



New Zealand Mental Health Consumers and their Outcomes

Professor Kathy Eagar

**Director, Centre for Health Service Development,
University of Wollongong, Australia**



The reports

- Trauer T and Eagar K (2004) *New Zealand Mental Health Consumers and their Outcomes*. Health Research Council of New Zealand: Auckland
- Plus separate outcomes reports for the 8 participating DHBs



NZ-CAOS Project

- **Funded under the NZ Mental Health Research & Development Strategy four priority areas:**
 - **Epidemiology**
 - **Outcomes**
 - **Quality and Best Practice**
 - **Casemix**
- **Planning for a casemix study commenced in 1995 with three services expressing an interest in the equivalent Australian MH-CASC project.**
- **Based on a recognition of the need for casemix information tools in mental health.**

Project Objectives

- **Primary Objective:** To develop the first version of a national casemix classification for specialist mental health services in New Zealand that builds on the classification developed in the Australian MH-CASC Project.
- **Secondary Objective:** To trial the introduction of outcome measurement into routine clinical practice.



8 participating DHB sites

- Northland
- Waitemata (forensics services only)
- Auckland
- Counties-Manukau (excl adult community)
- Waikato (excl adult community services)
- Lakes
- Capital & Coast
- Otago

NB: The study excluded all A&D and NGO services

Criteria for site selection:

1. The degree to which they were representative of New Zealand mental health services
2. The extent to which they provided a comprehensive range of services
3. Their combined capacity to provide a sufficient volume and mix of 'consumer cases'
4. The suitability of information infrastructure to collect the required clinical data, track service utilisation at the individual consumer level and assign costs; and
5. The readiness of the organisations to absorb the demands arising from a casemix study.



The basic idea behind the study design

- **Consumer needs** - as measured by clinical attributes and demographic characteristics
- drives **resource utilisation** - as measured by clinician time, length of inpatient stay, pharmacy, pathology and other service use.

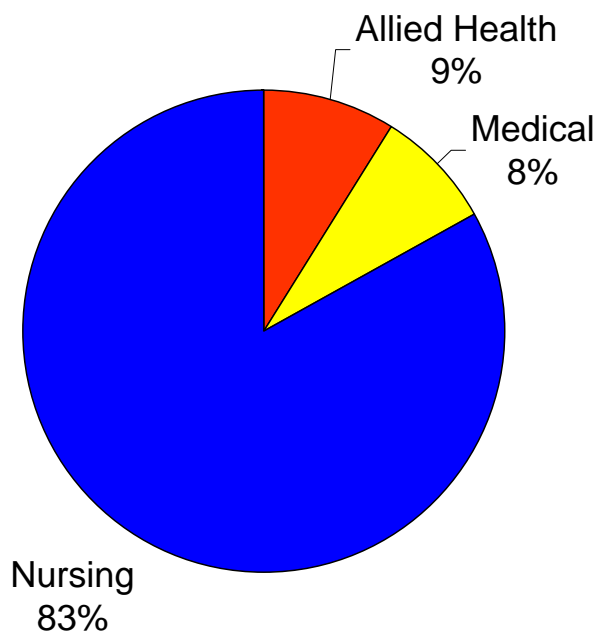
Consumer Profile

- **8 DHBs saw 12,576 consumers**
- **98% of consumers received services from only one DHB**
- **55% of consumers had one episode only, 39% had two, 6% had more than two**
- **53% male, 47% female**
- **19.8% Maori**
- **4.9% Pacific Island**
- **75.3% all other ethnicity groups**

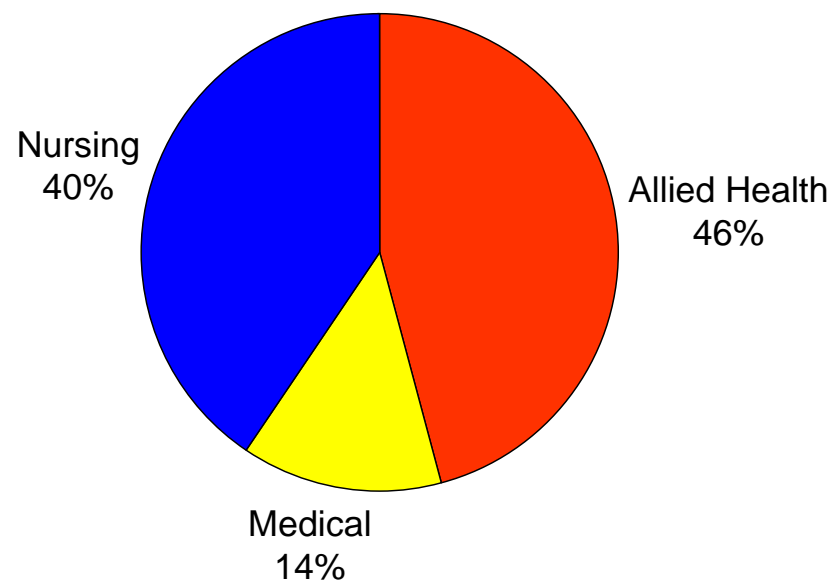
What clinical staff were involved?

