Data surveillance on the clinical data used for health system funding in Ontario

Lori Kirby, Maureen Kelly

Data Quality Department

Canadian Institute for Health Information

Canadian Institute for Health Information

Institut canadien d'information sur la santé

Outline

- About Data Quality at CIHI
- What is data surveillance?
- CIHI’s data surveillance pilot
- Big Data insights

Data Quality Is Fundamental to CIHI

Vision

Better data.
Better decisions.
Healthier Canadians.

Mandate

Deliver comparable and actionable information to accelerate improvements in health care, health system performance and population health across the continuum of care

Strategic goals

Be a trusted source of standards and quality data

Expand analytical tools to support measurement of health systems

Produce actionable analysis and accelerate its adoption

Priority themes and populations

Themes

Patient experience
Quality and safety
Outcomes
Value for money

Health system performance

Populations

Seniors and aging
Mental health and addictions
First Nations, Inuit and Métis
Children and youth

Foundation

Our people

Stakeholder engagement and partnerships

Privacy and security

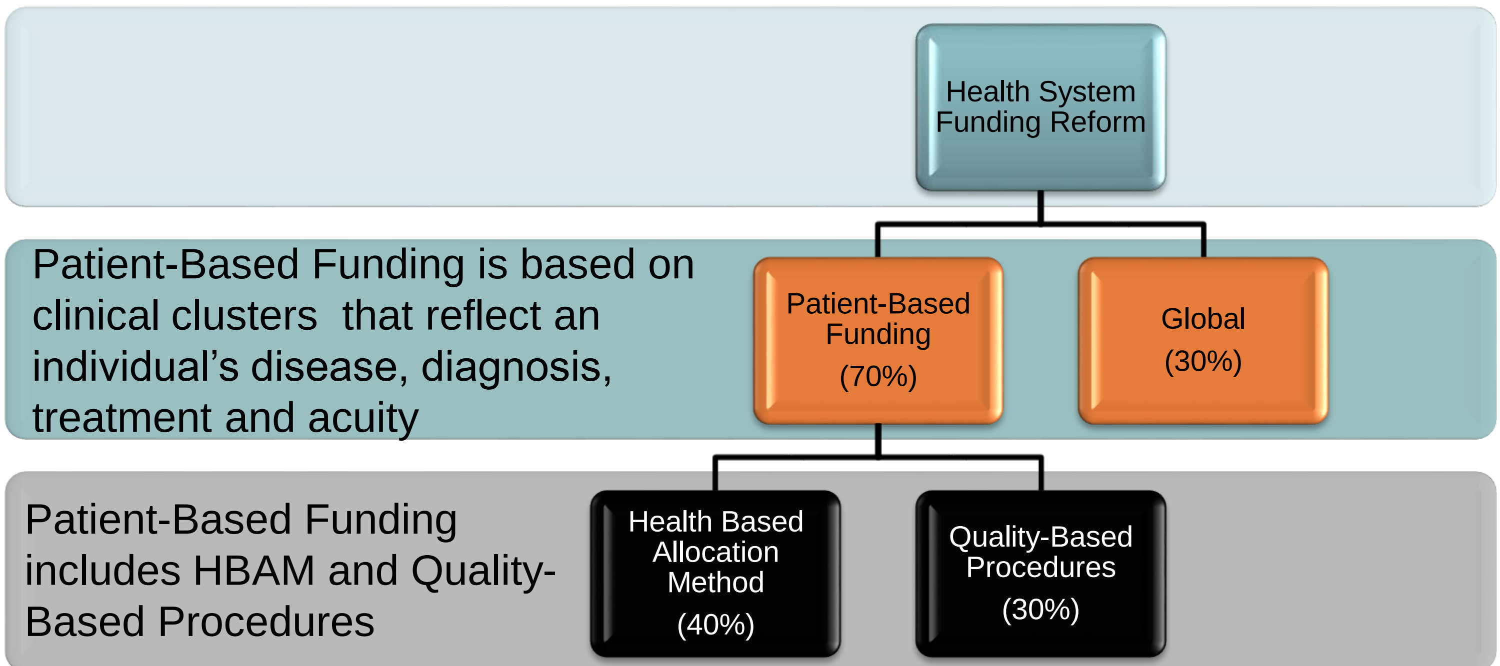
Information technology

Values

Respect • Integrity • Collaboration • Excellence • Innovation

Health System Funding Reform in Ontario

- Calls for increased focus on data quality
- Clinical administrative data being used to determine funding allocations to regions and hospitals



HBAM provides organizational-level allocations informed by case-mix utilization and aggregate cost, volume and types of patients and providers

Quality Based Procedures (QBPs) are clusters of patients with clinically related diagnoses or treatments that have been identified by an evidence-based framework as providing opportunity for process improvements, clinical re-design, improved patient outcomes, enhanced patient experience and potential cost savings

