Transfusion Medicine & Transformational Change Series
Part 3 Big Data

Presenter:
James R. Stubbs, M.D.
Chair, Division of Transfusion Medicine
Department of Laboratory Medicine and Pathology
at Mayo Clinic, Rochester, Minnesota
Disclosures

- None

Big Data

- 2009 – New flu virus discovered: H1N1
  - Spread quickly
  - Health agencies fear potential pandemic
  - No vaccine readily available
  - Only hope, slow the spread of the virus
    - Needed to know where it already was
Big Data

- H1N1 Outbreak
- US CDC requested that doctors report new flu cases
  - Picture of the pandemic: 1-2 weeks out of date
  - People might feel sick for days before seeing a doctor
  - Relaying information to central organizations took time
  - CDC tabulated numbers once a week
  - Rapidly spreading disease, 2-week lag time is a very long time

- Shortly before H1N1 became widely known
- Google paper in *Nature*
  - Predict the spread of winter flu in the United States, specific regions and states
  - Looking at what people were searching on Internet
    - 3 billion search queries every day
    - 50 million most common search terms typed by Americans
    - Compared list to CDC data on the spread of seasonal flu between 2003 and 2008
      - Identify areas infected by flu by what people were searching for on the Internet
Big Data

- Google and flu
  - Processed 450 million different mathematical models to test search terms
    - Compared predictions against actual flu cases from CDC data for 2007 and 2008
    - Found a combination of 45 search terms when used in their mathematical model – Strong correlation between their prediction and official nationwide figures
      - They could tell where the flu had spread, near real time, not a week or 2 after the fact
    - H1N1 crisis 2009: Google’s system more useful and timely than government statistics!

Big Data

- Data is no longer static
  - In the past, its usefulness was finished once the purpose for collecting it is finished
  - Now data can be reused for innovation, analyses, and productive change
  - Big data: things can be done on a large scale that cannot be done on a smaller scale (ie, size matters!)
    - Extract new insights, lead to changes
    - Change our way of thinking
      - Causality → Correlations
      - Why → What
  - Major Transformation!
Big Data

- Big data core: predictions
  - Machine learning
  - Applying math to huge quantities of data to infer probabilities
    - Email: the likelihood that the typed word “teh” is actually “the”
  - Systems perform well
    - Fed lots of data to feed predictions
  - Systems
    - Built to improve themselves over time
    - Keep tabs on best signals and patterns as more data is fed in to them

Big Data

- Future
  - Many aspects that today are the sole purview of human judgment are being replaced or augmented by computer systems
    - Amazon – Recommend the ideal book
    - Google – Rank the most relevant website
    - Facebook – Knows our likes
  - Same technologies
    - Diagnosing illnesses and recommending treatments
  - Big data is changing fundamental aspects of life by giving it a quantitative dimension it never had previously
Big Data

• 3 Shifts in the way Information is analyzed
  • Shift 1
    • Analyze much more data
    • Some scenarios will use all the data as it relates to a particular process or phenomenon
    • Since 19th century science has relied on sampling when faced with large numbers
      • Artifact of a period of information scarcity
      • Natural constraints of information in an analog era
      • Took sampling for granted prior to development of high-performance digital technologies

Big Data

• Shift 1
  • Analyze much more data
    • Using all the data
      • See details that could not be seen when limited to smaller quantities
      • Clear view of the granular
        • Subcategories that sampling cannot assess
Big Data

• Shift 2
  • Loosens up desire for precision
  • Less error from sampling allows more acceptance of measurement error
    • Sampling: count only the most important things
    • Function: small data environment
  • As the scale increases, the number of inaccuracies increases
  • Often big data is “messy”
    • Sense of general direction rather than precise knowledge

Big Data

• Shift 3
  • Move away from causality
  • Big data will not have to be focused on causality
  • Discover patterns and correlations offering novel and invaluable insights
    • Will not tell us “why” something is happening
    • Will tell us “that” something is happening
  • Many times that’s good enough
**Big Data**

- **Before big data**
  - Analyses
    - Limited to a small number of hypotheses
    - Defined before data collected
- **Let the data speak!**
  - Connections will be identified that we never thought existed
- **“There is a treasure hunt currently under way!”**
  - Driven by insights to be extracted from data
  - Value that can be obtained thru a shift from causation to correlation

---

**OR Datamat System Overview**

- **Real time**
  - OR check patients
  - OR check time
- **Historical**
  - All OR check patients
  - All OR check time
- **Integrated**
  - OR check patients from historical data
  - OR check time from historical data
- **MDM**
  - Multi-dimensional surgical data
  - OR check from real-time data
  - OR check from historical data

---

**Data Sources**
- Chart
  - Demographics, patient, fluids, allergies, procedures
- CPOE
  - Order entry
- Safe Trace
  - Transfusion order tracking
- SIRS
  - Sepsis, patient data
- ARIES
  - Lab data
- Soft
  - Other systems
- UDP
  - Blood products
- VAR
  - Varicose data

---

**Info Consumers**
- OR ALERT System
  - OR alert notifications
  - OR alert statistics
- Transfusion metrics Dashboard
  - OR check patients
  - OR check time
- POINT tool
  - OR check patients
  - OR check time
- AWARE
  - OR check patients
  - OR check time
Publications of Interest


Reference


Thank You for informing and inspiring my inquiries into Big Data!
Questions or requests…
Email to: MMLHotTopics@mayo.edu

For more information…
Visit MayoMedicalLaboratories.com or call Mayo Laboratory Inquiry at 800-533-1710