

# Big Data needs Little Data

Hospital Costing and Health Coding Forum 2015 Melbourne

**Heather Grain**

Global eHealth Collaborative and eHealth Education



# What is the problem?

## Morbidity Data Constrained

- IT capacity
- Record systems
- Data understanding

## Today....

- EHRs with shared structure
- Quality of structure
- Restricts utility
- Impacting classification





<http://www.informationweek.com/big-data/big-data-analytics/5-big-wishes-for-big-data-deployments/d/d-id/1109606?>

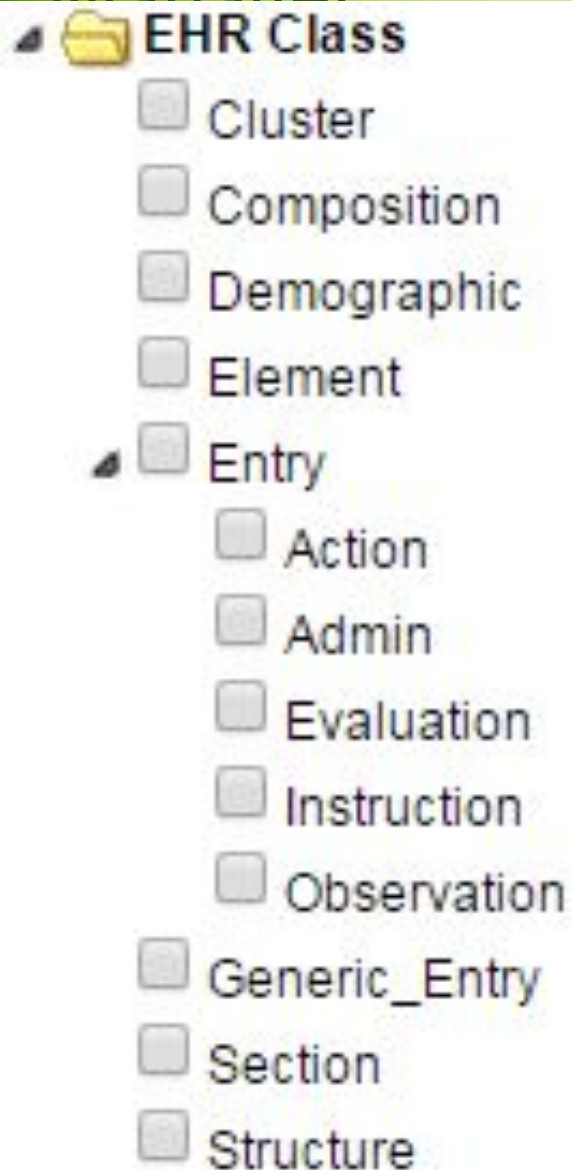


# EHR with structure – Information Model

Internationally accepted and increasingly the basis of systems and information exchange

