Artificial intelligence in clinical medicine
Science fiction, science fact
Science fiction, science fact
What do the experts say?

"Technology will replace 80% of what doctors do"
- Vinod Khosla


Executive summary
**Specialty** *n.* A specific sub discipline of medicine

What does a doctor do?

- Anaesthetist
- Family doctor
- General surgery
- Internist
- Neurologist
- Paediatrician
- Psychiatrist
- Radiologist
- …
**Specialty** *n.* A specific sub discipline of medicine

**What does a clinician do?**

- Mental health counsellor
- Nurse practitioner
- Occupational therapist
- Optometrist
- Paramedic
- Physical therapist
- Psychologist
- Speech-Language pathologist
- …
### Health Professionals per 100,000 Population

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NF</td>
<td>172</td>
<td>106</td>
<td>66</td>
<td>1,002</td>
<td>6</td>
<td>13</td>
<td>30</td>
<td>23</td>
<td>533</td>
<td>59</td>
</tr>
<tr>
<td>PE</td>
<td>128</td>
<td>75</td>
<td>52</td>
<td>903</td>
<td>5</td>
<td>31</td>
<td>43</td>
<td>36</td>
<td>448</td>
<td>75</td>
</tr>
<tr>
<td>NS</td>
<td>201</td>
<td>101</td>
<td>100</td>
<td>923</td>
<td>7</td>
<td>44</td>
<td>48</td>
<td>40</td>
<td>329</td>
<td>79</td>
</tr>
<tr>
<td>NB</td>
<td>152</td>
<td>90</td>
<td>63</td>
<td>974</td>
<td>6</td>
<td>33</td>
<td>35</td>
<td>33</td>
<td>294</td>
<td>81</td>
</tr>
<tr>
<td>QC</td>
<td>214</td>
<td>106</td>
<td>108</td>
<td>796</td>
<td>13</td>
<td>49</td>
<td>54</td>
<td>26</td>
<td>210</td>
<td>37</td>
</tr>
<tr>
<td>ON</td>
<td>180</td>
<td>85</td>
<td>95</td>
<td>697</td>
<td>23</td>
<td>56</td>
<td>60</td>
<td>19</td>
<td>225</td>
<td>60</td>
</tr>
<tr>
<td>MB</td>
<td>181</td>
<td>92</td>
<td>89</td>
<td>875</td>
<td>18</td>
<td>49</td>
<td>49</td>
<td>23</td>
<td>202</td>
<td>83</td>
</tr>
<tr>
<td>SK</td>
<td>154</td>
<td>91</td>
<td>62</td>
<td>835</td>
<td>16</td>
<td>27</td>
<td>34</td>
<td>22</td>
<td>151</td>
<td>91</td>
</tr>
<tr>
<td>AB</td>
<td>166</td>
<td>86</td>
<td>80</td>
<td>736</td>
<td>23</td>
<td>43</td>
<td>54</td>
<td>20</td>
<td>136</td>
<td>61</td>
</tr>
<tr>
<td>BC</td>
<td>195</td>
<td>106</td>
<td>88</td>
<td>681</td>
<td>18</td>
<td>44</td>
<td>65</td>
<td>21**</td>
<td>108*</td>
<td>59</td>
</tr>
<tr>
<td>YT</td>
<td>136</td>
<td>116</td>
<td>20</td>
<td>779</td>
<td>30</td>
<td>50</td>
<td>60</td>
<td>&quot;</td>
<td>&quot;</td>
<td>53</td>
</tr>
<tr>
<td>NT</td>
<td>112</td>
<td>69</td>
<td>43</td>
<td>1,027</td>
<td>&quot;</td>
<td>&quot;</td>
<td>71</td>
<td>&quot;</td>
<td>157</td>
<td>&quot;</td>
</tr>
<tr>
<td>NU</td>
<td>25</td>
<td>21</td>
<td>4</td>
<td>333</td>
<td>&quot;</td>
<td>&quot;</td>
<td>&quot;</td>
<td>&quot;</td>
<td>&quot;</td>
<td>&quot;</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>-----------------------</td>
<td>-----------------------------------</td>
<td>---------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NF</td>
<td>25</td>
<td>6</td>
<td>96</td>
<td>37</td>
<td>&quot;</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PE</td>
<td>&quot;</td>
<td>8</td>
<td>86</td>
<td>34</td>
<td>&quot;</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NS</td>
<td>26</td>
<td>7</td>
<td>100</td>
<td>45</td>
<td>&quot;</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NB</td>
<td>27</td>
<td>12</td>
<td>75</td>
<td>53</td>
<td>&quot;</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QC</td>
<td>1</td>
<td>35</td>
<td>77</td>
<td>43</td>
<td>&quot;</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ON</td>
<td>2</td>
<td>29</td>
<td>72</td>
<td>43</td>
<td>&quot;</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MB</td>
<td>2</td>
<td>37</td>
<td>78</td>
<td>46</td>
<td>89</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SK</td>
<td>22</td>
<td>11</td>
<td>109</td>
<td>52</td>
<td>103</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AB</td>
<td>1</td>
<td>32</td>
<td>10</td>
<td>96</td>
<td>54</td>
<td>38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BC</td>
<td>1</td>
<td>30</td>
<td>8</td>
<td>80</td>
<td>57</td>
<td>54</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YT</td>
<td>&quot;</td>
<td>20</td>
<td>10</td>
<td>86</td>
<td>&quot;</td>
<td>&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NT</td>
<td>&quot;</td>
<td>10</td>
<td>&quot;</td>
<td>43</td>
<td>&quot;</td>
<td>&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NU</td>
<td>&quot;</td>
<td>&quot;</td>
<td>6</td>
<td>&quot;</td>
<td>&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**
† Data are preliminary as of November 2001 and are subject to change. Rates per 100,000 population.
With the exception of physician and registered nurse data, personnel per 100,000 ratios for the Northwest Territories include Nunavut Territory data.
* Data for 2000 British Columbia LPNs are estimates.
**B.C dietitians include dietitians and nutritionists
" Not Available.
Note: due to updated population estimates frequencies may differ slightly from past publications.