Increased focus on data quality

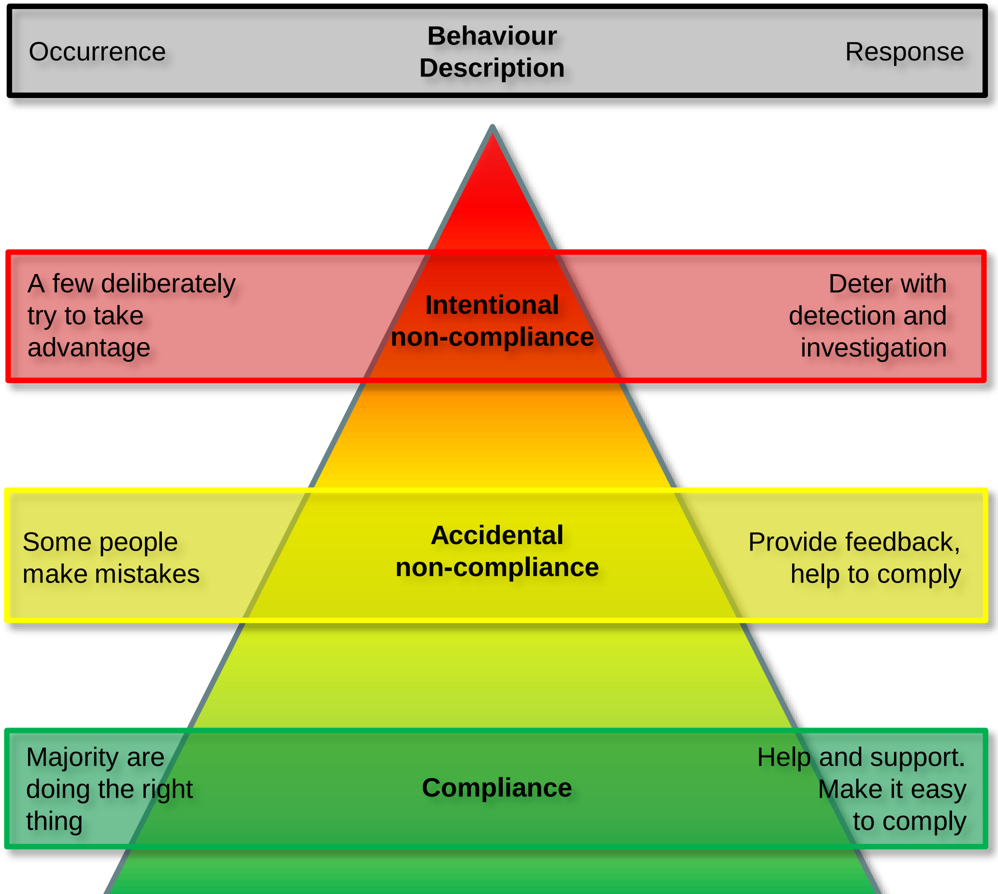
- Impacts on data quality can be both **positive** and **negative:**

  - Positive: People pay more attention to the data and its quality; more complete and timely submissions

  - Negative: Manipulation of data/coding/clinical practice to maximize funding (i.e. gaming)
- To prevent and minimize the impact on data quality, CIHI is exploring options for developing systems and processes - **Data Surveillance -** specifically **targeted** toward these issues

What do we mean by “data surveillance”?