Approx. 2,000 full time equivalent staff

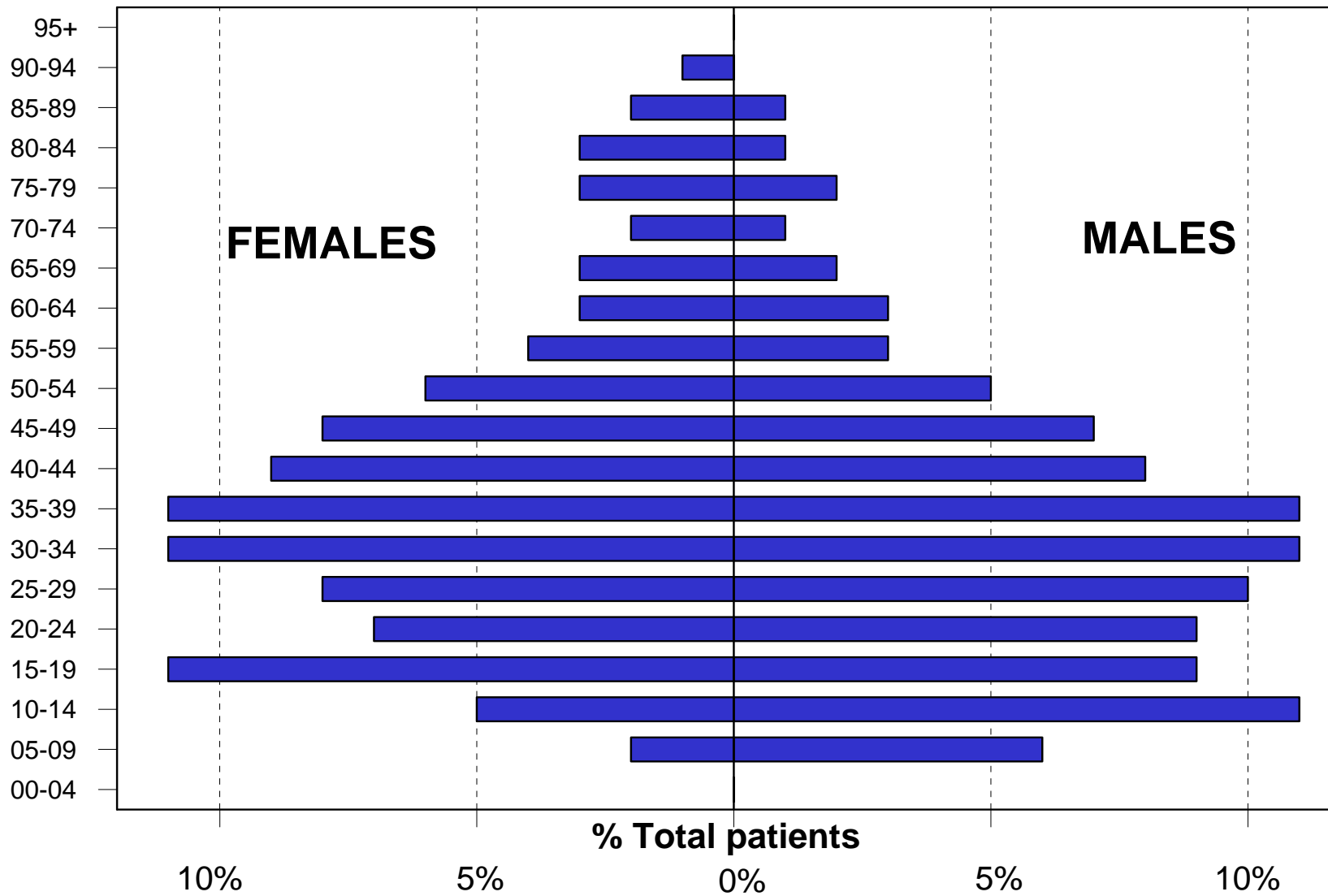
Inpatient Services



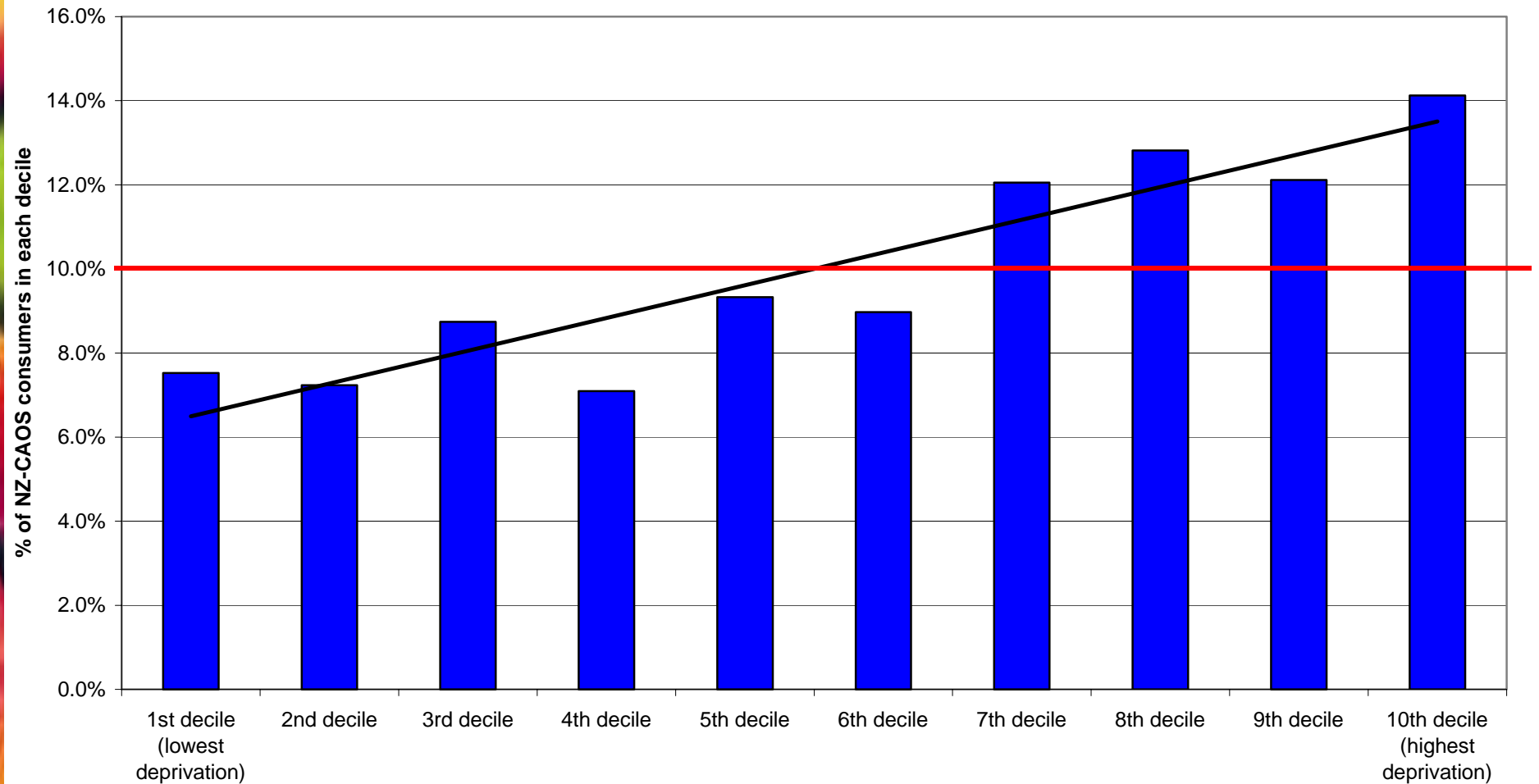
Community Services



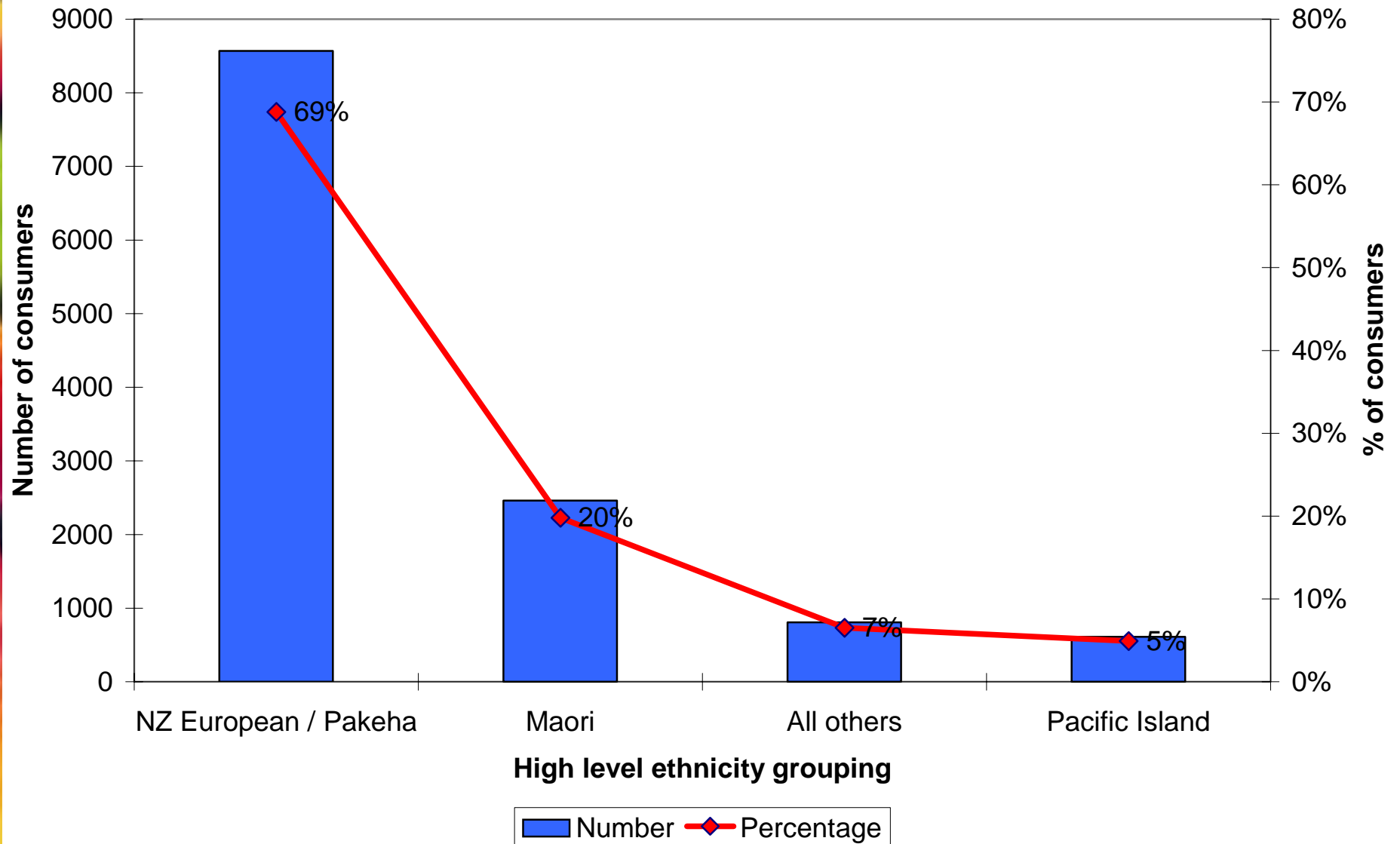
Age profile



Profile of social deprivation

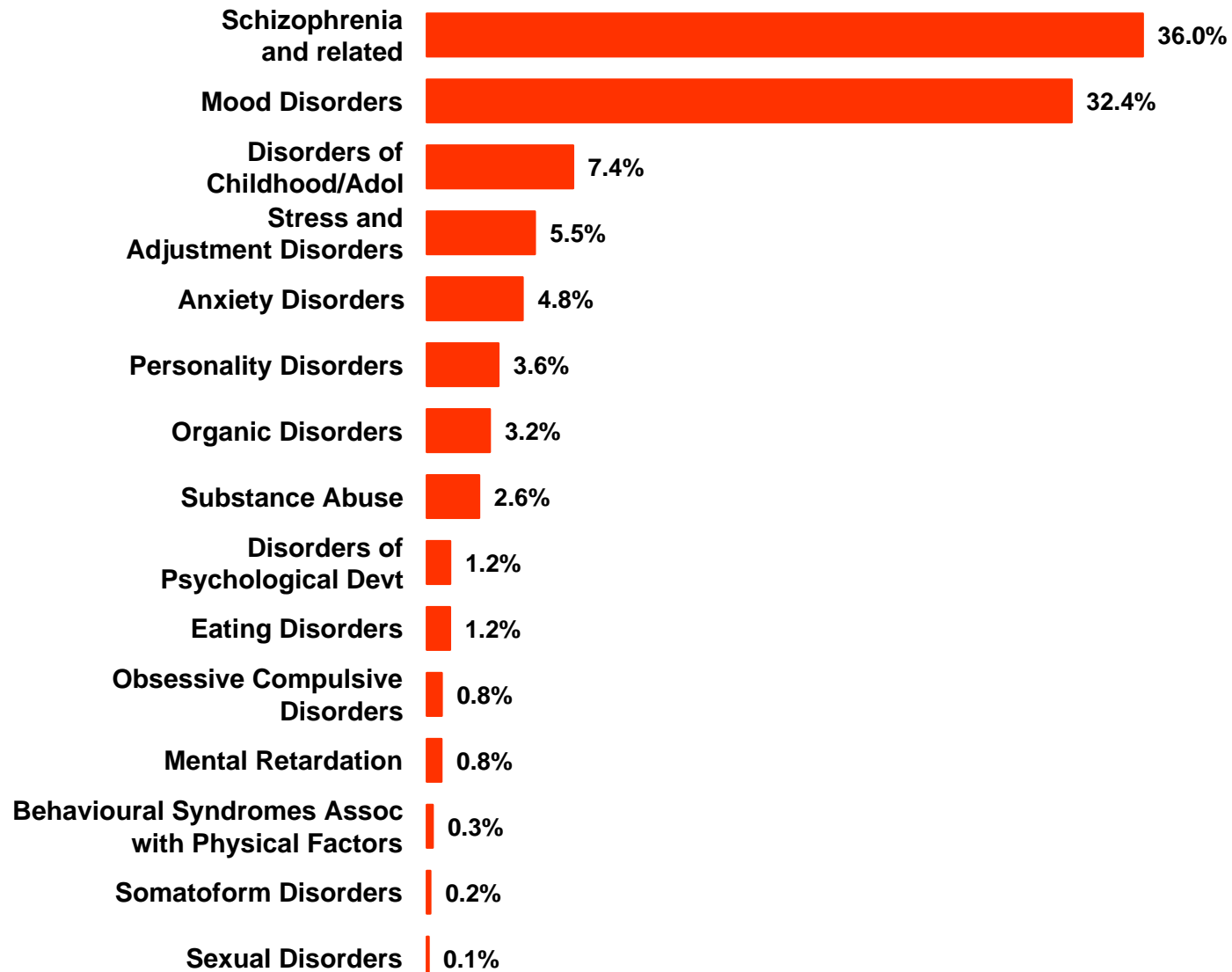


High level ethnicity profile



Episode Profile

Diagnosis, all Episodes





The Outcome Measures

- **Compliance and completion**

The outcome measures

Consumer Population	Clinical ratings of severity	Levels of functioning
Adults	<p><u>Health of the Nation Outcome Scales (HoNOS)</u></p> <p>The HoNOS is rated on a scale of 0 to 4. The higher the score, the greater the severity of the problem.</p>	<p><u>Life Skills Profile (LSP-16)</u></p> <p>The LSP is rated on a scale of 0 to 3. The higher the score, the greater the level of functional impairment.</p>
Consumers over the age of 65 years	<p><u>Health of the Nation Outcome Scales for consumers over the age of 65 (HoNOS 65+)</u></p> <p>The HoNOS 65+ is rated on a scale of 0 to 4. The higher the score, the greater the severity of the problem.</p>	<p><u>Resource Utilisation Group Activities of Daily Living (RUG-ADL)</u></p> <p>The RUG-ADL is rated on a scale from 1 to 4. The higher the score, the greater the level of functional impairment.</p>
Children and adolescents	<p><u>Health of the Nation Outcome Scales for Children and Adolescents (HoNOSCA)</u></p> <p>The HoNOSCA is rated on a scale of 0 to 4. The higher the score, the greater the severity of the problem.</p>	<p><u>Children's Global Assessment Scale (CGAS)</u></p> <p>The CGAS is a single rating on a scale of 0 to 100. The lower the score, the lower the level of function</p>