**Source:**
Southam Medical Database, CIHI.
Health Personnel Database, CIHI.
Registered Nurses Database, CIHI.
How Physicians Spend Their Work Time

For more than a decade, the Canadian Medical Association has asked physicians about how they spend their work time. The graph below shows the results—the average number of hours worked per week by full and part-time Canadian physicians between 1982 and 2001, excluding time spent on-call. Annual averages range from a low of just under 47 hours in 1993 to a high of 54 hours in 1999. Most of this time was spent on direct patient care. Other activities included indirect patient care (e.g. other phone calls or charting), administration, research, teaching, continuing medical education (CME), or other activities. Figures prior to 1993 are based on census surveys of all physicians (including family doctors and specialists) in Canada; later figures come from sample surveys.

Source: Physician Resource Questionnaire, Canadian Medical Association
Decisions under uncertainty

- Humans, like our computer cousins, routinely have to make decisions based on imperfect input…
  - Family history
  - Oral reports from patients
  - Lab results
- … and rely on imperfect knowledge.
  - Changes to practice
  - Basic retrieval of information
Ontario consent flowchart, Adapted by Hébert P from Sunnybrook Health Sciences Centre Consent Guidelines, Toronto Notes
Decision criteria

- **Maximin**: “best worst” payoff < minimum outcome
- **Maximax**: “best best” payoff ≤ best possible
- **Laplace**: “best average” payoff
- **Minimax regret** (opportunity loss): “best worst” regret minimizes difference between realized payoff and best payoff for each future condition
To err is human

Humans (and other primates) are notoriously bad with information.

- Patients often misread or miscommunicate their own symptoms.
  - Nearly half of American adults have difficulty understanding and acting upon health information (IOM, 2004).

- Faulty memory; skill obsolescence; cognitive biases; cognitive/time limitations; recency biases; other human biases.

- Diagnoses correlate with advertising and media exposure.

  - Winters et al. (2012) showed that ~40,500 patients die in ICU, in the USA, each year due to misdiagnosis.
  - Non-fatal diagnostic errors cost up to $300,000 per malpractice claim.

http://www.nap.edu/openbook.php?record_id=10883&page=1
To err is human

Graber et al. (2005) studied 100 cases of diagnostic error involving internists.

- **System-related factors** (e.g., poor processes, team problems, miscommunication) contributed to 65% of cases;
- **Cognitive factors** contributed to 74% of cases.

  - Most common cause: ‘premature closure’.

  Eddy (1990) showed surgeons descriptions of surgical problems. *Should the patient have surgery?*

  - 50% said **Yes**, 50% said **No**.
  - 40% gave conflicting answers upon retesting.

To compute is divine

Bennett and Hauser (2013) compared patient outcomes between doctors and *sequential decision-making algorithms* using 500 randomly selected patients.

- Estimated AI cost: $189; Human cost: $497.
- Outcomes up to 50% better using AI.

Enlitic showed lung CT scans to their *deep learning system* and to four top *human radiologists*, to diagnose cancer.

**Humans:** FN 7%; FP 66%.

**AI:** FN 0%; FP 47%.

A Profound Example of Digital Health Getting it So Right (mygihealth.io)
What to expect in this course

The road ahead

- **Week 1**: Healthcare, EMRs, context
- **Week 2**: Language
- **Week 2.5**: Images/video
- **Week 3**: Machine learning and human-computer interaction
- **Weeks 4-9**: PROJECTS!
- **Week 10**: bioethics/implementation
- **Weeks 11-12**: Student presentations
Not on the menu

Surgical robotics

Robots of any kind, really
Not on the menu

Computational biology
Project structure

- Teams of 2-4 students

- **Proposals** (0.5-1 page) due **21 September**
  - Each week starts with **group updates** hereafter