Surveillance is targeted to those trying to taking advantage of the system



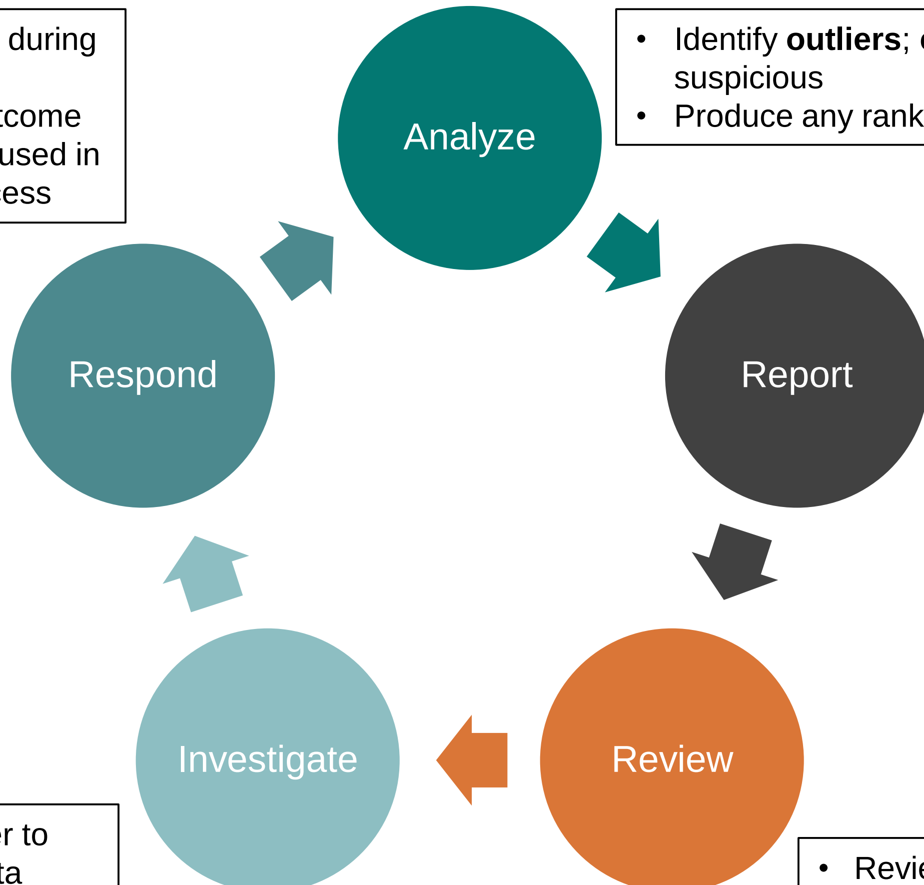
Focus of a surveillance system

Focus of traditional DQ activities

Adapted from Australia's Medicare Compliance Model

The Surveillance Process: Data into Action

- Carry out actions identified during investigation
- Document process and outcome so that information can be used in future iterations of the process



- Identify **outliers**; **changes** in data that look suspicious
- Produce any ranking/prioritization scores

- Summarize analysis and produce reports for review

- Review current results and any other information
- Determine which cases warrant further investigation

- Follow-up with data provider to determine how and why data anomalies exist
- Determine next steps

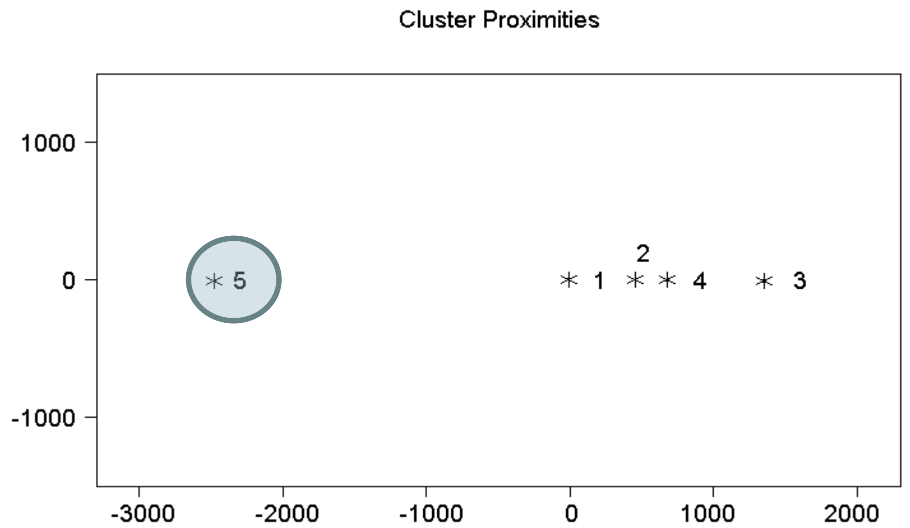
CIHI's Data Surveillance Pilot

Pilot Overview

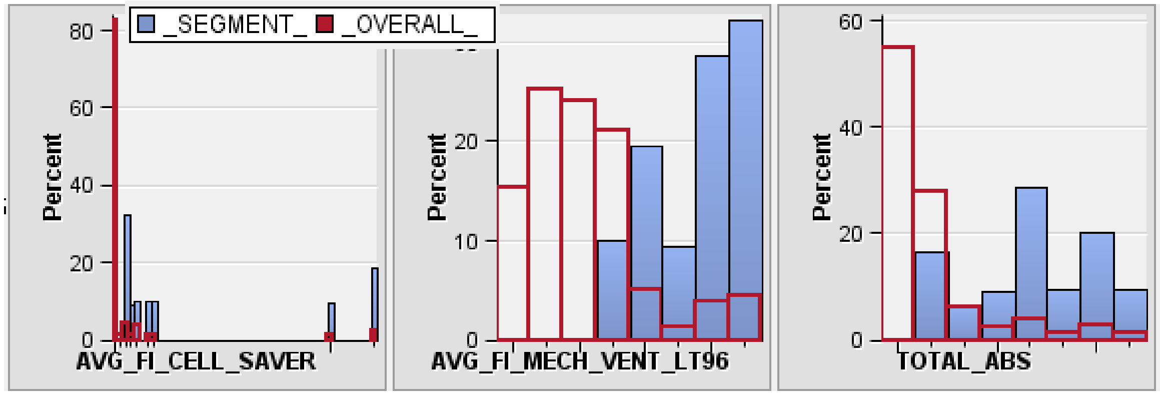
- **Objective:** Identify outliers in Ontario acute care data from CIHI's Discharge Abstract Database
- **Outcome:** Produce an **overall data quality score** to prioritize which facilities may warrant further analysis and investigation
- **Focus:** Multiple elements that impact Ontario's funding formula:
  - Special Care Units (SCU)
  - Discharge to Home Care
  - Quality Based Procedures (QBP's)
  - Comorbidities
- **Methods:**
  - Applied 3 different analytical techniques using SAS Enterprise Miner to identify outliers