Plus 2 measures to aid interpretation - Focus of Care and FIHS

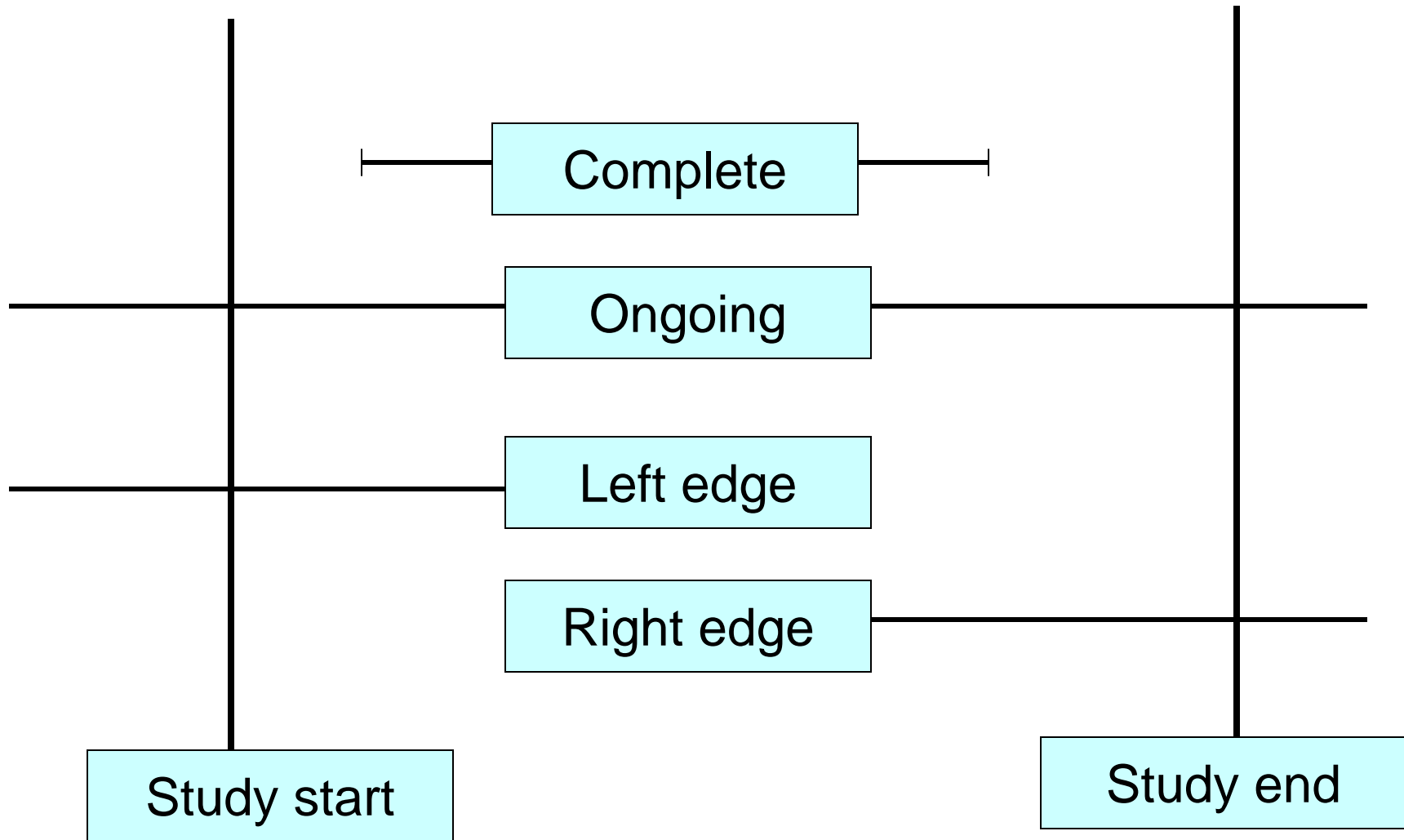


Number of clinical ratings (compliance)

All ratings	Start	End	Total
Both start and end ratings	18,655	18,655	37,310
Start or end, but not both	10,206	993	11,199
Total	28,861	19,648	48,509

Instrument	Start	End	% total episodes
HoNOS	7,805	7,805	57.5%
LSP	7,140	7,140	52.6%
HoNOSCA	1,716	1,716	52.6%
CGAS	1,994	1,994	61.2%
Total	18,655	18,655	

Compliance affected by CAOS being a snapshot survey



Percentage of ratings with no missing items (completion)

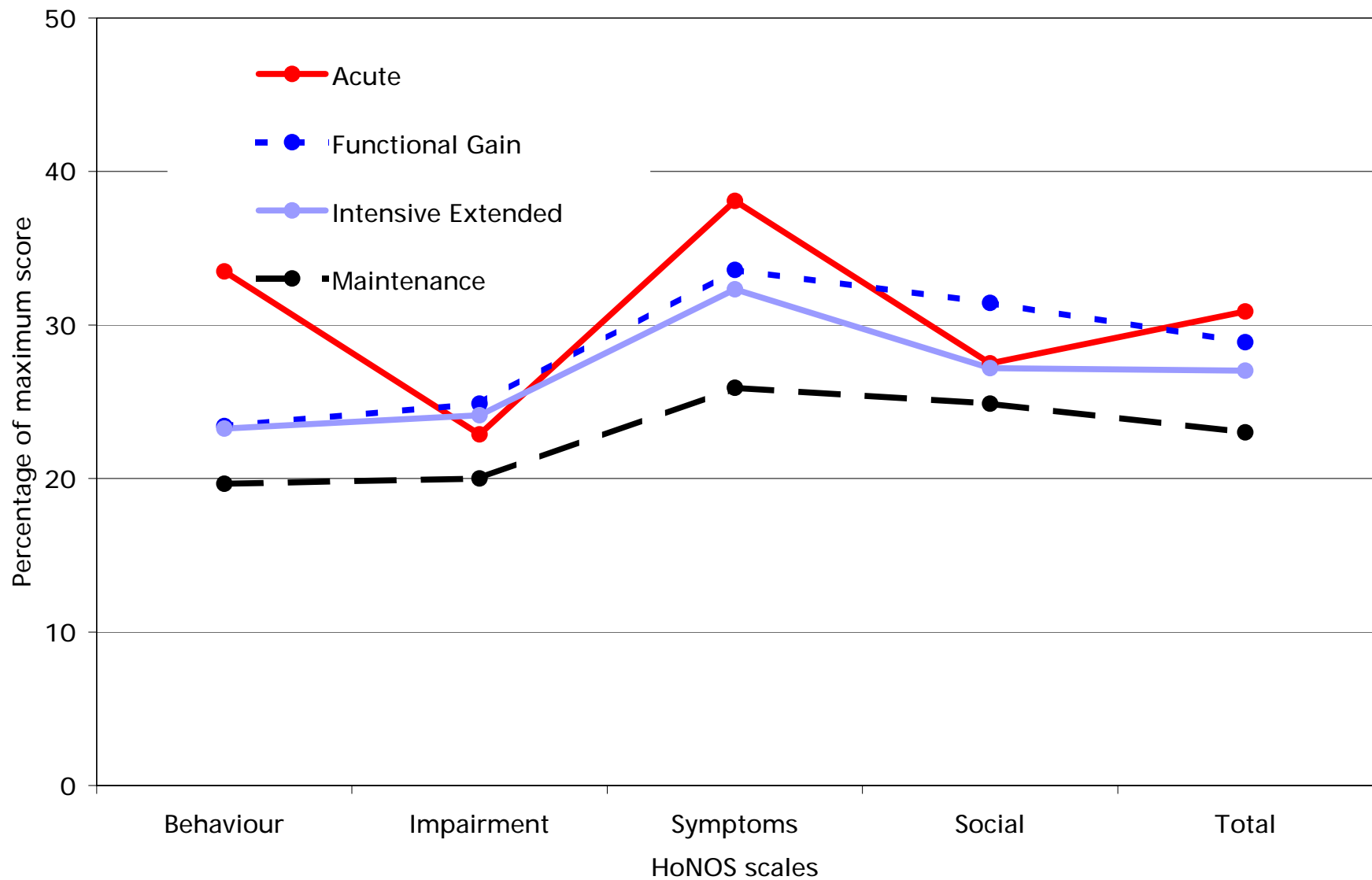
Instrument	Start	End
HoNOS	90.1%	92.7%
LSP-16	98.1%	98.7%
HoNOSCA	95.8%	95.0%
FIHS	94.1%	94.3%



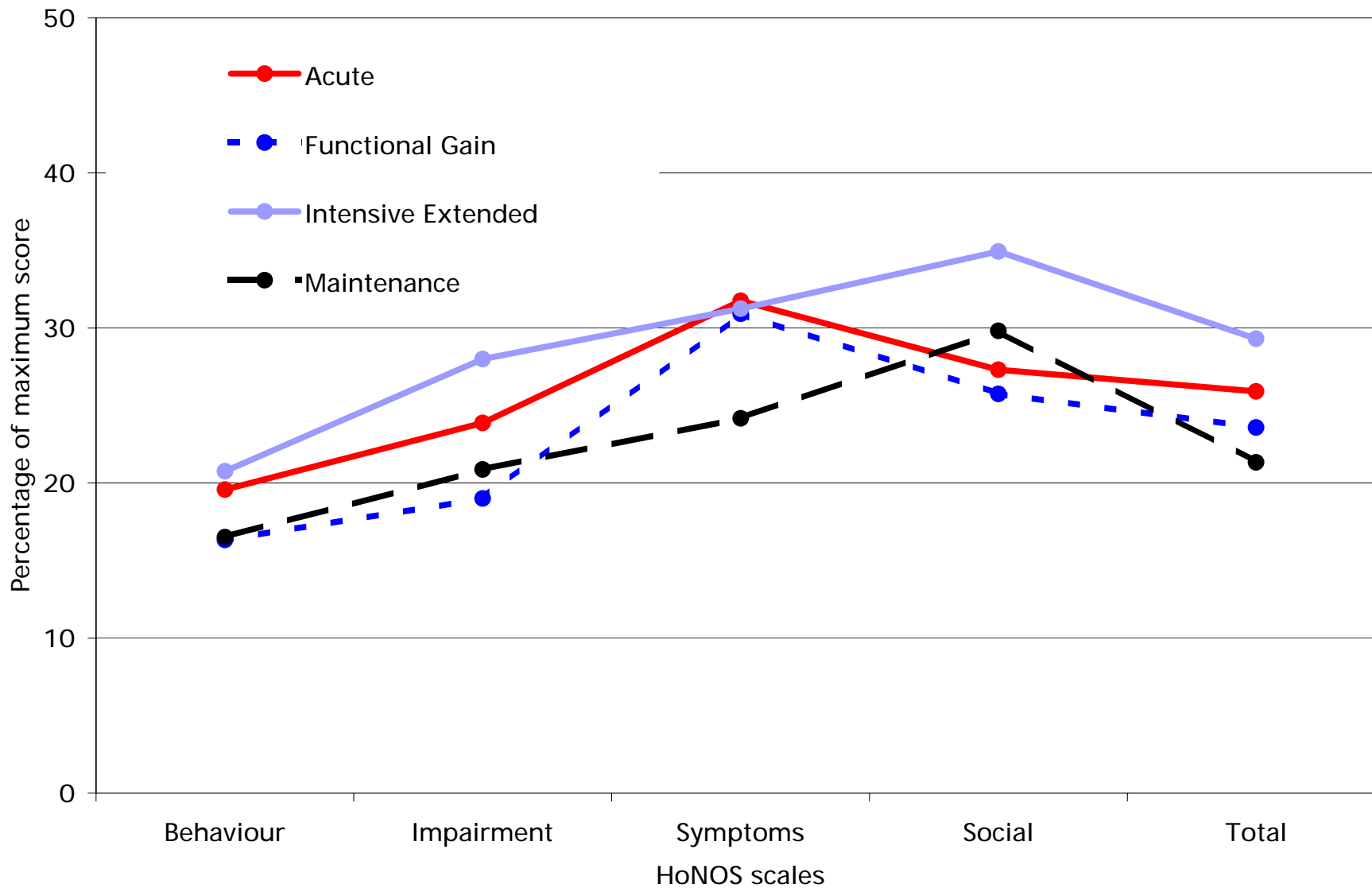
Do the Focus of Care and Factors Influencing Health Status help to interpret outcome ratings?