- **Literature review** (3-5 pages) due **5 October**

- **Data & Methodology** (3-5 pages) due **19 October**

- **Preliminary results** (1-2 pages) due **2 November**

- **Student presentations** during **21-30 November**

- **Final reports** due **9 December**
Evaluation

- **Two quizzes** during class time, done individually: **5% each**
- **Participation** (i.e., showing up during workshop weeks): **5%**
- **Project**: **85%**
  - **Oral presentation**: **10%**
  - **Report**: **90%**
    - Data analytics: **15%**
    - Code: **15%**
    - Experiments/analysis: **30%**
    - Literature review: **15%**
    - Technical quality: **15%**
    - Overall presentation: **10%**

**Stand up**: Every Monday morning, someone from your team will tell everyone:

1) What you’ve accomplished last week or What challenges you’ve had
2) What you’ll work on next
Your project

Where to make impact?
Mechanisms of Influence

Entire Person Health Continuum

- Well
- Researching
- Preparing
- Treatment
- Recovering
- Evaluating
- Maintaining

- Behavioral Insights
- Findability
- Predictive Analytics
- Precision Marketing
- Targeted Content
- Advanced Personalization
- Collaborative Interactions
- Self-Management Tools
- Compelling Calls to Action
- Well-Researching-Preparing-Treatment-Recovering-Evaluating-Maintaining

- Well
- Researching
- Preparing
- Treatment
- Recovering
- Evaluating
- Maintaining
Spending

**Ontario**
- Hospitals: 34%
- Physicians: 23%
- Other health spending: 19%
- Other institutions: 8%
- Drugs: 10%
- Capital: 1%
- Other professionals: 1%

**United States**
- Hospitals: 31%
- Physicians: 21%
- Other health spending: 19%
- Other institutions: 10%
- Other medical products: 7%
- Gov & private admin: 4%
- Research: 3%
- Bldgs & Equip.: 3%
- Public health: 3%
- Home health: 6%
- Nursing home: 6%
- Prescription drugs: 10%
- Gov & private admin: 7%
- Other professionals: 10%

National Health Expenditure Trends, Canadian Institute for Health Information, 2010

Department of Health and Human Services, 2011
### Health and social service institutions revenue and expenditures, by province and territory (Quebec, Ontario, Manitoba, Saskatchewan)

<table>
<thead>
<tr>
<th></th>
<th>2009</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Canada</td>
<td>Que.</td>
<td>Ont.</td>
<td>Man.</td>
</tr>
<tr>
<td></td>
<td>$ thousands</td>
<td>$ thousands</td>
<td>$ thousands</td>
<td>$ thousands</td>
</tr>
<tr>
<td>Total revenue</td>
<td>80,103,978</td>
<td>19,735,747</td>
<td>26,937,123</td>
<td>3,283,411</td>
</tr>
<tr>
<td>Own source revenue</td>
<td>12,107,664</td>
<td>2,306,378</td>
<td>5,207,403</td>
<td>296,715</td>
</tr>
<tr>
<td>Sales of goods and services</td>
<td>10,124,989</td>
<td>2,128,615</td>
<td>4,085,143</td>
<td>278,113</td>
</tr>
<tr>
<td>Investment income</td>
<td>208,214</td>
<td>27,895</td>
<td>111,664</td>
<td>3,257</td>
</tr>
<tr>
<td>Other own source revenue</td>
<td>1,774,461</td>
<td>149,868</td>
<td>1,010,596</td>
<td>15,346</td>
</tr>
<tr>
<td>Transfers from other levels</td>
<td>67,996,314</td>
<td>17,429,369</td>
<td>21,729,720</td>
<td>2,986,695</td>
</tr>
<tr>
<td>of government</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Federal government</td>
<td>7,361</td>
<td>2,399</td>
<td>507</td>
<td>..</td>
</tr>
<tr>
<td>Provincial governments</td>
<td>67,686,684</td>
<td>17,426,970</td>
<td>21,628,816</td>
<td>2,959,046</td>
</tr>
<tr>
<td>Local governments</td>
<td>302,269</td>
<td>..</td>
<td>100,396</td>
<td>27,649</td>
</tr>
<tr>
<td>Other health and social</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>services</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total expenditures</td>
<td>80,826,966</td>
<td>19,703,929</td>
<td>27,744,375</td>
<td>3,363,120</td>
</tr>
<tr>
<td>Health</td>
<td>71,435,800</td>
<td>15,990,403</td>
<td>26,052,730</td>
<td>3,021,344</td>
</tr>
<tr>
<td>Hospital care</td>
<td>37,336,376</td>
<td>8,351,064</td>
<td>15,472,861</td>
<td>1,290,216</td>
</tr>
<tr>
<td>Medical care</td>
<td>16,558,004</td>
<td>3,164,202</td>
<td>5,323,829</td>
<td>725,733</td>
</tr>
<tr>
<td>Preventive care</td>
<td>1,636,130</td>
<td>688,918</td>
<td>660,658</td>
<td>..</td>
</tr>
<tr>
<td>Other health services</td>
<td>15,905,290</td>
<td>3,786,219</td>
<td>4,595,383</td>
<td>1,005,395</td>
</tr>
<tr>
<td>Social services</td>
<td>9,112,216</td>
<td>3,483,525</td>
<td>1,658,292</td>
<td>338,930</td>
</tr>
<tr>
<td>Social assistance</td>
<td>3,299</td>
<td>..</td>
<td>484</td>
<td>..</td>
</tr>
<tr>
<td>Other social services</td>
<td>9,108,917</td>
<td>3,483,525</td>
<td>1,657,808</td>
<td>338,930</td>
</tr>
<tr>
<td>Debt charges</td>
<td>278,949</td>
<td>230,002</td>
<td>33,352</td>
<td>2,846</td>
</tr>
<tr>
<td>Protection of persons and</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>property</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Housing</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>Surplus or deficit</td>
<td>-722,988</td>
<td>31,818</td>
<td>-807,252</td>
<td>-79,709</td>
</tr>
</tbody>
</table>
# The future of care delivery