Methodology 1 – Segmentation Model using Cluster Analysis

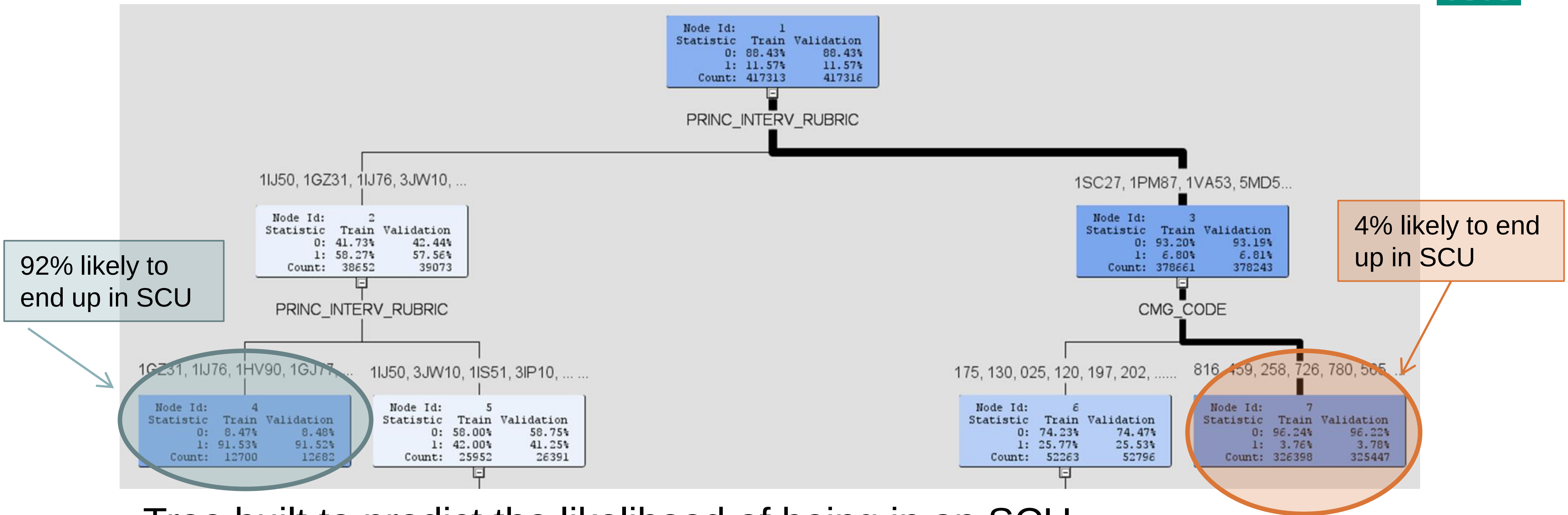
- Identify “segments” – **facilities** – with similar patterns of SCU data
- Facilities are grouped based on distribution of all variables considered
- Identified outlier group – segment 5



Individual variables in segment 5 were different compared with the rest of the facilities



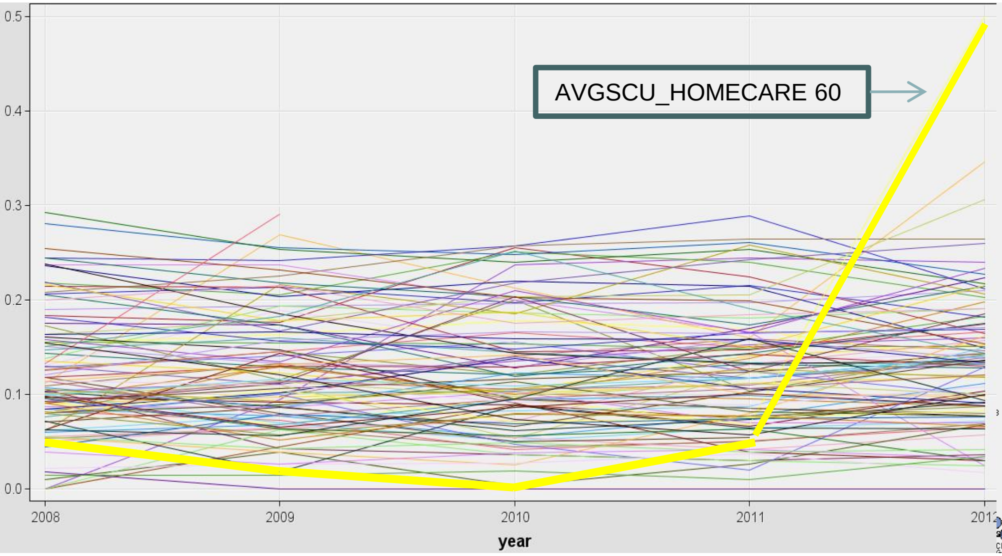
Methodology 2 – Predictive Model using Decision Trees

