Sharing Information to Improve Outcomes

Routine Outcome Measurement: What's happening next?

Philip Burgess

November 2008

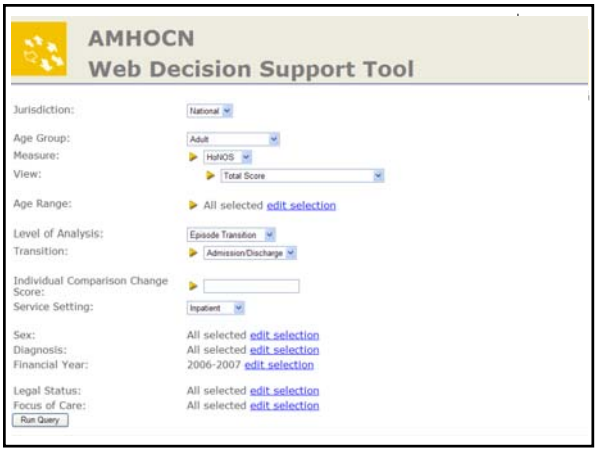
A joint Australian, State and Territory Government initiative

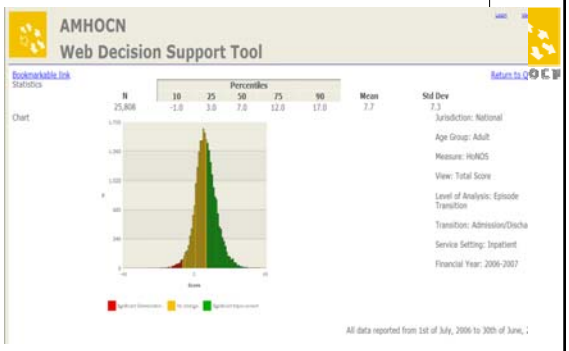
Background

- Very much work in progress – time to share some more ideas (as well as some anxieties);
- Last few years, much effort in the design, analysis & reporting of NOCC Data – cf WDST, Reports Portal, etc;
- Stakeholders becoming increasingly sophisticated in using these products;
- We need to build on these efforts to further enhance the utility of the NOCC.

A very brief update …

- Much debate and discussion regarding the enhancement of Decision Support Tools to assist in service delivery….
- Specifically, looking to enhance both the visual representation and implement also a series of clinical prompts…





A very brief update …

- Even more debate regarding how to report the 'effectiveness' of mental health services;
- Various contexts – both from the perspective of individual consumers and aggregate reporting;
- Simplistically, perhaps, but how many consumers are getting better, staying the same or getting worse?

A brief update …

- And there are many debates within these debates –
  - What method to use (Reliable Change Index or Effect Size);
  - What level of certainty (95%, 90% or 80% Confidence Intervals)
  - What 'episodes' should be considered 'in-scope' (e.g., Change of Setting from Ambulatory to Inpatient)

And the answers are:

- "Everyone's right and everyone gets a prize"
- "The NOCC data are only a fraction of the overall picture" – housing outcomes, vocational outcomes, 'quality of life';
- "And, whose outcome is it anyway?" The consumer, the service, the community

Adult Completed Episodes

	FY	N	Better	Same	Worse
Ambulatory	2004-2005	6351	53.2	41.1	5.7
	2005-2006	7609	53.9	39.4	6.7
	2006-2007	8446	53.9	40.2	5.9
Inpatient	2004-2005	16208	71.6	23.1	5.3
	2005-2006	18756	72.2	22.2	5.6
	2006-2007	21181	73.4	21.8	4.8

Where to now …

- If we accept that more debate and discussion is still to be had, arguably the next steps forward are to begin to consider how we can add further value to the reporting of outcomes;
- In particular, to what extent can our reporting take into account various factors that we well expect to influence outcome – diagnosis, legal status, sex, age & Focus of Care

Is this Casemix? Yes & No

- Typical applications of casemix are to designs models that explain the variation in costs per treated episode;
- That's what MH-CASC did in 1996 & NZ-CAOS did in 2002;
- But, other applications relate to 'risk adjustment' – there are reasonable candidate factors that could explain outcomes

Risk factors

- Insurers are particularly good at identifying risk factors – diet, smoking status, age, alcohol, physical exercise;
- "Premiums" reflect the insurer's "risk";
- As mental health service providers, understanding 'risk factors' assists in the planning and delivery of optimal services;
- How do we do this?