## Sample Use Cases

<table>
<thead>
<tr>
<th>When Care is Delivered</th>
<th>Where Care is Delivered</th>
<th>By whom Care is Delivered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retina Selfie – Retina self-imaging allows patients to monitor for diseases such as multiple sclerosis and detect early warning signs</td>
<td>Healogram – Mobile platform that helps providers remotely monitor patients post-surgical procedure</td>
<td>iDAvatars – 2D avatar, Sophie, uses artificial intelligence and natural language processing to remotely monitor patients</td>
</tr>
</tbody>
</table>
Case study: BIDMC and HMS
The future of the patient

Sample Use Cases

Self service
Tyto Care – handheld device that patients can use to self-examine their mouth, throat, eyes, heart, lungs, skin, and temperature

New ways to engage
Deloitte Cognitive Engagement – designed to boost patient involvement in care and expand the type of alerts and interactivity offered online

The quantified self
Ginger.io – aggregates cellphone data to monitor patient mental health & alert caregivers when symptoms are problematic

Healogram
USC Institute for Creative Technologies
fitbit
Google

From Deloitte LLC 2015
Gamification

Brown et al. (1997) designed a game to help kids manage diabetes and led to a 77 percent reduction in urgent care visits.

Accenture reports 7 key elements to gamification:
1. status,
2. milestones,
3. competition,
4. rankings,
5. social connectedness,
6. immersion reality, and
7. personalization.
The future of operations

Sample Use Cases

Improved productivity

Aethon TUG Robots – Smart, autonomous robots substitute for the labor needed to haul and transport materials & clinical supplies

Workforce augmentation

Evena – Technician glasses provide high-definition, real-time images of vascular anatomy to enable fast, precise IV access

Research

IBM and Mayo Clinic – using artificial intelligence to more accurately match patients with appropriate clinical trials

From Deloitte LLC 2015
Case study: Figure 1
Project ideas

- **Intelligent EMR**: differential diagnosis
- **Model-building** from unstructured data
  - Signs of neuropsychiatric issues in social media (e.g., Reddit)
- **HealthKit**-enabled applications
  - Smartwatch application (e.g., COPD)
- **Search**: A better ‘Dr. Google’
  - CARE-RATE — specifically designed for carers of people with dementia
- …
Resources 1

- Apple ResearchKit (free) or HealthKit (not free)
- IBM Watson
- Toronto Notes (1.3K pages)
- UpToDate (not free)
- emotiv EEG
Resources 2

- EMR/mySQL
- ‘10K patients’, 36K admissions, and 11M lab observations
- i2b2 challenges around smoking, obesity, and medication (67MB)
- ...
# Resources 3

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Link</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>i2b2 Informatics for Integrating Biology &amp; the Bedside</td>
<td><a href="https://www.i2b2.org/NLP/DataSets/Main.php">https://www.i2b2.org/NLP/DataSets/Main.php</a></td>
<td>Clinical notes used for clinical NLP challenges&lt;br&gt;  • 2006 Deidentification and Smoking Challenge&lt;br&gt;  • 2008 Obesity Challenge&lt;br&gt;  • 2009 Medication Challenge&lt;br&gt;  • 2010 Relations Challenge</td>
</tr>
<tr>
<td>Computational Medicine center</td>
<td><a href="http://www.people.vcu.edu/~btmcinnes/projects/icd9cm.html">http://www.people.vcu.edu/~btmcinnes/projects/icd9cm.html</a></td>
<td>Classifying Clinical Free Text Using Natural Language Processing</td>
</tr>
<tr>
<td>Texas Hospital Inpatient Discharge</td>
<td><a href="https://www.dshs.texas.gov/thcic/hospitals/Inpatientpudf.shtm">https://www.dshs.texas.gov/thcic/hospitals/Inpatientpudf.shtm</a></td>
<td>Patient: hospital location, admission type/source, claims, admit day, age, icd9 codes + surgical codes</td>
</tr>
<tr>
<td>National (Nationwide) Inpatient Sample (NIS)</td>
<td><a href="https://www.hcup-us.ahrq.gov/nisoverview.jsp">https://www.hcup-us.ahrq.gov/nisoverview.jsp</a></td>
<td>NIS is the largest publicly available all-payer inpatient health DB&lt;br&gt; Unweighted, it contains more than 7 million hospital stays each year.&lt;br&gt; Weighted, it estimates more than 35 million hospitalizations nationally.</td>
</tr>
</tbody>
</table>
Start on your report

_pick a venue for your report and build your report towards it._

E.g., The BMJ, NIPS conference

If your goal is _an app on an app-store_, go for it.

