Big Data in Peripheral Vascular Disease

Stefan Verheye, MD, PhD

January 25, 2017
Disclosure Statement of Financial Interests

Co-founder and CMO of Hilbert Paradox NV
1. What is “Big Data”

2. Machine Learning

3. Solutions and Catalysts

4. Peripheral Artery Disease

5. Future/Precision Medicine
• “The potential for big data analytics to improve cardiovascular quality of care and patient outcomes is tremendous.” (Rumsfeld et al. Nature 2016)
  - However: at a nascent stage and evidence of improving care and outcomes is scant
  - Effective use of data is essential in order to realize an optimal learning health-care system
  - Health-care system today:
The problem

Excessive data:

- Activity
- Financial Means
- Family Obligations
- Vocational Factors
- Diet
- Mental Wellbeing
- Personal Goals
- Insurance
- Clinical Resources
- Education
- Community Support
- On-line Resources
- Government Resources

Source: Frost & Sullivan
Definition

- Very large datasets (petabytes), too large to handle by traditional applications, and/or,
- Very complex datasets, too complex to be analyzed by (desktop) statistics and visualization packages
- Also refers to the use of predictive analytics or other advanced data analytics methods
- Requiring massive parallel software running on thousands of servers

Big Data: Expanding on 3 fronts at an increasing rate.
Expert systems work the way an ideal medical student would: they take general principles about medicine and apply them to new patients.

Machine learning, conversely, approaches problems as a doctor progressing through residency might: by learning rules from data.
Machine learning has become ubiquitous and indispensable for solving complex problems in most sciences:

- **Astronomy**: galaxies, supernovas
- **Biomedicine**: predict protein structure and function from genes
- **Medicine**: read cortical activity, transmit signals and restore motor control

**Machine learning from streaming data (BigML):** periodic re-training with a batch algorithm
Diagnosing Anatomic and Functionally-Significant CAD

<table>
<thead>
<tr>
<th>ANATOMY</th>
<th>FUNCTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify obstructive CAD</td>
<td>Identify lesion-specific ischemia that may benefit from PCI</td>
</tr>
</tbody>
</table>

### Invasive
- ![Image of invasive procedure](image)

### Noninvasive
- ![Image of noninvasive procedure](image)

---

**FFRct Validation**

[HeartFlow logo]
Vessel Lumen Segmentation From Coronary CTA Using Deep Learning: Validation With OCT Data

Validated in 4 clinical trials and over 1200 patients

Grady and Taylor et al. 2016
“FFRct will soon be enabled for <1 hour turnaround time by harnessing automated image analysis based on deep learning with less need for human editing and continual improvement” ...C. Taylor
Hilbert Paradox (HPX) is a global, cloud-based architecture with storage, retrieval, and data processing efficiencies. HPX accelerates market access, creates value, and realizes a vision for the digital health future.
Computing Power

The HPX platform is designed to ingest millions of events per second with unprecedented scalability. It can process massive amounts of data produced by connected and diagnostic devices, websites and applications in the field of genomics, biosensors, wearables and Micro-Electro-Mechanical Systems (MEMS).
Data Analytics

By enabling predictive analytics, HPX can assist in demonstrating better health outcomes and in identifying the most valuable digital health technologies for use in patient care.
Artificial Intelligence

AI will assist doctors and healthcare professionals in making faster and more accurate diagnostics. By employing AI, HPX augments the value of the analysis, leading to better patient treatments and reduced healthcare costs.
Data Engineering: Raw Data → Knowledge Engineering → Data Engineering → HPX model

HPX Connect → HPX Align → HPX Explore → HPX Explain → HPX Learn → HPX Inform

Data Equity:
- Decision making
- Correlations & patterns
- Metadata & ontologies
- Exploratory analysis
- Normalized data
- Raw data
Practice-Based Evidence: Profiling the Safety of Cilostazol by Text-Mining of Clinical Notes

Outcome analysis in the CHF subgroup comparing patients with a history of CHF and taking Cilostazol to a matched control of CHF patients not taking Cilostazol

Leeper et al. PLOS 2013
Peripheral Artery Disease

N = 1755 patients - data derived from GenePAD study
Presented for elective coronary angiography

- Employment of multiple supervised machine learning algorithms
- Used various information and build models
- Compare to standard stepwise linear regression models

Peripheral Artery Disease

Area under the curve (AUC) for stepwise logistic regression and machine-learned penalized regression model for identification of patients with PAD

0.76 vs 0.87


Area under the curve (AUC) for stepwise logistic regression and machine-learned random forest model for prediction of mortality

0.65 vs 0.76
Peripheral Artery Disease

Precision Medicine - Paradigm Shift

**Yesterday**
- Symptom-based
- Intuition Medicine

**Today**
- Pattern-based
- Evidence-based Medicine

**Tomorrow**
- Algorithm-based
- Precision Medicine

Application of rules, algorithms and reference databases enables ACTIONABLE clinical decision support & PRECISE/EFFICIENT care.

- Personal Molecular Fingerprints (PMF, multi-omics)
- Molecular Imaging
- Non-molecular content (waveforms)
- Big Data Analytics
- Reference Databases (Rules & Algorithms)
- Mobile Information Communication Technologies
- Precision Medicine (prevention, diagnosis, treatment)
Correlation of diverse, patient centric, anonymous, normalized data
The ability to transform data into knowledge will disrupt at least three areas of medicine:

1. Machine learning will dramatically improve the ability of health professionals to establish a prognosis
   → More data, better data allowing models to use thousands of rich predictor variables (time scale: in next 5 years)

2. Machine learning will displace much of the work of radiologists and anatomical pathologists
   → Massive imaging data sets, combined with recent advances in computer vision, will drive rapid improvements in performance, and machine accuracy will soon exceed that of humans (time scale: years)

3. Machine learning will improve diagnostic accuracy
   → Algorithms will soon generate differential diagnoses, suggest high-value tests, and reduce overuse of testing (time scale: next decade)
And why is that...

Traditional image analysis algorithms plateau whereas Deep Learning algorithms continue to improve with additional data.
Future perspective

Machine learning will transform clinical medicine!