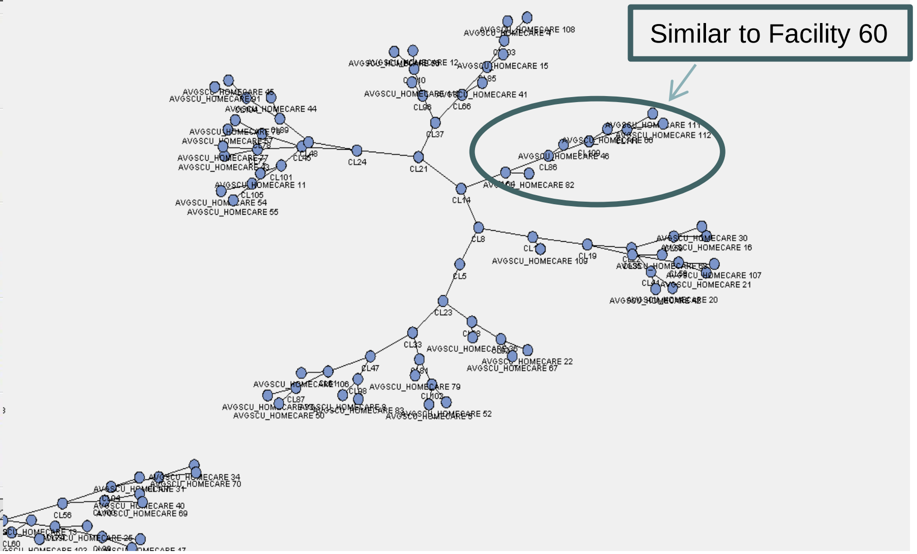


- Tree built to predict the likelihood of being in an SCU
- Model based on 2009-2010 data and then applied to 2013-2014 data
- Most Important Variables:
  - Principle Intervention Code (ruberik level)
  - Case Mix Group Code
  - Number of Intervention Episodes before SCU
- Calculate ratio of how many SCU occurred (observed) vs how many predicted by the model (expected)

Methodology 3 – Time Series Model

- Using Time Series to compare facility-level volumes for discharge to home care
- Identify the facility with the most significant change over time, set it as a target
- Identify the facilities that are most similar to the target facility
- Process will be repeated for other variables:
    - Volumes of Quality Based Procedures
    - Number of comorbidities coded