Severity Adjusted Effect Size

- Effect Size is a standardized method for reporting the magnitude of pre-post change;
- The Severity Adjusted Effect Size provides an estimate of effect size after adjusting for differences in case mix and severity of symptoms;
- It's also well known that one of the strongest predictors of outcome is severity status at baseline / admission to care;

Baseline Ratings & Outcomes

	Effect Size Index - Improved Difference Medium - 3 classes			
Baseline HoNOS Quartiles	Better	Same	Worse	Total
Bottom 5%	23.0%	68.8%	8.2%	100.0%
Q1 <25%	47.2%	44.8%	8.0%	100.0%
Q2 25-50%	64.7%	30.8%	4.5%	100.0%
Q3 50-75%	66.9%	29.2%	3.9%	100.0%
Q4 >75%	61.8%	34.5%	3.8%	100.0%
Total	55.9%	38.4%	5.8%	100.0%

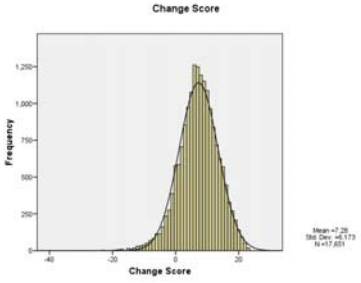
Some initial data considerations …

- Are the data 'good' enough?
- Are the data sufficiently complete?
- We know in any data set, extreme observations affect sample means & variability – what should we do about them?
- If the purpose is to measure change, what do we do about cases where expectation of change is minimal?

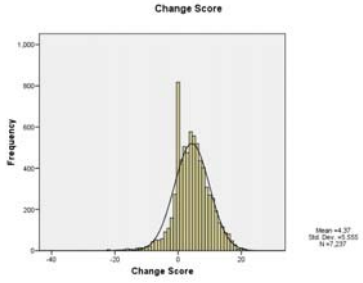
Some decisions…

- Remove extreme observations – more or less the top and bottom 5% - using IQR – a standard method used in classification analysis;
- Check baseline and change score distributions – hmmm – some issues but not easily solved by either 'transforming' the data or more complex rules for exclusion;

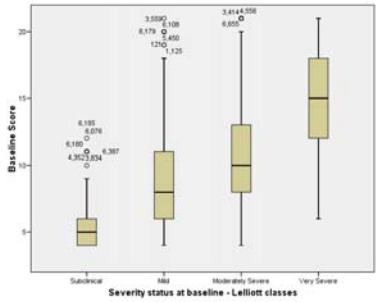
Change Scores – Adult Inpatient



Change Scores – Adult Ambulatory



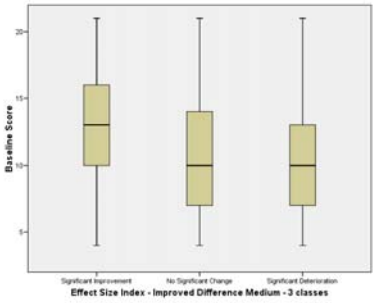
Baseline Score x Lelliott – Adult Ambulatory



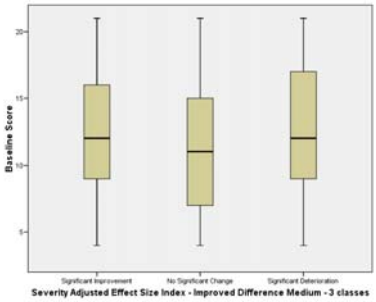
Analysis Strategy…

- Model both simple & complex risk adjustment models;
- Simple models only use Baseline Severity as an 'adjuster';
- Complex models include a range of other 'risk factors' – age, sex, legal status, diagnosis & focus of care

Unadjusted Effect Size – Adult Ambulatory



Severity Adjusted Effect Size – Adult Ambulatory



Pondering time …

- What's going on?
- Lelliott's classes look ok – but of course they are confounded with the baseline measure …
- The Effect Size charts look suspicious – they're a little more intuitive for Inpatient Episodes (not shown) but not dramatically;
- So, what's it mean ….

Unadjusted vs. Adjusted Effect Size Classes – Adult Ambulatory

	Severity Adjusted Effect Size			Total
Unadjusted Effect Size	Better	Same	Worse	
Better	63.4%	36.6%	-	100.0%
Same	-	73.1%	26.9%	100.0%
Worse	-	-	100.0%	100.0%
Total	35.4%	48.5%	16.1%	100.0%

Understanding the logic …

- If there is no adjustment for severity, then the threshold for 'significant change' is FIXED;
- For example, for Adult Inpatient and Ambulatory, the critical threshold is '4';
- So, if consumer has a start of 8 & an end of 4
  - Unadjusted significant change
  - Severity adjusted no significant change…
  - Because there was a larger expectation of change when baseline is accounted for…

Where to next?

- Severity adjusted outcomes probably better represent 'actual' change;
- It's obvious too that services see a different mix of consumers in terms of severity and complexity;
- Profiling individual consumer outcomes can mask the mix of cases – we need to develop tools that gives relative 'weightings' to outcomes achieved by services…

Outcomes for 3 MHSOs

	Significant Improvement Classification		
MHSO	Better	Same	Worse
A	33.0	53.8	13.1
B	21.6	42.6	35.8
C	36.6	56.1	7.3
Total	31.1	51.5	17.3

Severity Profiles for 3 MHSOs

	Severity Class - Quartiles				
MHSO	<25%	25-50%	50-75%	>75%	Episodes
A	44.3	25.8	25.8	4.1	221
B	38.3	25.9	27.8	8.0	162
C	55.1	24.9	17.6	2.4	205
Total	46.4	25.5	23.5	4.6	588

Severity Weights for 3 MHSOs

MHSO	Unweighted Severity	Weighted Severity
A	1.000	0.988
B	1.000	1.193
C	1.000	0.860
Total	1.000	1.000