If you intend to involve or _recruit_ other humans — let me know asamfp!
How data are currently stored and used

Electronic medical records

How humans and computers currently share medical data
Data deluge

Sources of the data deluge

- Mobile
- EMRs
- Paper/Text
- Social Media
- Images
- Videos
- Sensors & Devices

40-50% Annual growth in digital data volume*
62% Annual growth in unstructured data*
~9X of unstructured data vs. structured data by 2020**

Advances in computation

- Smarter algorithms
- Faster devices
- Improved visualization

Advances in **analytical** and **computing** techniques coupled with the explosion of data in healthcare can help uncover leading clinical practices, shrink research discovery time, streamline administration, and offer new personalized engagement at an industrial scale that align people’s decisions.

* HP Autonomy, *Transitioning to a new era of human information*, 2013
** Steve Hagan, *Big data, cloud computing, spatial databases*, 2012
Recall: (aka sensitivity, true positive rate) the proportion of positives that are correctly identified as such.

Precision: the proportion of identified elements that are true positives.

Specificity: (aka true negative rate) the proportion of negatives that are correctly identified as such.
The **International Classification of Diseases (ICD)** is a hierarchical terminology of diseases, signs, symptoms, and procedure codes maintained by the **World Health Organization (WHO)**.

In the US, most people use ICD-9, and the rest of world use ICD-10.

**Pros:** Universally available.

**Cons:** medium recall and medium precision for characterizing patients.

- (250) Diabetes mellitus
  - (250.0) Diabetes mellitus without mention of complication
  - (250.1) Diabetes with ketoacidosis
  - (250.2) Diabetes with hyperosmolarity
  - (250.3) Diabetes with other coma
  - (250.4) Diabetes with renal manifestations
  - (250.5) Diabetes with ophthalmic manifestations
  - (250.6) Diabetes with neurological manifestations
  - (250.7) Diabetes with peripheral circulatory disorders
  - (250.8) Diabetes with other specified manifestations
  - (250.9) Diabetes with unspecified complication
The **Current Procedural Terminology (CPT)** was created by the American Medical Association for **billing purposes**.

- CPT is similar to ICD, except it identifies services rendered, not the diagnosis.
  - Category I: evaluation, anesthesia, surgery, radiology,…
  - Category II: no value (e.g., patient safety).
  - Category III: **emerging technology**

**Pros**: High precision

**Cons**: Low recall

---

**Codes for Evaluation and Management**: 99201-99499
- (99201 - 99215) office/other outpatient services
- (99217 - 99220) hospital observation services
- (99221 - 99239) hospital inpatient services
- (99241 - 99255) consultations
- (99281 - 99288) emergency dept services (99291 - 99292) critical care services
...
Lab results

 riches for labs:

Many lab systems still use local dictionaries for encoding
Diverse numeric scales on different labs
Often need to map to normal, low, or high ranges in order to be useful for analytics
Missing data

not all patients have all labs

The order of a lab test can be predictive, for example, BNP indicates high likelihood of heart failure.

<table>
<thead>
<tr>
<th>Time</th>
<th>Lab</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996-03-15 12:50:00.0</td>
<td>BUN</td>
<td>16.0</td>
</tr>
<tr>
<td>1996-03-15 12:50:00.0</td>
<td>HDL-C</td>
<td>37.0</td>
</tr>
<tr>
<td>1996-03-15 12:50:00.0</td>
<td>K</td>
<td>4.5</td>
</tr>
<tr>
<td>1996-03-15 12:50:00.0</td>
<td>Cl</td>
<td>102.0</td>
</tr>
<tr>
<td>1996-03-15 12:50:00.0</td>
<td>Glue</td>
<td>86.0</td>
</tr>
<tr>
<td>1996-03-15 12:50:00.0</td>
<td>CO2</td>
<td>29.0</td>
</tr>
</tbody>
</table>
Clinical notes

- Clinical notes contain rich and diverse source of information.
- Challenges for handling clinical notes include:
  - Ungrammatical, short phrases; abbreviations; typos; Semi-structured
- Structured template: SOAP notes: Subjective, Objective, Assessment, Plan