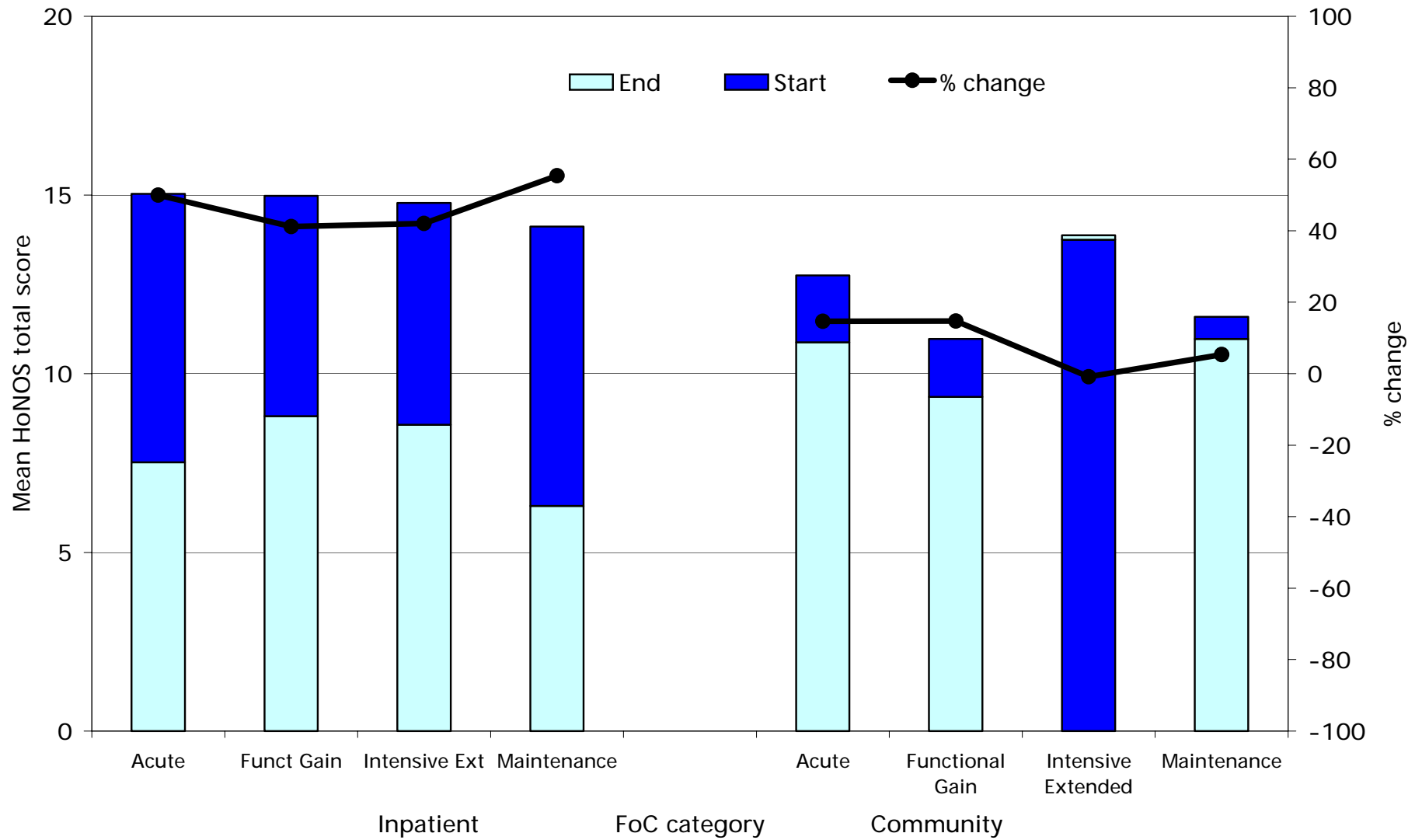
The HoNOS (at episode start) and FOC - inpatient



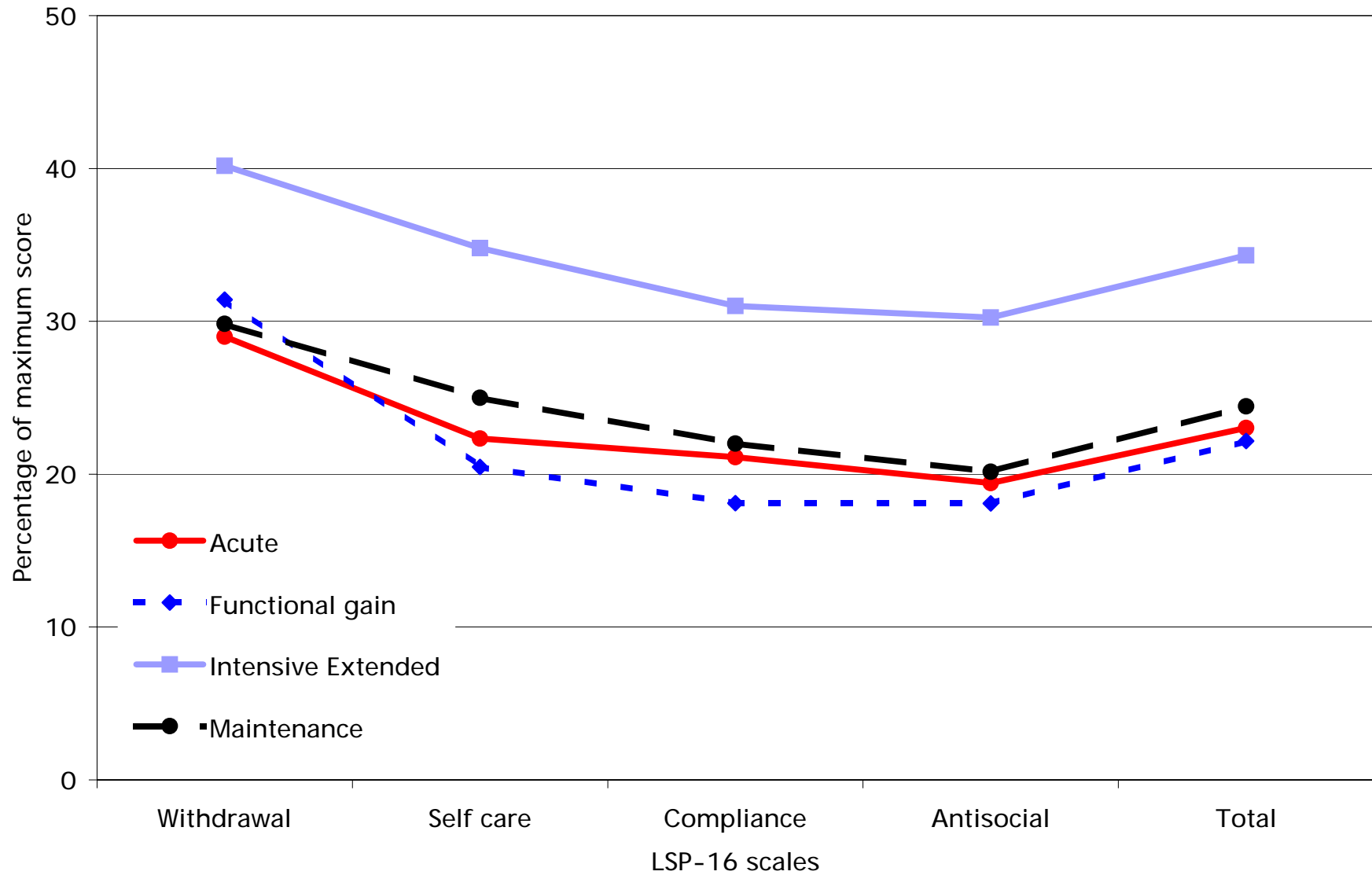
The HoNOS at episode start and FOC - community