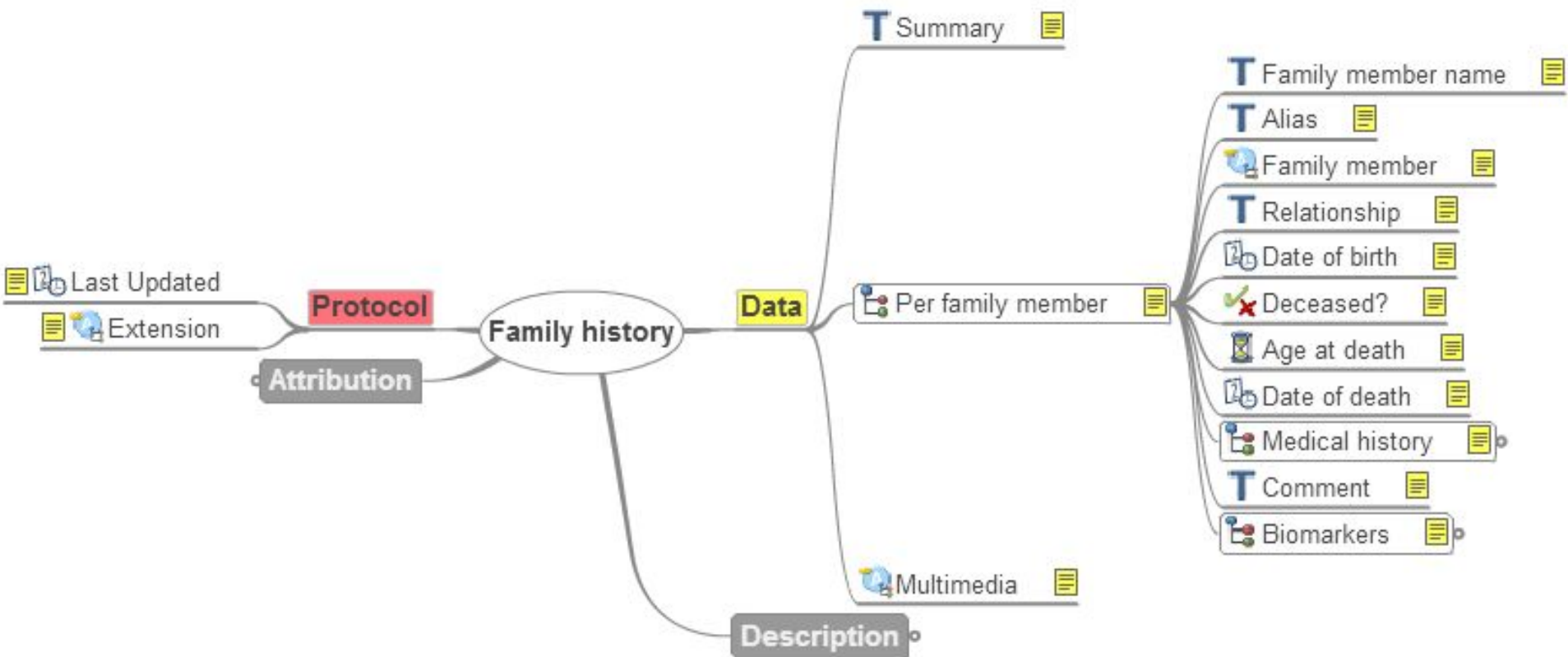
- More accurate over time in the record
- Easier to exchange (report) and design (cheaper)
- Supports queries of the data
- Clinically agreed
- Suited to ICD-11's approach

# Example: openEHR Clinical Knowledge Manager



- Record components – not forms
  - Clinically validated
  - Maximal data model
    - pick what you need

# Example of specific data



# How this impacts systems

- Select data like lego
  - Makes screens
  - Makes reports
  - Available for data extraction
- Easier, faster and cheaper IT systems
- Standards structure
  - vendors can focus on functionality

# Coding and Quality Data

- Known desire
- However – current structure breaks quality rules
  - 1 Semantic space per field
  - Subset of classification or terminology for purpose
  - Associate field with definition of single purpose



# Quality of Structure - Semantic Scope

- Data collected in a field restricted to a single consistent meaning
  - Example:
    - Marital status: civil status of each individual in relation to the marriage laws or customs of the country.
      - never married,
      - married,
      - widowed and not remarried,
      - divorced and not remarried,
      - married but legally separated,
      - defacto union
  - Does not include living arrangements (though many data collections do include this in the classification)

Glossary of statistical terms:

