</td>
<td>84%</td>
</tr>
<tr>
<td>Billing Diagnosis</td>
<td>8,433,956</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Lab Results</td>
<td>3,684,253</td>
<td>29%</td>
<td>100%</td>
</tr>
<tr>
<td>Medications</td>
<td>5,926,451</td>
<td>71%</td>
<td>93%</td>
</tr>
<tr>
<td>Allergies</td>
<td>193,168</td>
<td>18%</td>
<td>18%</td>
</tr>
<tr>
<td>Physical Signs</td>
<td>5,040,003</td>
<td>96%</td>
<td>100%</td>
</tr>
<tr>
<td>Medical Procedures</td>
<td>1,567,059</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Referrals</td>
<td>705,486</td>
<td>0%</td>
<td>80%</td>
</tr>
<tr>
<td>Risk Factors</td>
<td>709,534</td>
<td>0%</td>
<td>90%</td>
</tr>
<tr>
<td>Vaccines</td>
<td>1,110,673</td>
<td>46%</td>
<td>93%</td>
</tr>
<tr>
<td>Patients w Index Disease</td>
<td>136,304</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index Disease Cases</td>
<td>204,817</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disease Indicators</td>
<td>1,765,140</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continuous Quality Improvement

- New cleaning algorithms being developed constantly
- All algorithms are validated by 2 independent reviewers, reconciled by consensus
- Case finding definitions for 8 chronic diseases have been validated for this database
- Case finding algorithms have been validated through chart reviews in physician offices

Future Studies and Applications

- Pharmaco-surveillance
- Infectious disease surveillance
- Additional chronic disease surveillance
- Apply predictive analytic algorithms to CPCSSN data for advising patients about health risks and educating them about effective interventions
- Integrate into mobile applications for patients
- Link with other databases for system-wide studies

Current Studies within CPCSSN

- Association Study 9%
- Attitudes 2%
- Audit and feedback 2%
- Case control study 7%
- Clinical Quality improvement 2%
- Continuity of Care 2%
- Data quality 20%
- De-identification 2%
- Denominator 2%
- Descriptive study 2%
- EMR Adoption 2%
- Feasibility 2%
- Intervention assessment 2%
- Medication 2%
- Practice profile 4%
- Resource use 7%
- Surveillance 18%
- SES study 4%
- Treatment pattern 4%
- Validation 4%

N=53

CPCSSN data is being re-identified at one clinical site for use as a disease registry – testing on-going