<table>
<thead>
<tr>
<th>OBJECTIVE</th>
<th>SUBJECTIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>250.00 DM, CONTROLLED, TYPE II</td>
<td>ANXIETY STATE NOS 300.00</td>
</tr>
<tr>
<td>585.3 Kidney DZ, Chronic (GFR&gt;30-59), STG III</td>
<td>DEPRESSIVE DISORDER NEC311</td>
</tr>
<tr>
<td>412 OLD MYOCARDIAL INFARCT</td>
<td>ATRIAL FIBRILLATION 427.31</td>
</tr>
<tr>
<td>715.09 GENERAL OSTEOARTHRROSIS</td>
<td>OLD MYOCARDIAL INFARCT 412</td>
</tr>
<tr>
<td>427.31 ATRIAL FIBRILLATION</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ASSESSMENT</th>
<th>PLAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>BP 122/68</td>
<td>Continue present medication(s)</td>
</tr>
<tr>
<td>PULSE 78</td>
<td>Referral(s) to: eye</td>
</tr>
<tr>
<td>TEMP ( Src) 98.1 (oral)</td>
<td>Injection(s) ordered: b12</td>
</tr>
<tr>
<td>Resp 22</td>
<td>Schedule labs: Labs on return</td>
</tr>
<tr>
<td>Wt 227 lbs</td>
<td></td>
</tr>
<tr>
<td>Abdomen: soft, non-tender, obese and no masses or organomegaly</td>
<td></td>
</tr>
<tr>
<td>Back: No CVA tenderness</td>
<td></td>
</tr>
<tr>
<td>Extremeties: No edema</td>
<td></td>
</tr>
</tbody>
</table>

More later…
Medication

- The standard code is the **National Drug Code (NDC)** by Food and Drug Administration (FDA), which gives a unique identifier for each drug.
  - Not used universally by EMR systems.
  - Quite specific: drugs with the same ingredients but different brands have different NDC.
- **RxNorm**: a normalized naming system for generic and branded drugs by National Library of Medicine.

- Medication data can vary in EMR systems
  - Can be in both structured or unstructured forms.

- **Availability** and **completeness** of medication data vary.
  - Inpatient medication data are complete, but outpatient medication data are not.
  - Unclear whether patients actually filled those prescriptions.
# Qualities of EMR data

<table>
<thead>
<tr>
<th></th>
<th>ICD</th>
<th>CPT</th>
<th>Lab</th>
<th>Medication</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Availability</strong></td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td><strong>Recall</strong></td>
<td>Medium</td>
<td>Poor</td>
<td>Medium</td>
<td>Inpatient: High</td>
<td>Outpatient: Mixed</td>
</tr>
<tr>
<td><strong>Precision</strong></td>
<td>Medium</td>
<td>High</td>
<td>High</td>
<td>Inpatient: High</td>
<td>Outpatient: Mixed</td>
</tr>
<tr>
<td><strong>Format</strong></td>
<td>Structured</td>
<td>Unstructured</td>
<td>Mostly structured</td>
<td>Both</td>
<td>Unstructured</td>
</tr>
<tr>
<td><strong>Pros</strong></td>
<td>Easy to work with; approximates disease</td>
<td>Easy to work with; high precision</td>
<td>High data validity</td>
<td>High data validity</td>
<td>More detail about thought process</td>
</tr>
<tr>
<td><strong>Cons</strong></td>
<td>Can be applied only in screening</td>
<td>Missing data</td>
<td>Data normalization and ranges</td>
<td>Prescriptions not necessarily filled/taken</td>
<td>Difficult to process</td>
</tr>
</tbody>
</table>
Interacting with EMRs

![EMR Image]

More later…
## Certified EMRs in Ontario

<table>
<thead>
<tr>
<th>Offering</th>
<th>Vendor</th>
<th>Type</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABELMed EHR - EMR/PM v12</td>
<td>ABELMed Inc.</td>
<td>LOCAL</td>
<td>v12.9</td>
</tr>
<tr>
<td>Accuro EMR CMS4 - ASP</td>
<td>QHR Technologies Inc.</td>
<td>ASP</td>
<td>CMS4 2014.5</td>
</tr>
<tr>
<td>Accuro EMR CMS4 - Local</td>
<td>QHR Technologies Inc.</td>
<td>LOCAL</td>
<td>CMS4 2013.9</td>
</tr>
<tr>
<td>Clinic Information System (CIS) - Complete EMR v8.0 - Clinic Edition</td>
<td>P&amp;P Data Systems Inc.</td>
<td>LOCAL</td>
<td>v8.0.068 VCN Available</td>
</tr>
<tr>
<td>Clinic Information System (CIS) - Complete EMR v8.0 - Enterprise Edition</td>
<td>P&amp;P Data Systems Inc.</td>
<td>LOCAL</td>
<td>v8.0.068 VCN Available</td>
</tr>
<tr>
<td>EMR Advantage 3.2</td>
<td>Canadian Health Systems Inc</td>
<td>LOCAL</td>
<td>v3.21.18</td>
</tr>
<tr>
<td>GlobeMed v2.0</td>
<td>Alpha Global iT Inc.</td>
<td>LOCAL</td>
<td>v2.07 VCN Available</td>
</tr>
<tr>
<td>Indivicare 4</td>
<td>Indivica Inc.</td>
<td>ASP</td>
<td>v4.0.1</td>
</tr>
</tbody>
</table>
Certified EMRs in Ontario (cont.)

