Big Data in Global Health:
The Global Burden of Disease Study

Peter Speyer
@peterspeyer
March 29, 2014
Institute for Health Metrics and Evaluation

• Independent research center at the University of Washington
• Core funding by Bill & Melinda Gates Foundation and State of Washington
• 185 faculty, researchers and staff
• Providing independent, rigorous, and scientific measurement
• “Our goal is to improve the health of the world’s populations by providing the best information on population health”
Big data in global health

Variety
• Surveys
• Censuses
• Disease registries
• Vital registration
• Verbal autopsy
• Mortuaries / burial sites
• Police records

Volume
• Hospital / ambulatory / primary care records
• Claims data
• Surveillance systems
• Administrative data

Velocity
• Literature reviews
• Sensor data
• Social media
• Quantified self
IHME input data catalog at GHDx.org
Example: Global Burden of Disease Study

• A **systematic scientific** effort
to quantify the **comparative** magnitude of
**health loss** due to diseases, injuries and risk factors

• GBD 2010 results published in
*The Lancet* in 2012
  – 291 causes, 67 risk factors
  – 187 countries
  – 1990-2010
  – By age and sex

• GBD 2013 Update in process
Accessing the data

• Systematic identification of all relevant data sources
  – Lit reviews
  – Data Indexers

• Key challenges
  – Frequently changing websites
  – Data on paper/PDF or in obsolete formats
  – Unwillingness to share data
  – Confidentiality / de-identification
  – Data cleaning, e.g. garbage codes
  – Cross-walks, e.g. ICD versions
Analyzing the data

• Synthesizing all available data
  – Including covariates: indicators related to quantity of interest

• Using relevant modeling approaches
  – Predictive validity testing (CODEm)
  – Bayesian meta-regression (DisMod)
  – Small Area Estimation

• Applying appropriate corrections, e.g. causes of death to match all-cause mortality

• Quantifying uncertainty

• Review: 1000+ experts, peer-reviewed publication
Demo: Tobacco Viz

United States male smoking prevalence (%)
Sharing the data

- Researchers
- Analysts
- Domain experts
- Policy & decision makers
- Interested public

Audiences

- Academic papers
- Policy reports
- Data search engine
- Data visualizations
  - Input data
  - Exploratory
  - Declarative

Content & tools
Big Data in Health:
The Global Burden of Disease Study

Contact me
speyer@uw.edu
@peterspeyer