<https://stats.oecd.org/glossary/detail.asp?ID=1597>

9

# Semantic Scope of Diagnosis Field

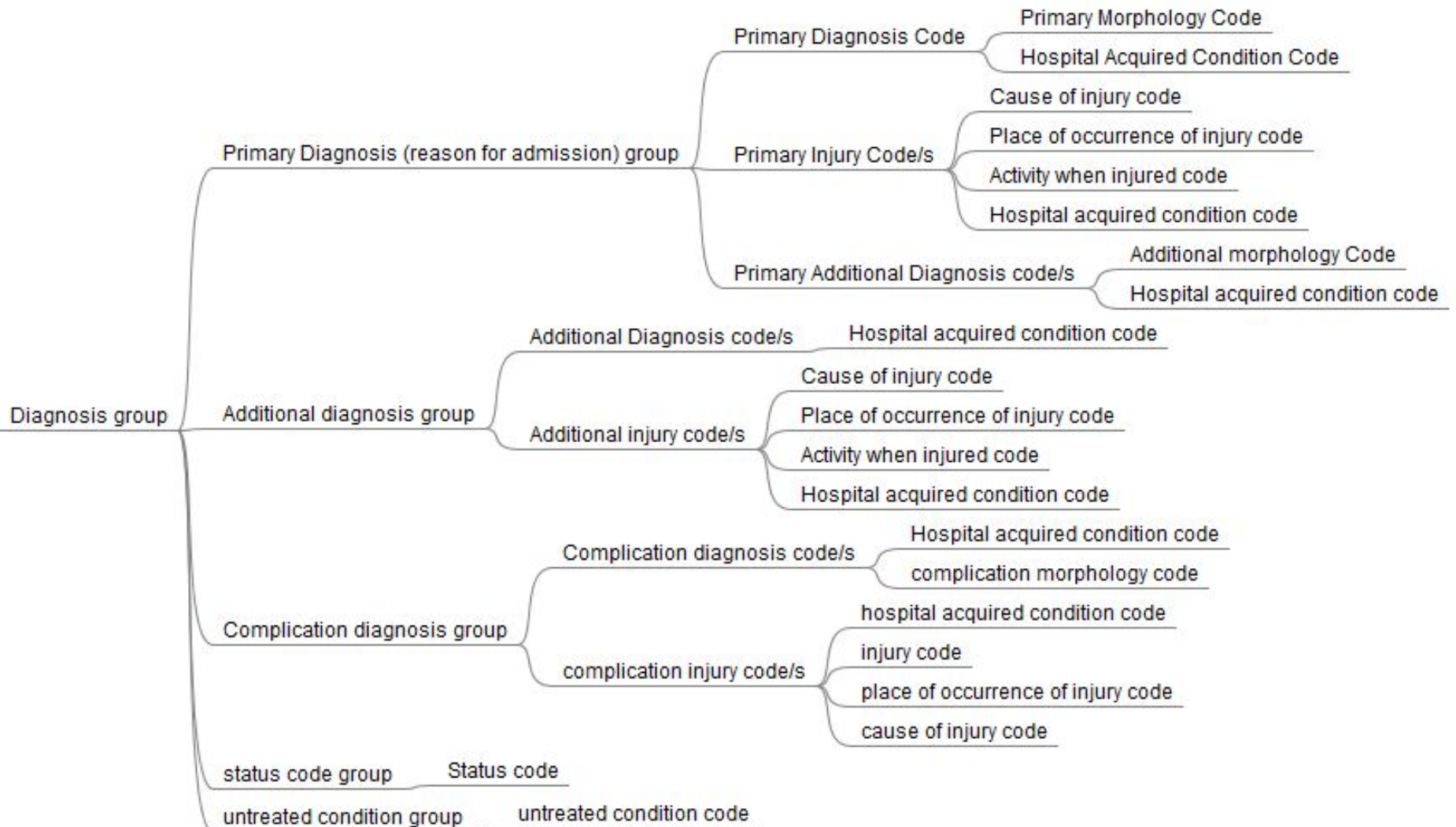
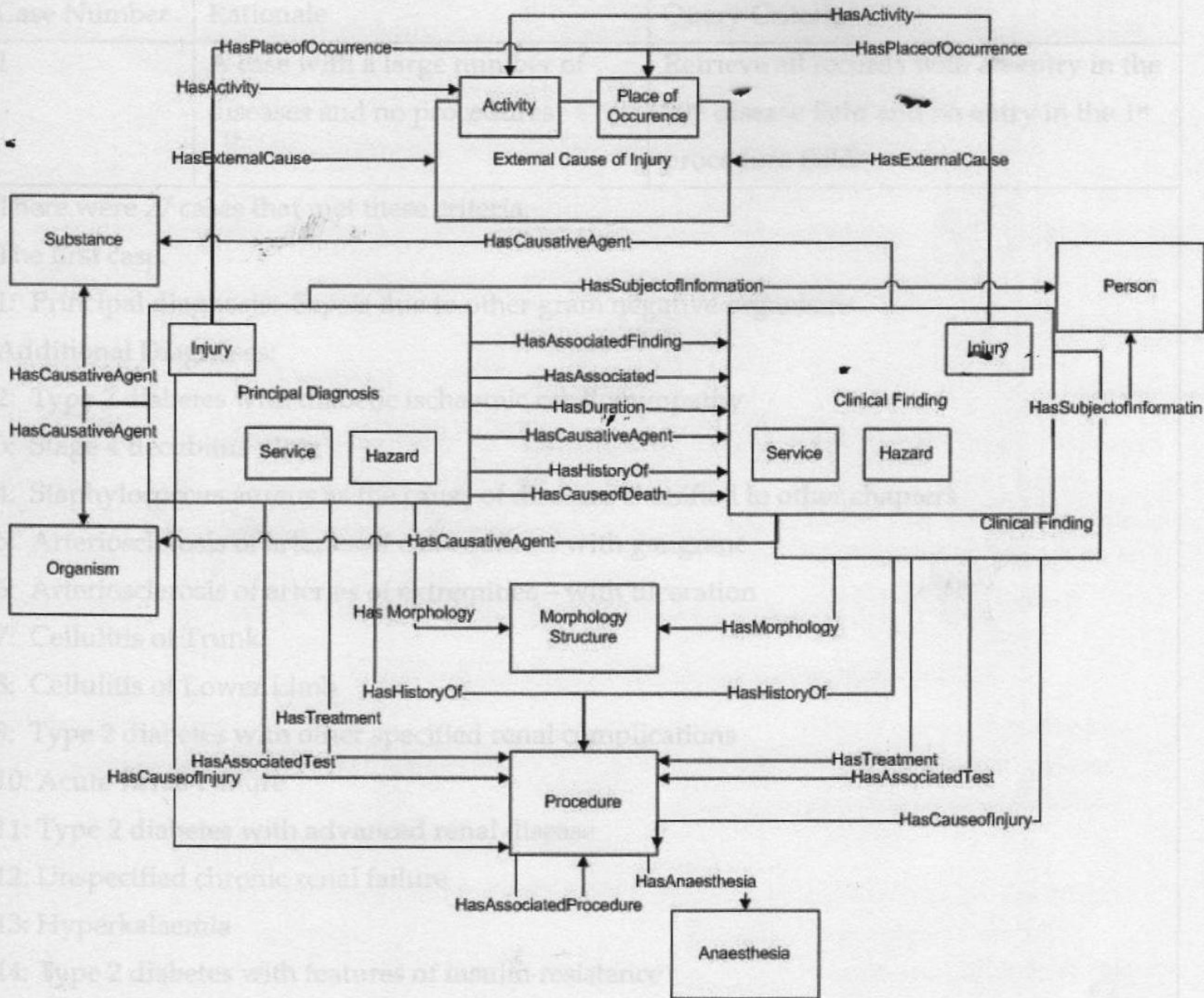


Figure 5-1 Model of morbidity data



Extract:  
Investigation of the ontology and information model of morbidity reporting in the EHR environment – Grain. 2008



# Structure

- aids retrieval
- supports knowledge acquisition



Imagine trying to write software to retrieve information by writing an instruction for a computer.....

M51.2, T81.2, Y83.6, Y92.22, 40300-00, 40330-00, 90011-02

The utility of the morbidity information could be much more.....



# Records and Systems changing....

What about morbidity classification and recording

- Don't change the classification change the field
  - If you want conditions present but not treated
    - This is a different context, not a different code! (ICD-10-AM V9)
- ICD-11 is coming – a logical time for change
- Build rules to improve quality of data capture
  - Standards need to include additional documentation
    - If X then do Y to support computation as well as manual application
- We need to think about how computers could help us do it better

# Development of test system

- GeHCO (NFP) are developing Simulator to:
  - Serve as education resource for clinical coding (over 1500 records with answers and feedback)
  - Test alternative structures for collecting morbidity data
  - Seeking interest and input

# Contact Details

If you have questions feel free to contact

- [heather.grain@gehco.org](mailto:heather.grain@gehco.org)
- Visit the web site at  
[www.gehco.org](http://www.gehco.org)