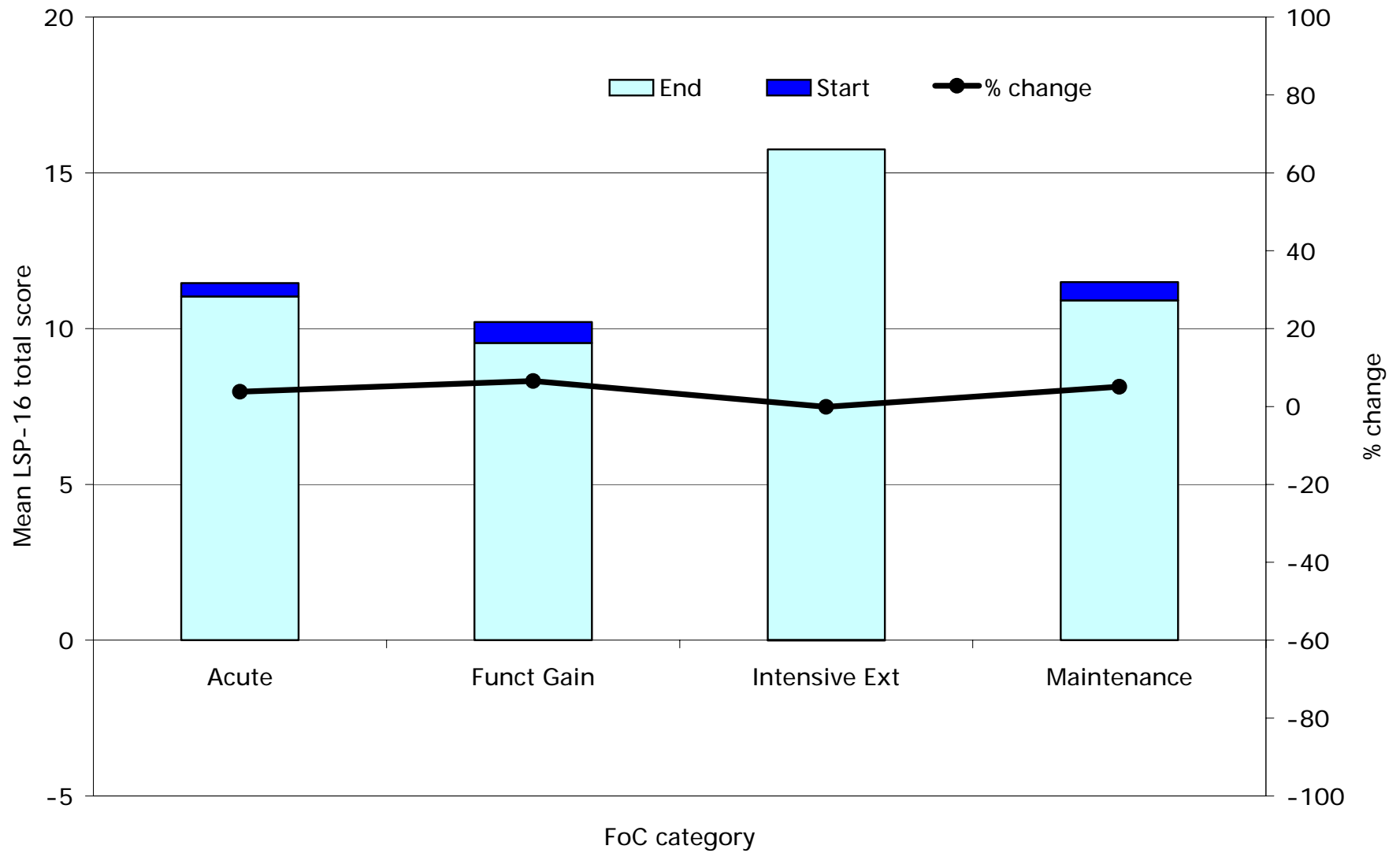
Change on the HoNOS



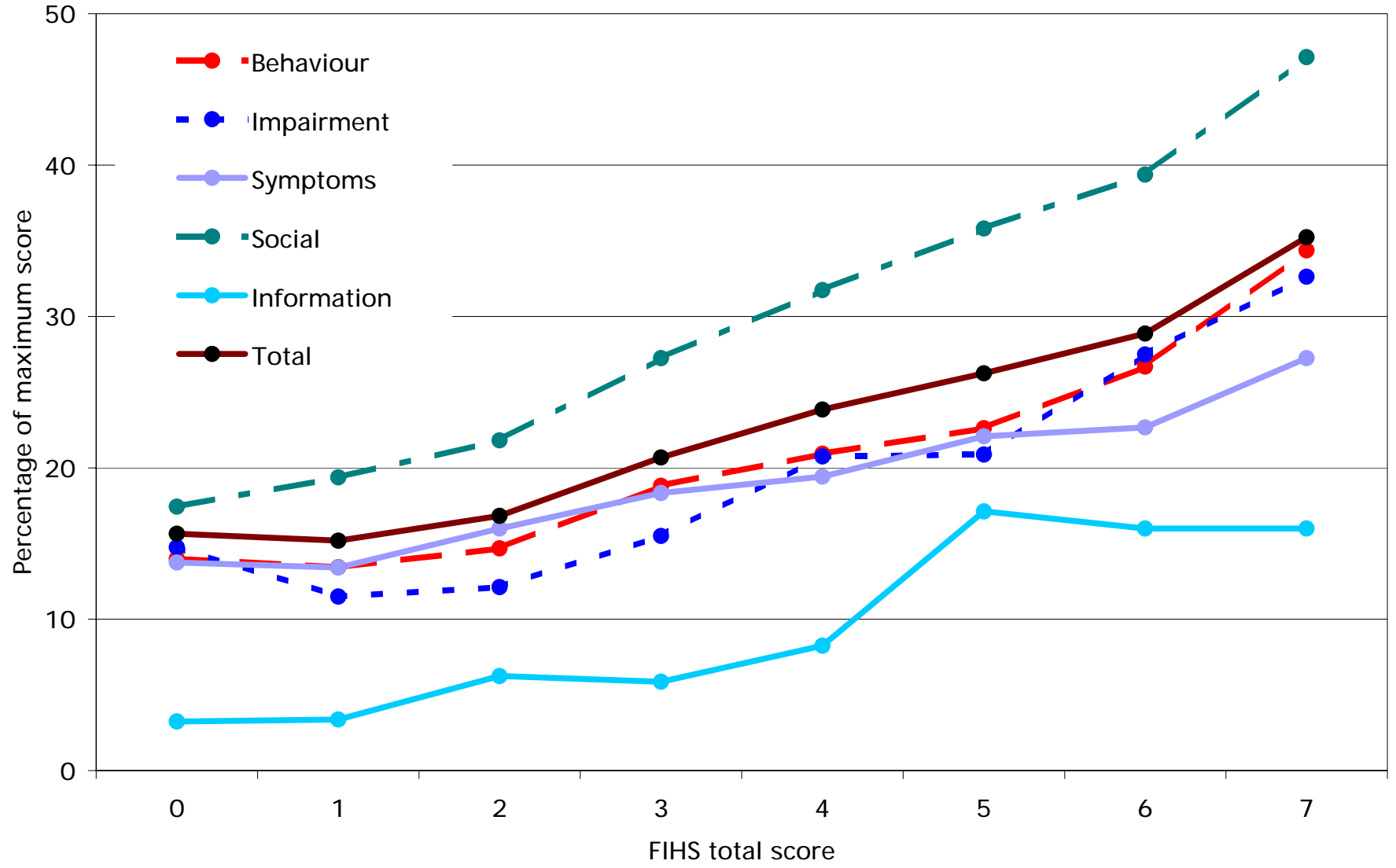
The LSP-16 at episode start and FOC - community



Change on the LSP-16



The HoNOSCA and FIHS - community



Implications

- **FIHS aids interpretation of the Child and Youth measures**
- **With the exception of ‘Intensive extended’ in the community, Focus of Care doesn’t help much**



Relationships between clinical ratings and other variables

- Age, sex, diagnosis, legal status

Relationship between HoNOS scores & other variables



- Total score unrelated to age.
- No clinically significant sex differences.
- Highest HoNOS total scores - Substance misuse, Personality disorder & Organic disorders.
- Involuntary status - associated with higher scores on the Behaviour, Impairment, Social subscales and the total score.
 - Involuntary - higher scores on the Hallucinations / delusions item
 - Voluntary - higher scores on the depression related items.

Relationship between LSP-16 scores & other variables



- Only 1 item significantly correlation with age
 - Item 16: What sort of work is this person generally capable of?
- Males - worse ratings on 15 of 16 items, all subscales and total score. However, actual differences small on all 16 items.
- Higher scores - Organic, Substance misuse, Schizophrenia & Personality Disorders
- Lower scores - Mood, Anxiety, Obsessive-Compulsive, Stress and Eating Disorders
- Involuntary - higher scores on all four LSP-16 subscales and the total score.

Relationship between HoNOSCA scores & other variables



- Total score unrelated to age, but differences at the item level.
- On average, males total scores about one point higher than females.
- Higher scores - Retardation, Schizophrenia, Developmental and Substance Misuse
- Lower scores - Eating Disorders.

Relationship between CGAS, FIHS & other variables



- CGAS scores unrelated with age.
- Males - higher (better) ratings in inpatient settings and
- Females - higher ratings in community.
- Lower (worse) CGAS ratings - Retardation, followed by diagnoses of Schizophrenia and Developmental disorder.
- Negligible correlation between the FIHS total score and age.
- Most frequent problem - Item 4 (Problems related to primary support group, including family circumstances)
- 6 of the 7 FIHS items more common for females
- Highest (worse) FIHS scores - Substance misuse and Stress disorders.



Painting by Serena Young



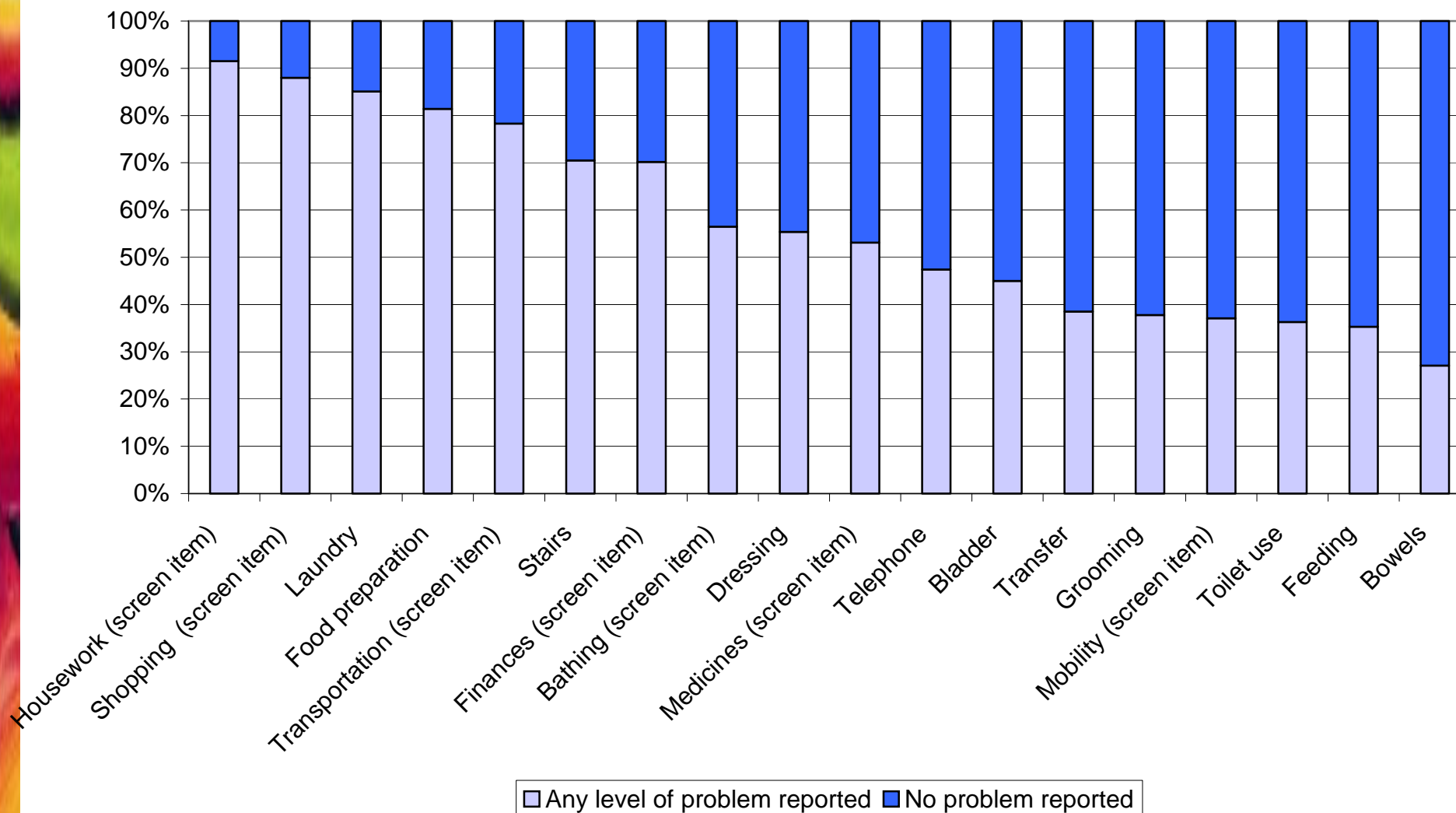
The problems with the LSP-16 and the RUG-ADL for older people