<table>
<thead>
<tr>
<th>Offering</th>
<th>Vendor</th>
<th>Type</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Med Access EMR v4.3</td>
<td>TELUS Health Solutions</td>
<td>LOCAL</td>
<td>v4.5 SP32</td>
</tr>
<tr>
<td>Nightingale On-Demand v9.0 - ASP</td>
<td>Nightingale Informatix Corporation</td>
<td>ASP</td>
<td>v9.2.4</td>
</tr>
<tr>
<td>Nightingale On-Demand v9.0 - Local</td>
<td>Nightingale Informatix Corporation</td>
<td>LOCAL</td>
<td>v9.2.3.1</td>
</tr>
<tr>
<td>OSCAR v15*</td>
<td>OSCAR EMR</td>
<td>LOCAL</td>
<td>v15</td>
</tr>
<tr>
<td>PS Suite v5.2 - ASP</td>
<td>TELUS Health Solutions</td>
<td>ASP</td>
<td>v5.2.532</td>
</tr>
<tr>
<td>PS Suite v5.2 - Local</td>
<td>TELUS Health Solutions</td>
<td>LOCAL</td>
<td>v5.2.550</td>
</tr>
<tr>
<td>YES EMR v2.0</td>
<td>YES Medical System</td>
<td>LOCAL</td>
<td>v2.0 BLD Sept 16.2013</td>
</tr>
<tr>
<td>YMS EMR v8.7 + 4.5</td>
<td>YMS</td>
<td>LOCAL</td>
<td>v8.73 + 4.5.3.0</td>
</tr>
</tbody>
</table>
Nightingale, e.g.
OSCAR, e.g.
Ethics in EMRs

- New US regulations passed under the Health Insurance Portability and Accountability Act (HIPAA) permit patients to request and receive their medical records electronically.

- These regulations can help patients by reducing the cost of obtaining accessible records, including otherwise ‘hidden’ information.

- 45 CFR §164.524 is effective as of 25 March 2013 and requires that providers who maintain their patient records in electronic form must, on request by the patient, provide the records in an electronic format acceptable to the patient if the record is “readily producible” in that format.

- The provider may charge a reasonable, cost-based fee, which may include the device (CD-ROM or USB device), labour costs, and postage.
Some artificial intelligence

Clinical decision support systems (CDSSs)

Typical ways intelligent systems augment EMRs
Definition of CDSS

A working definition has been proposed by Dr. Robert Hayward of the Centre for Health Evidence:

"Clinical Decision Support systems link health observations with health knowledge to influence health choices by clinicians for improved health care"
General structure of CDSSs

Information extraction
- Structured EMR
- Feature analysis
- Unstructured data

Modelling
- Patient representation
- Context
- Knowledge
- Feature selection

Prediction & inference
- Classification
- Regression
- Clustering

Information extraction

Modelling

Prediction & inference
Types of CDSS

Knowledge-based:
- Most CDSS consist of three parts:
  - the **knowledge base** contains the rules and associations of compiled data, which traditionally took the form of IF-THEN rules (e.g., IF drug(X) AND drug(Y) THEN alert(patient)),
  - an **inference engine** combines the rules from the knowledge base with the patient’s data, and
  - a **mechanism to communicate** to show results and accept new input and rules.

Non-knowledge-based:
- As you might expect, the knowledge base and inference engines are typically replaced by either **neural networks** or **genetic algorithms**.
Effectiveness 1

- A 2005 systematic review by Garg et al. of 100 studies concluded that CDSSs improved practitioner performance in 64% of the studies. The CDSSs improved patient outcomes in 13% of the studies.

- The CDSS is integrated into the clinical workflow rather than as a separate log-in or screen.

- I.e., the CDSS provides decision support at the time and location of care, rather than prior to or after the patient encounter.

- The CDSs provides (active voice) recommendations for care, not just assessments.
**Effectiveness 2**

- **Maintenance**: A core challenge is the difficulty in incorporating an extensive quantity of clinical research on an ongoing basis.
- Each year, >>10K clinical trials are published.
- Currently, each of these studies must be manually read, evaluated for legitimacy, and incorporated accurately.
Effectiveness 3

- **Evaluation:** CDSSs must demonstrably improve clinical workflow or outcome.
- The evaluation depends on the CDSS’s *purpose*.
  - E.g., a *diagnostic* CDSS may be rated by the **accuracy** of its classification of disease (as compared to physicians or other CDSSs).
  - E.g., an *evidence-based medicine* system may be rated by a high incidence of patient improvement, or financial considerations of care.
Examples

- **Zynx Health** – the most prominent organization in the CDSS marketplace, whose CDSS is linked to a statistically significant percentage of hospital discharges nationwide.