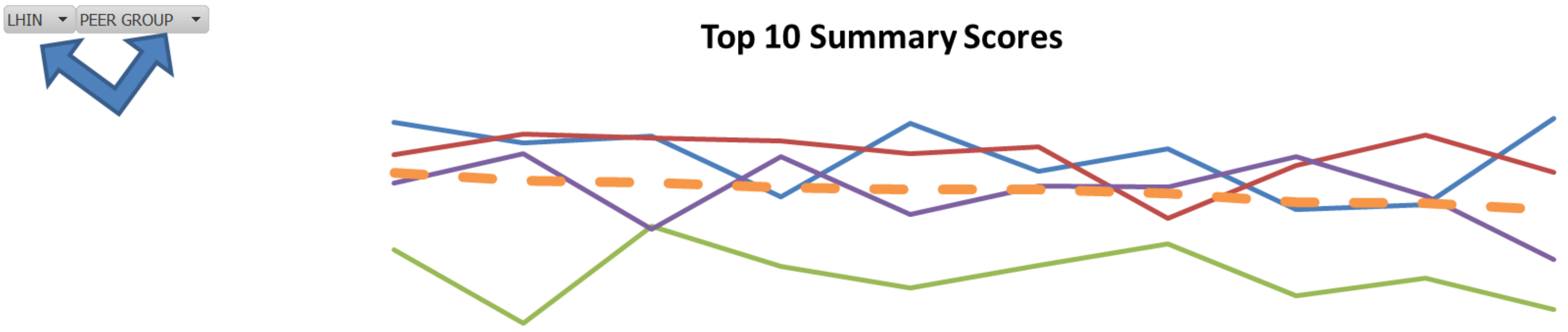

RESULTS

Overall Score Card

Facility Name	LHIN	PEER GROUP	Total Score	SCU Score	QBP Score	Discharge to HC Score	Comorbidity Score
Hosptial 237	LHIN H	Teaching	8.29	**9.87**	8.85	5.86	**7.96**
Hosptial 217	LHIN M	Teaching	8.05	**9.24**	**9.49**	3.53	**8.89**
Hosptial 128	LHIN M	Teaching	7.97	**9.43**	**9.38**	6.58	6.49
Hosptial 208	LHIN L	Teaching	7.82	7.52	**9.29**	5.33	**8.79**
Hosptial 104	LHIN K	Teaching	7.75	**9.85**	8.88	4.64	6.97
Hosptial 164	LHIN F	Teaching	7.75	8.33	9.09	5.35	**7.86**
Hosptial 163	LHIN F	Teaching	7.64	**9.04**	6.85	6.03	**7.84**
Hosptial 209	LHIN J	Teaching	7.36	7.11	8.53	4.40	**8.79**
Hosptial 166	LHIN F	Teaching	7.33	7.28	**9.47**	4.94	7.55
Hosptial 213	LHIN M	Teaching	7.12	**10.00**	8.30	3.97	5.54
Hosptial 126	LHIN M	Teaching	7.05	6.60	8.57	2.62	**9.43**
Hosptial 205	LHIN K	Large Community	7.02	**8.51**	6.70	6.59	6.03
Hosptial 236	LHIN D	Teaching	6.96	6.06	9.08	4.70	**7.94**
Hosptial 119	LHIN K	Large Community	6.93	**9.44**	7.77	7.61	3.42
Hosptial 240	LHIN D	Teaching	6.57	**9.15**	2.89	4.57	7.78
Hosptial 138	LHIN L	Small	6.31	6.74	4.16	6.78	7.00
Hosptial 150	LHIN D	Small	6.29	6.70	3.32	7.38	7.13
Hosptial 159	LHIN N	Large Community	6.25	7.36	8.04	6.84	3.55
Hosptial 258	LHIN C	Large Community	6.25	5.46	**9.46**	6.18	4.93
Hosptial 101	LHIN L	Large Community	6.20	7.64	7.72	7.81	2.69