- And why a better measure is required

Functional hierarchy - early loss and late loss ADLs

- **People lose functional abilities in the opposite order to which they acquire them**
- **‘Early loss’ ADLs like housework, transport, handling money, managing medicines (domestic functioning) are gained last and lost first**
- **‘Late loss’ ADLs like dressing, toileting, feeding and bed mobility (self-care) are gained 1st and lost last**
 - the RUG-ADL measures only very late loss ADLs
 - the LSP-16 does not capture where consumers sit on this functional hierarchy

Profile of the HACC population on the functional assessment





Painting by Serena Young



Implications

- The measures

The measures - implications

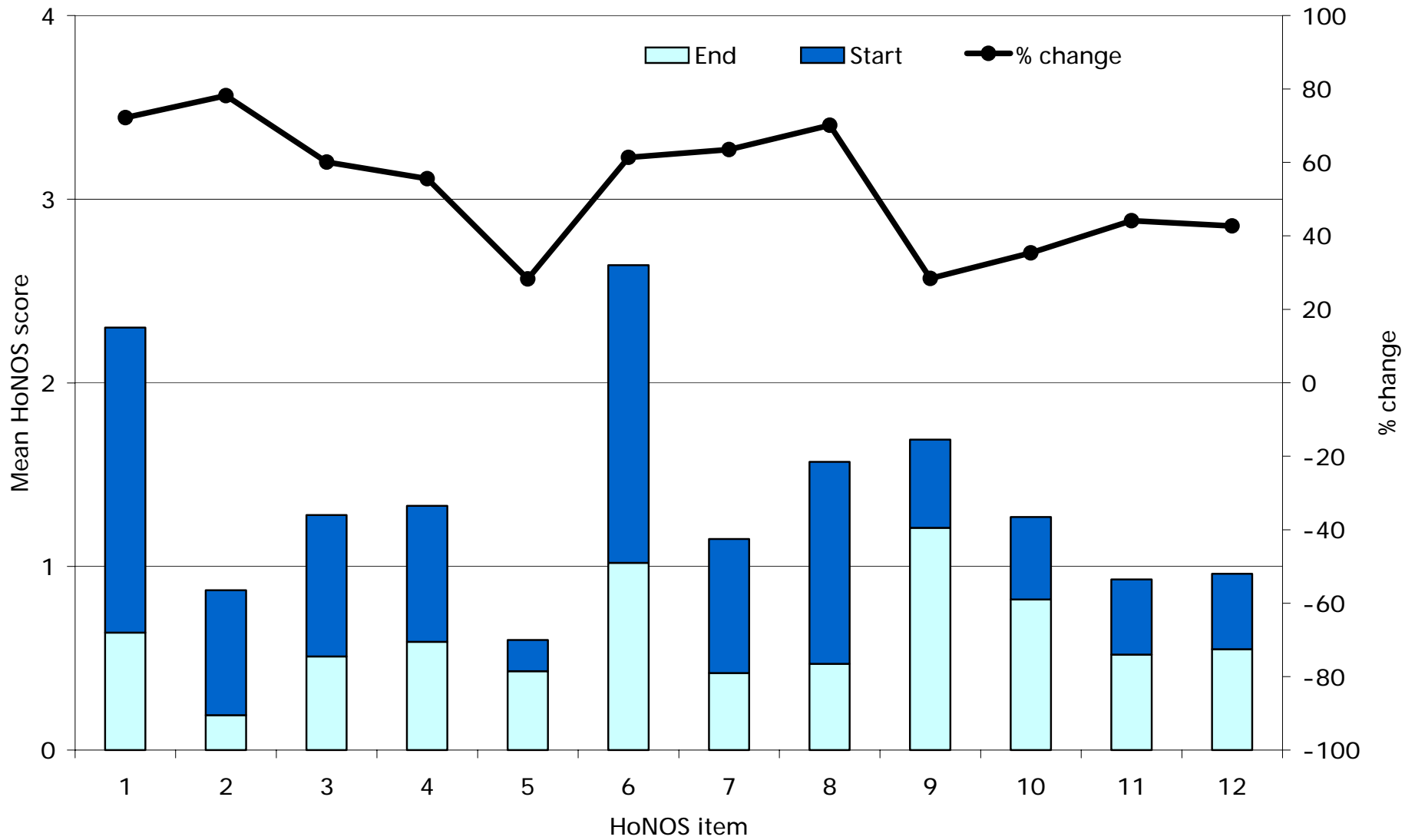
- **The HoNOS, the LSP-16, the HoNOSCA and the CGAS all work OK technically**
- **LSP-16 not acceptable**
 - **an alternate measure of function is required**
 - **consider both mainstream and mental health specific measures of function**
- **Work needed on the Focus Of Care**
 - **refine definitions – Acute, Functional Gain, Intensive Extended and Maintenance.**
- **Training, training, training!**
- **Need complementary consumer-rated and family/carer-rated measures.**



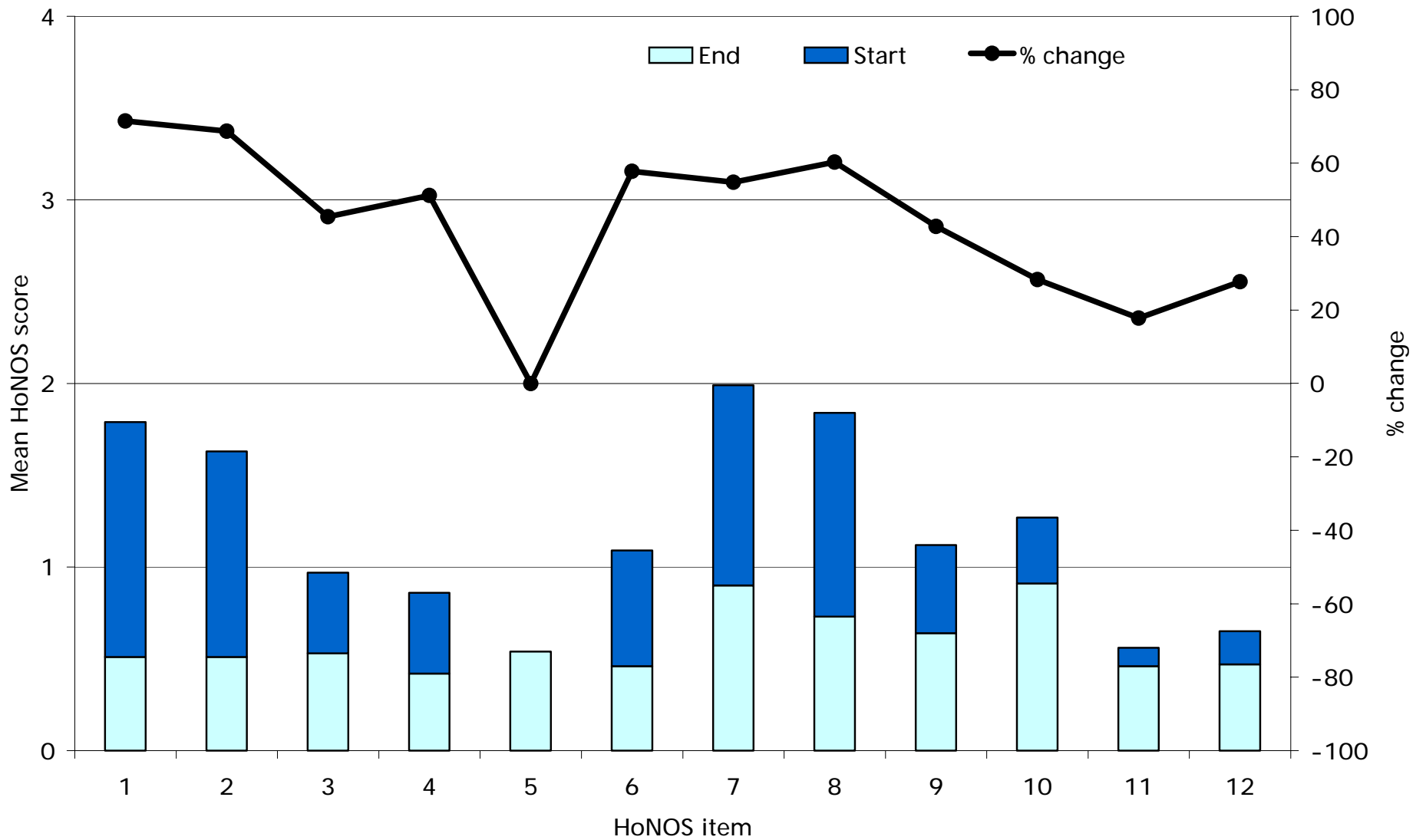
Do consumers change over the course of an episode?