- **MYCIN**, one of the first expert systems to be developed in the 1970s, it does ethiological diagnoses of bacterial diseases.

- **CADUCEUS**, a medical expert system that could diagnose 1000 diseases.

- **Internist-I**, a computer-assisted diagnostic tool.
General structure of CDSSs

Information extraction
- Structured EMR
- Unstructured data
- Feature analysis

Modelling
- Patient representation
- Context
- Feature selection
- Knowledge

Prediction & inference
- Classification
- Regression
- Clustering

Knowledge extraction
- Modelling
- Prediction & inference
Risk analysis

If we want to regress on data \( x \in \mathbb{R}^{n \times p} \) (with \( n \) observations and \( p \) covariates) and add explicit ‘knowledge-driven’ features \( x \in \mathbb{R}^{n \times q} \) to a corresponding response \( y \in \mathbb{R}^n \), we want to measure the fitness \( F \) of \( \alpha \):

\[
F(\alpha) = \frac{1}{2} \|y - x\alpha\|^2 + \frac{\beta}{4} \left[ \sum_{i=1}^{p} \sum_{j=1}^{p} (\alpha_i \alpha_j x_i^T x_j)^2 + \sum_{i=1}^{p} \sum_{j=p+1}^{p+q} (\alpha_i \alpha_j x_i^T x_j)^2 \right] + \lambda \|\alpha\|_1
\]

- Model error
- Correlation among data-driven features
- Correlation among data- and KB-driven features
- Sparse penalty

AUC significantly improves as complementary data driven risk factors are added into existing knowledge based risk factors.

A significant AUC increase occurs when we add first 50 data driven features.
Challenges to adoption

Things to consider for your projects
Initializing the procedure

- In the short term, **intelligent decision support** systems will **continue** to be used, but perhaps only to **limited degrees**.

- Modern AI deals with **unstructured** data (e.g., images, text in medical records, patient speech) where 'noise' almost **guarantees** errors.

- There are also institutional barriers to overcome...
In the US, the FDA recognizes medical devices and permits their sale.

- About 99% of new devices are cleared if they are “substantially equivalent” to existing devices.
- Otherwise, despite guidance released in 2012, new devices must go through very rigorous “premarket approval”, sometimes requiring clinical trials. Devices then fall into three classes:

  - **Class I** devices are low risk; they do not support or sustain life.
    - E.g., dental floss
  
  - **Class II** devices do not cause harm if used as intended.
    - E.g., acupuncture needles, power wheelchairs.
  
  - **Class III** devices are high risk and subject to the highest scrutiny
    - E.g., replacement heart valves.
Regulation time

IBM has been lobbying in Washington for years to convince regulators that Watson does not need approval "because doctors make a final diagnosis"

H.R.6 - "21st Century Cures Act" passed House 344-77 (10 July 2015), read twice in Senate, passed to Committee…
Privacy

- One-fifth of smartphone owners had health apps in 2012.
- 7% of primary care physicians recommended a health app.
- The US Food and Drug Administration has approved the prescription of some apps.
- Health apps can transmit sensitive medical data, including disease status and medication compliance. Privacy risks and the relationship between privacy disclosures and practices of health apps are understudied.
Blenner et al. (2016) identified 271 diabetes apps and chose a random sample of 75 for analysis.

Within 6 months, 60 apps became unavailable, leaving 211 apps in the sample and 65 apps in the subset.

Most of the 211 apps (81%) did not have privacy policies. Of the 41 apps (19%) with privacy policies, not all of the provisions actually protected privacy (eg, 80.5% collected user data and 48.8% shared data).

Only 4 policies said they would ask permission to share data. These authorized collection and modification of sensitive information, including tracking location (17.5%), activating the camera (11.4%), activating the microphone (3.8%), and modifying or deleting information (64.0%).

In the subset, sensitive health information from diabetes apps (eg, insulin and blood glucose levels) was routinely collected and shared with third parties, with 56 of 65 apps (86.2%) placing tracking cookies; 31 of the 41 apps (76%) without privacy policies, and 19 of 24 apps (79%) with privacy policies shared user information. Of the 19 apps with privacy policies that shared data with third parties, 11 apps disclosed this fact, whereas 8 apps did not.

the hippocratic oath

I swear to fulfill, to the best of my ability and judgment, this covenant:

- I will respect the hard-won scientific gains of those physicians in whose steps I walk, and gladly share such knowledge as is mine.

- I will apply, for the benefit of the sick, all measures which are required, avoiding those twin traps of over-treatment and therapeutic nihilism.

- I will remember that there is art to medicine as well as science, and that warmth, sympathy, and understanding may outweigh the surgeon's knife or the chemist's drug.

- I will respect the privacy of my patients, for their problems are not disclosed to me that the world may know...

- I will remember that I remain a member of society, with special obligations to all my fellow human beings, ...

...