Score:		
		close to zero
		range of caution
		top outliers

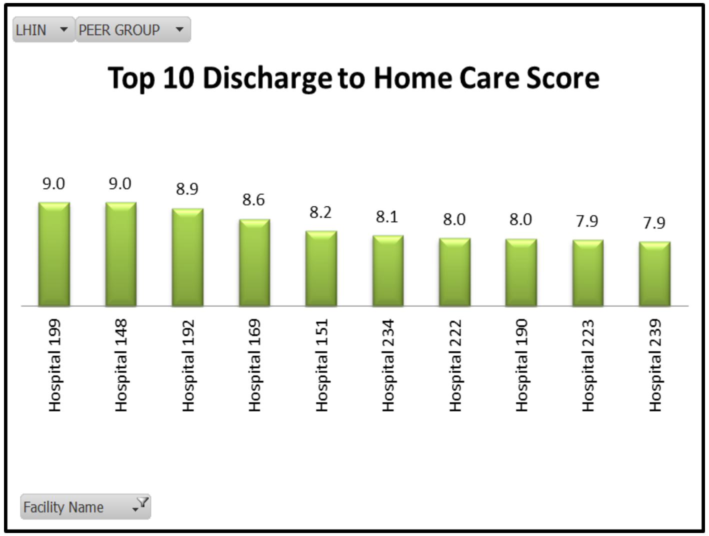
Dashboard Summary

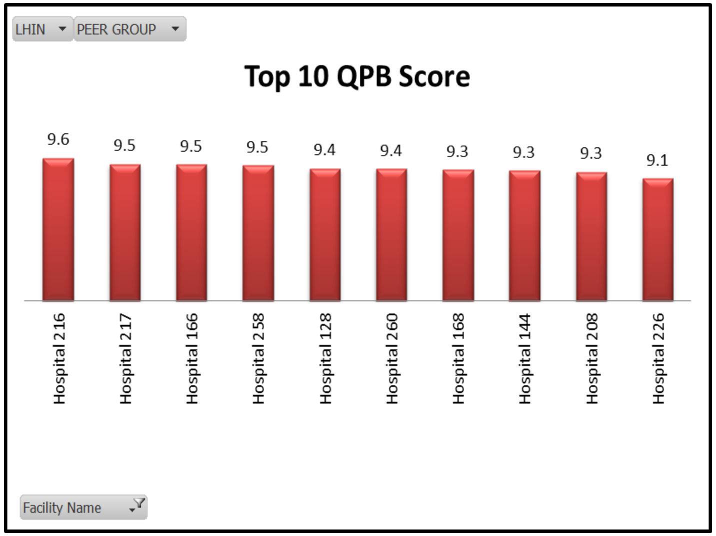


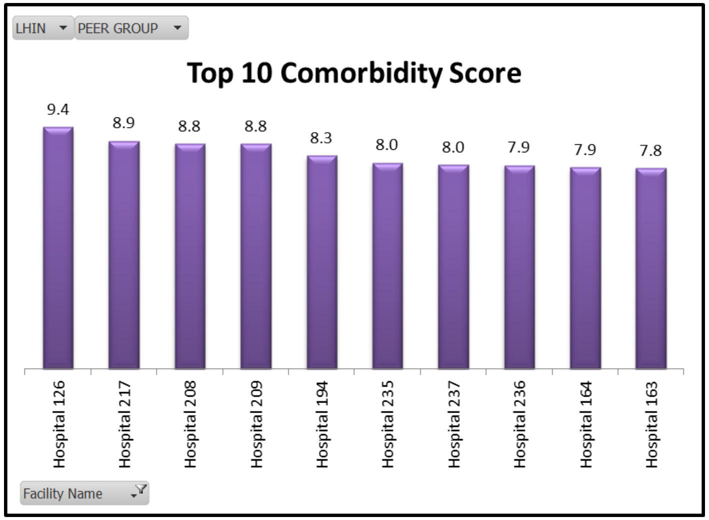
	Hospital 237	Hospital 217	Hospital 128	Hospital 208	Hospital 104	Hospital 164	Hospital 163	Hospital 209	Hospital 166	Hospital 213
Total SCU Score	9.87	9.24	9.43	7.52	9.85	8.33	9.04	7.11	7.28	10.00
Total QBP Score	8.85	9.49	9.38	9.29	8.88	9.09	6.85	8.53	9.47	8.30
Total Discharge to HC Score	5.86	3.53	6.58	5.33	4.64	5.35	6.03	4.40	4.94	3.97
Total Comorbidity Score	7.96	8.89	6.49	8.79	6.97	7.86	7.84	8.79	7.55	5.54
Overall Score	8.29	8.05	7.97	7.82	7.75	7.75	7.64	7.36	7.33	7.12