- Yes, but amount of change depends on the measure and the setting

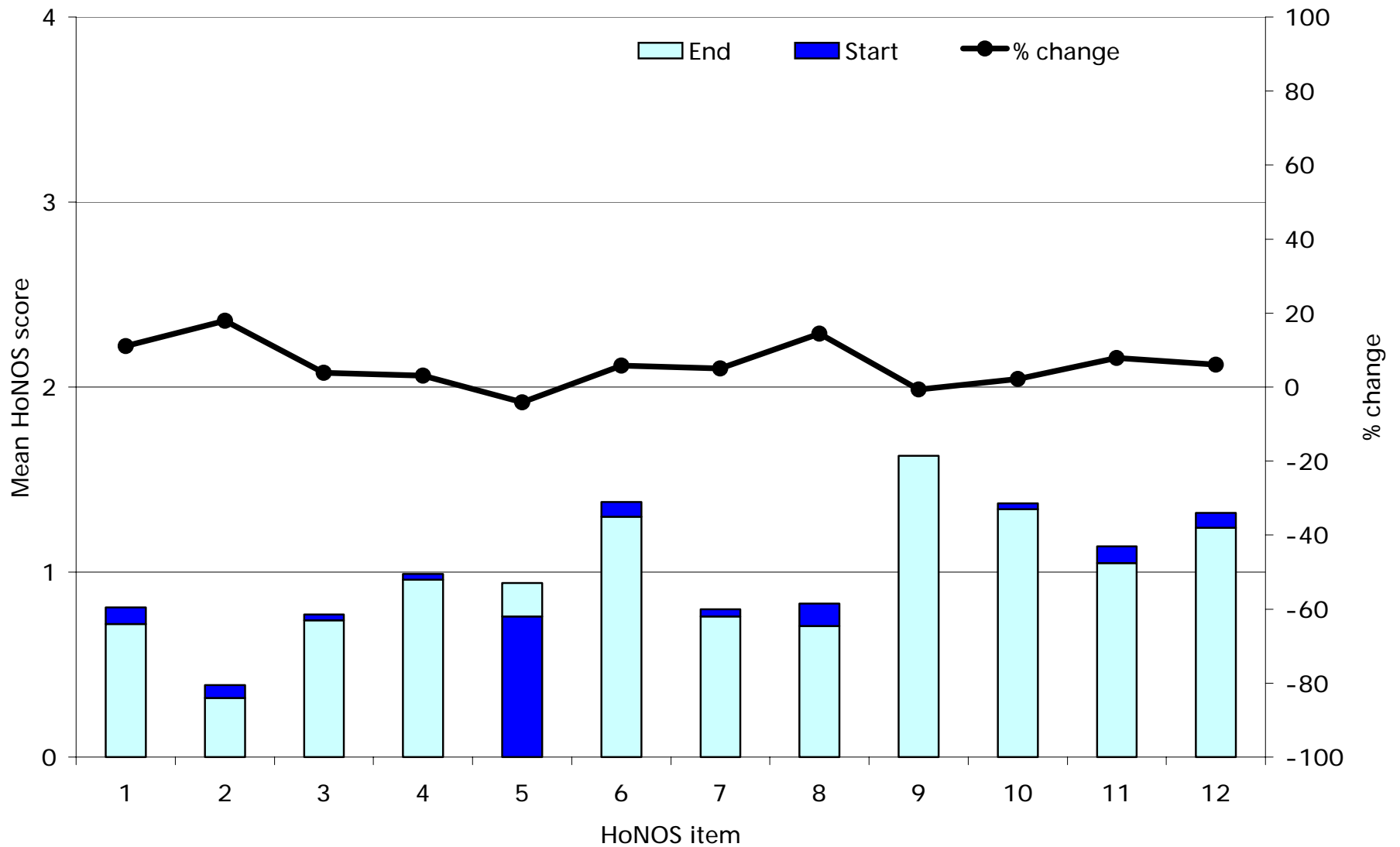
HoNOS, inpatient, schizophrenia



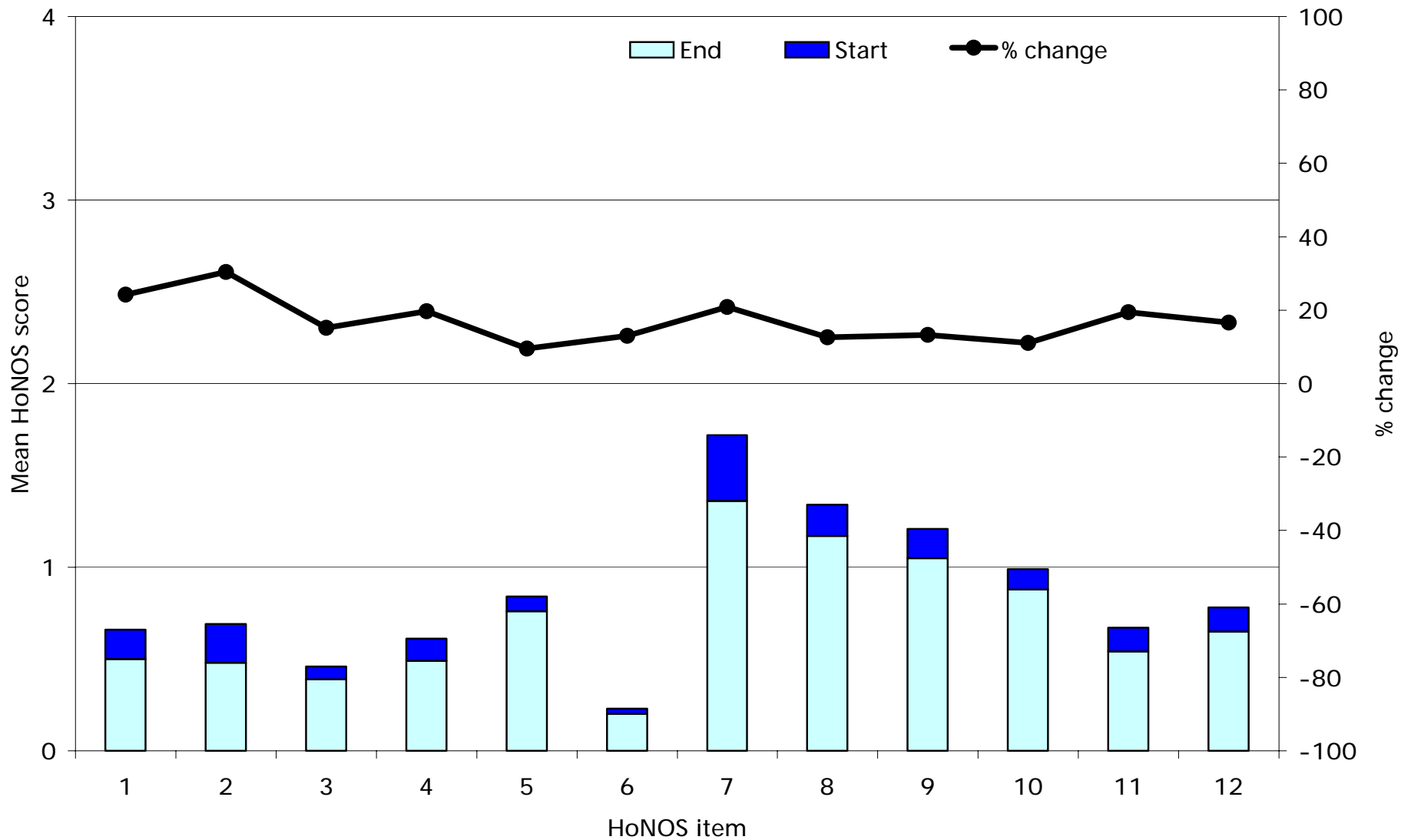
HoNOS, inpatient, mood



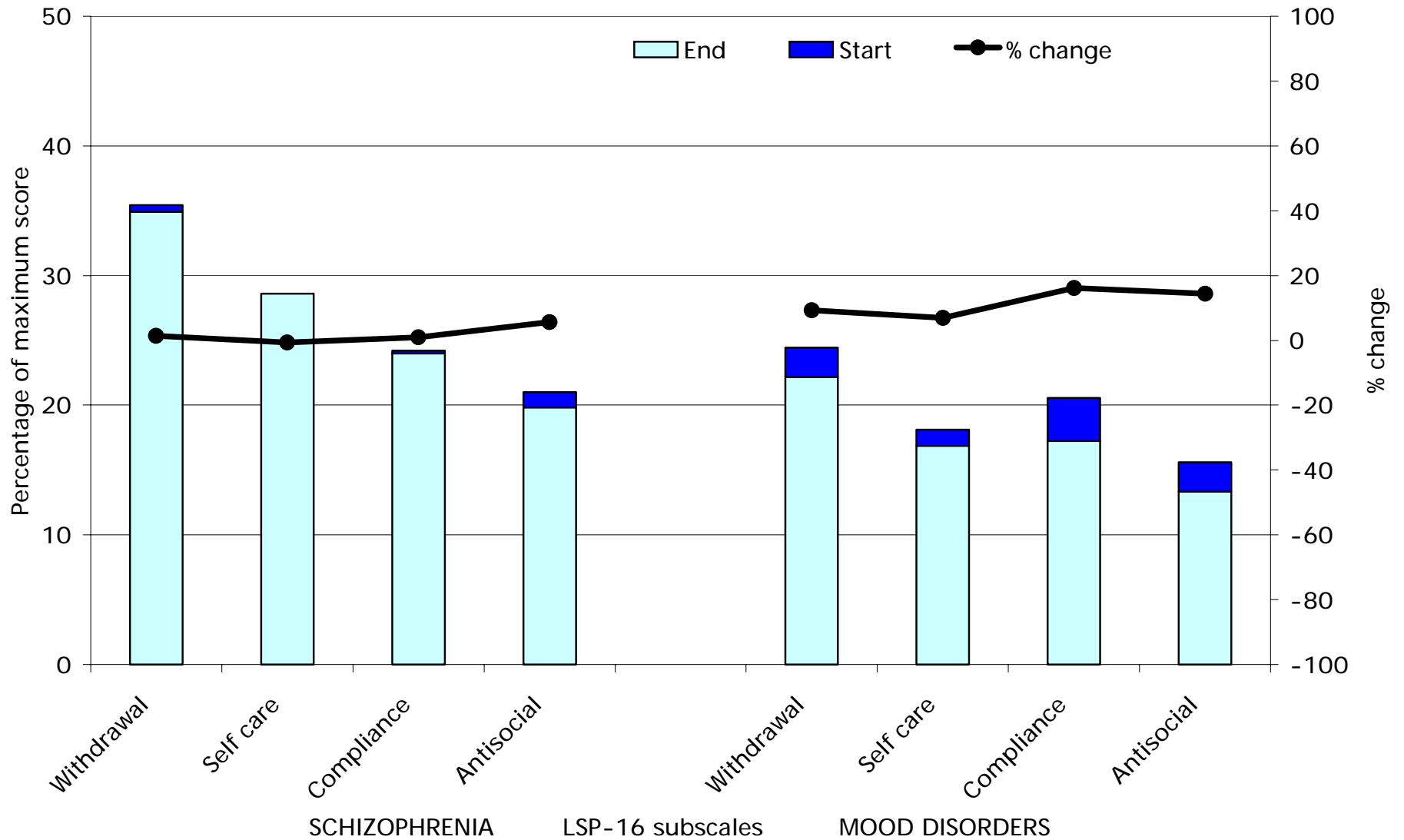
HoNOS, community, schizophrenia



HoNOS, community, mood



LSP-16, community, schizophrenia and mood



Sensitivity to change

Measure	Average percentage change from episode start to episode end
HoNOS in inpatient settings	50%
HoNOS in community settings	11%
LSP-16 in community settings	6%
HoNOSCA in community settings	17%
CGAS in community settings	9%