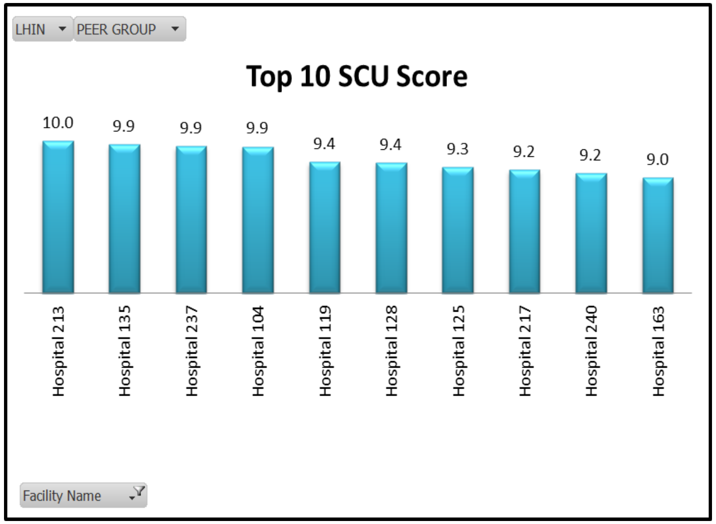
Facility Name

Dashboard Summary

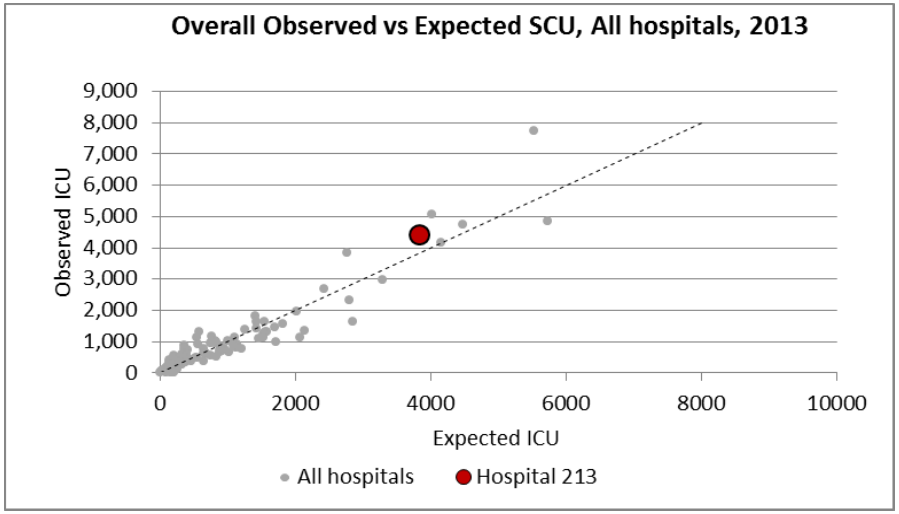


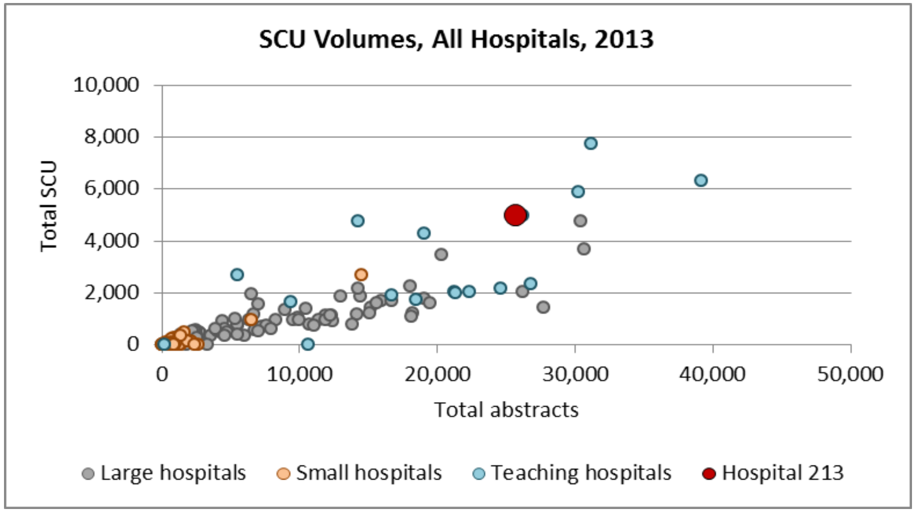


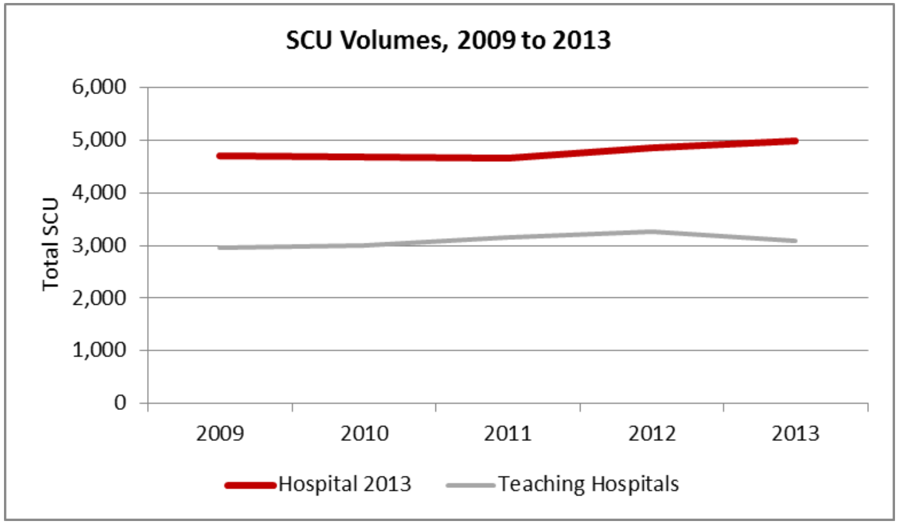


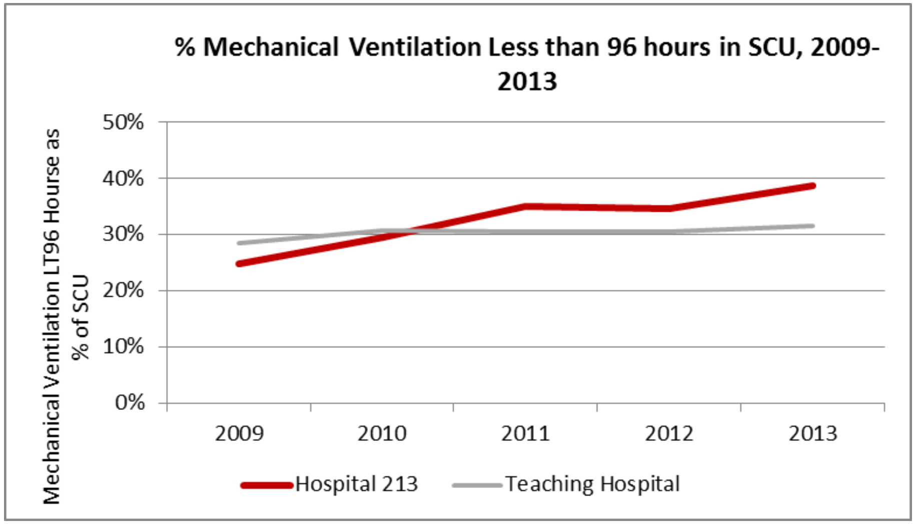


Facility Level Drilldown









Where do we go from here?

Next Steps

- Found anomalies in the data; we need to understand why they exist
- Continue to work collaboratively with Ontario Ministry of Health and Long-Term Care to ensure that this work adds value and can be used to improve the quality of the data used in the funding formula
- Apply knowledge and tools to other jurisdictions and areas in CIHI (health system performance indicators)

Big Data Insights

- Techniques are useful if used correctly
  - Techniques can identify lots of anomalies; needs to be targeted and have insight into which issues are important
  - Conclusions can only be as good as the models they are based on: need to assess model efficacy and robustness
- Data mining software (SAS Enterprise Miner)
  - Significantly increased staff productivity in developing and refining models
  - Easy to use interface, but need to know what you are doing
- Don’t forget about the power of simple statistics
- Need to be able to describe methods in plain language

Questions?

Thank you!

- For more information email:

  dataquality@cihi.ca

  LKirby@cihi.ca

  MKelly@cihi.ca