Implications

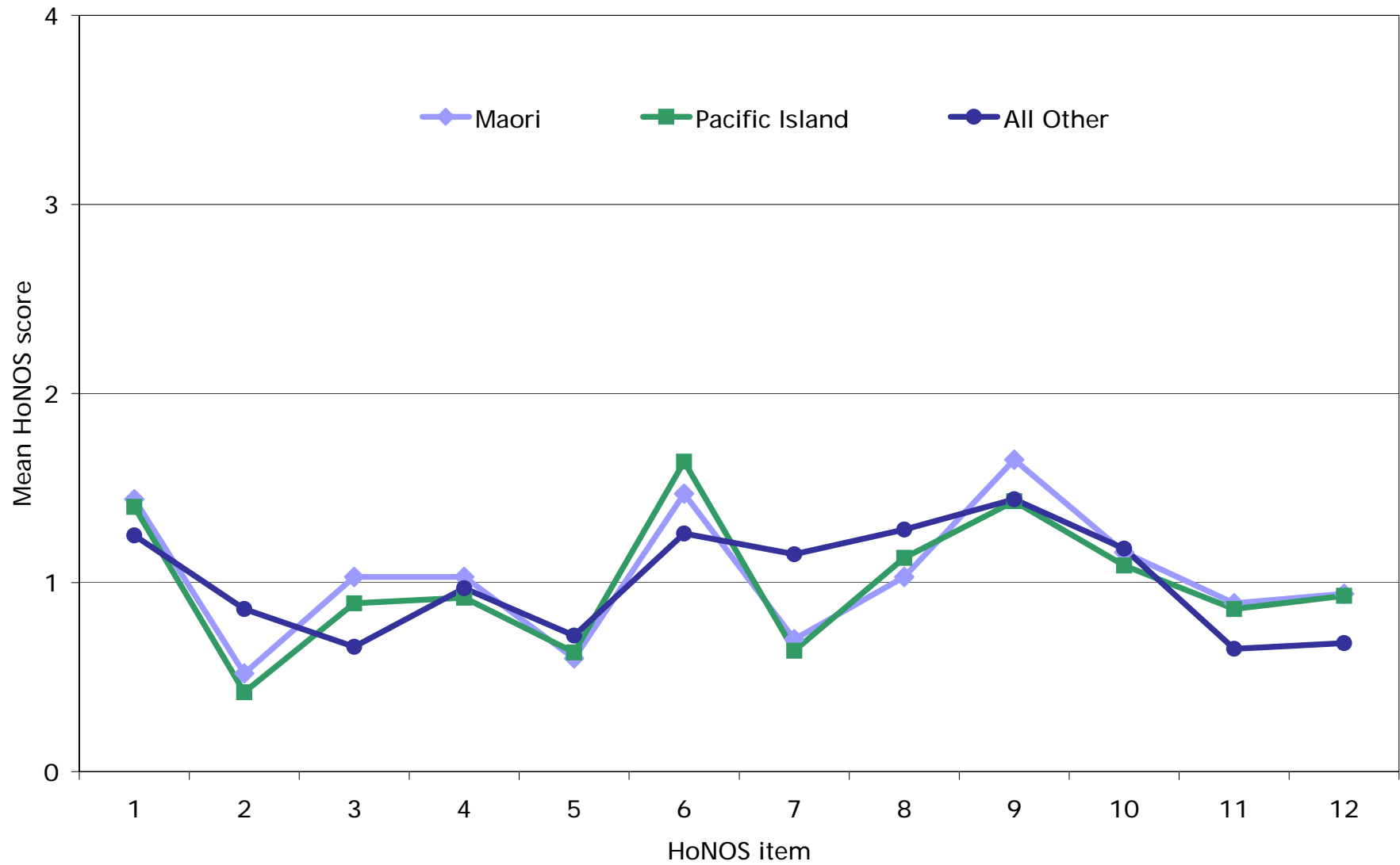
- **How much change you get depends on where you start.**
- **The measures are sensitive to change and appear suitable for use in benchmarking.**
- **However, outcomes achieved by different mental health services need to be casemix-adjusted to take into account the unique mix of consumers at each DHB.**



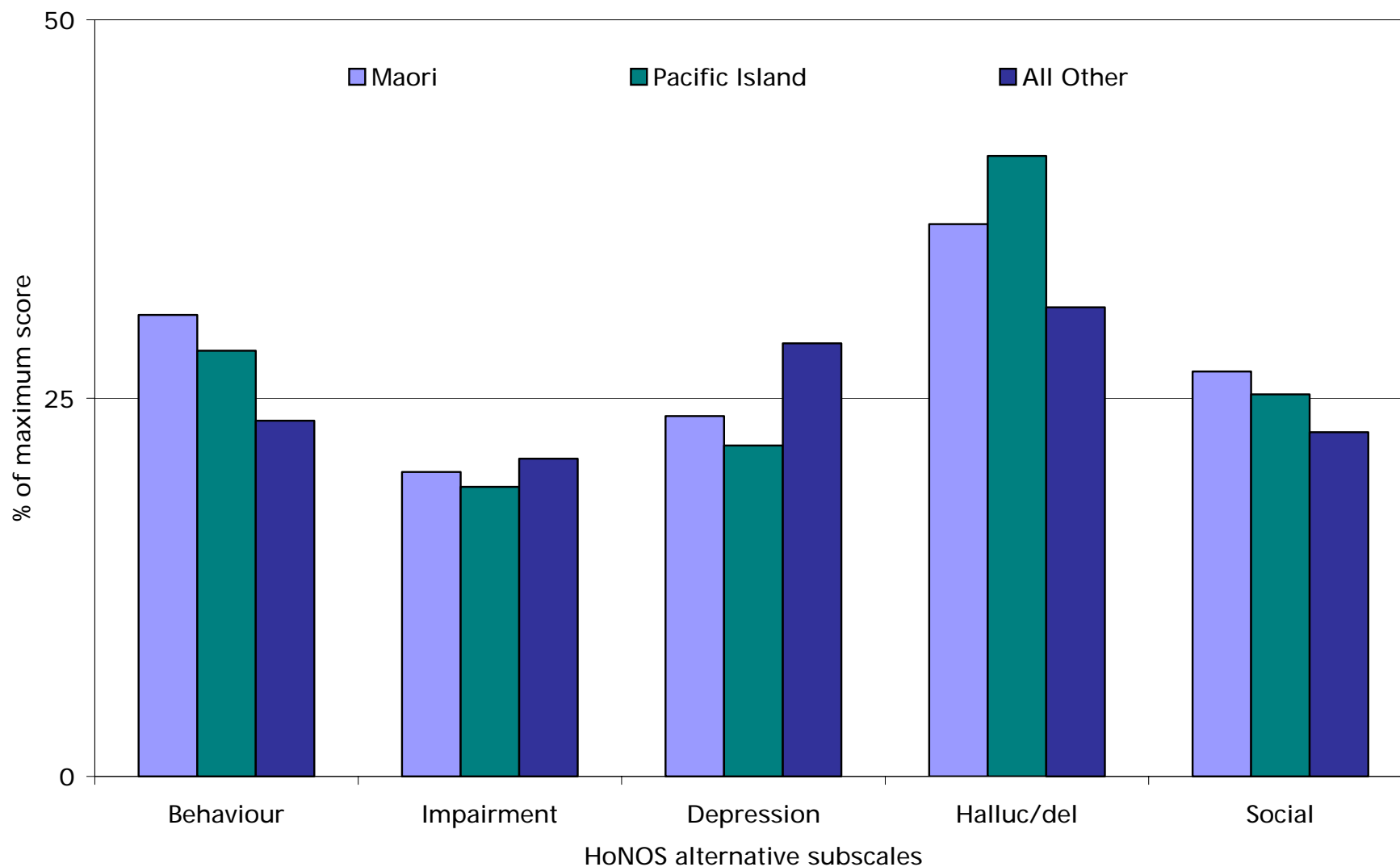
Comparison between outcomes for Maori, Pacific and Other Consumers

- But, first, a caveat on the data!

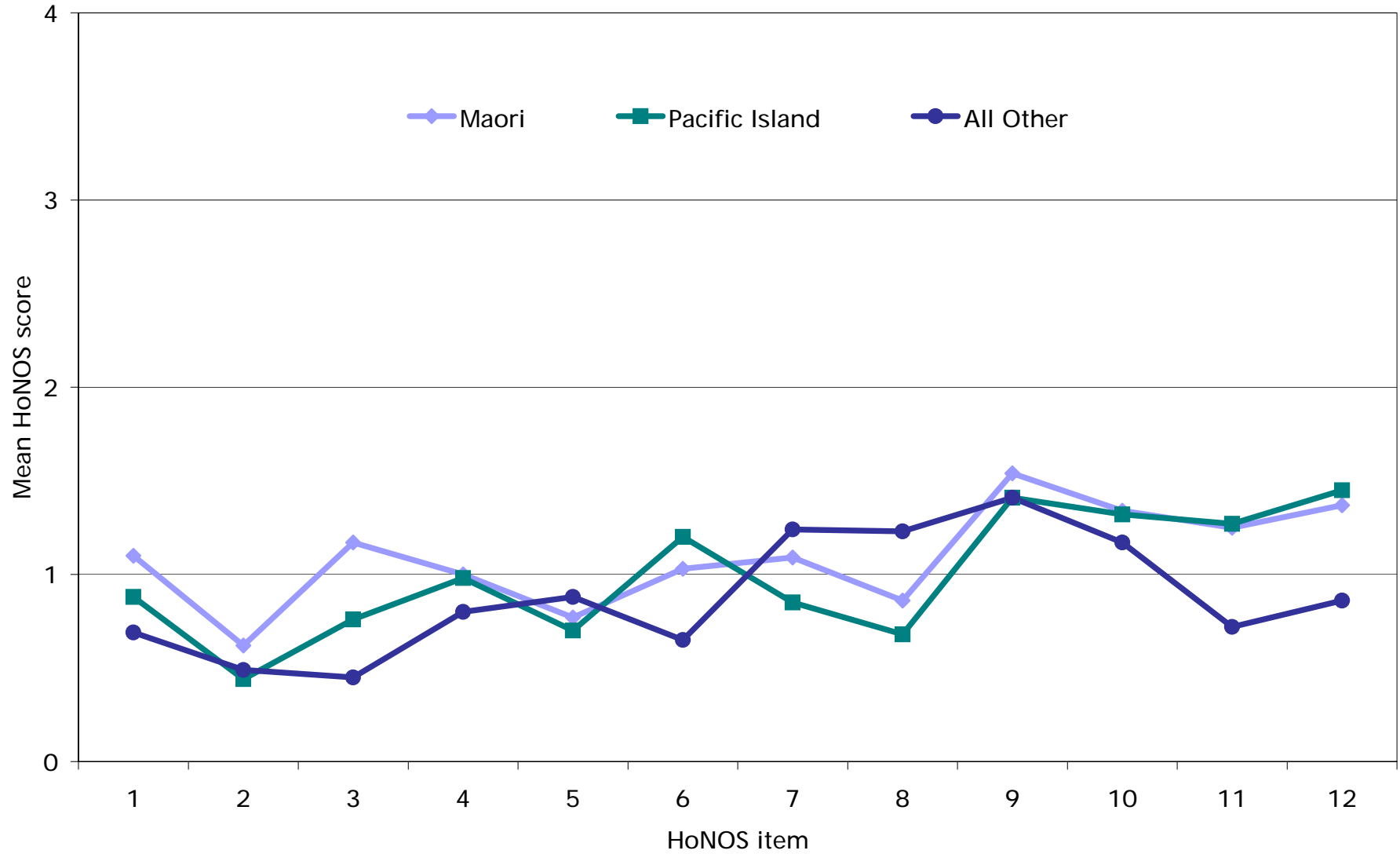
HoNOS, inpatient



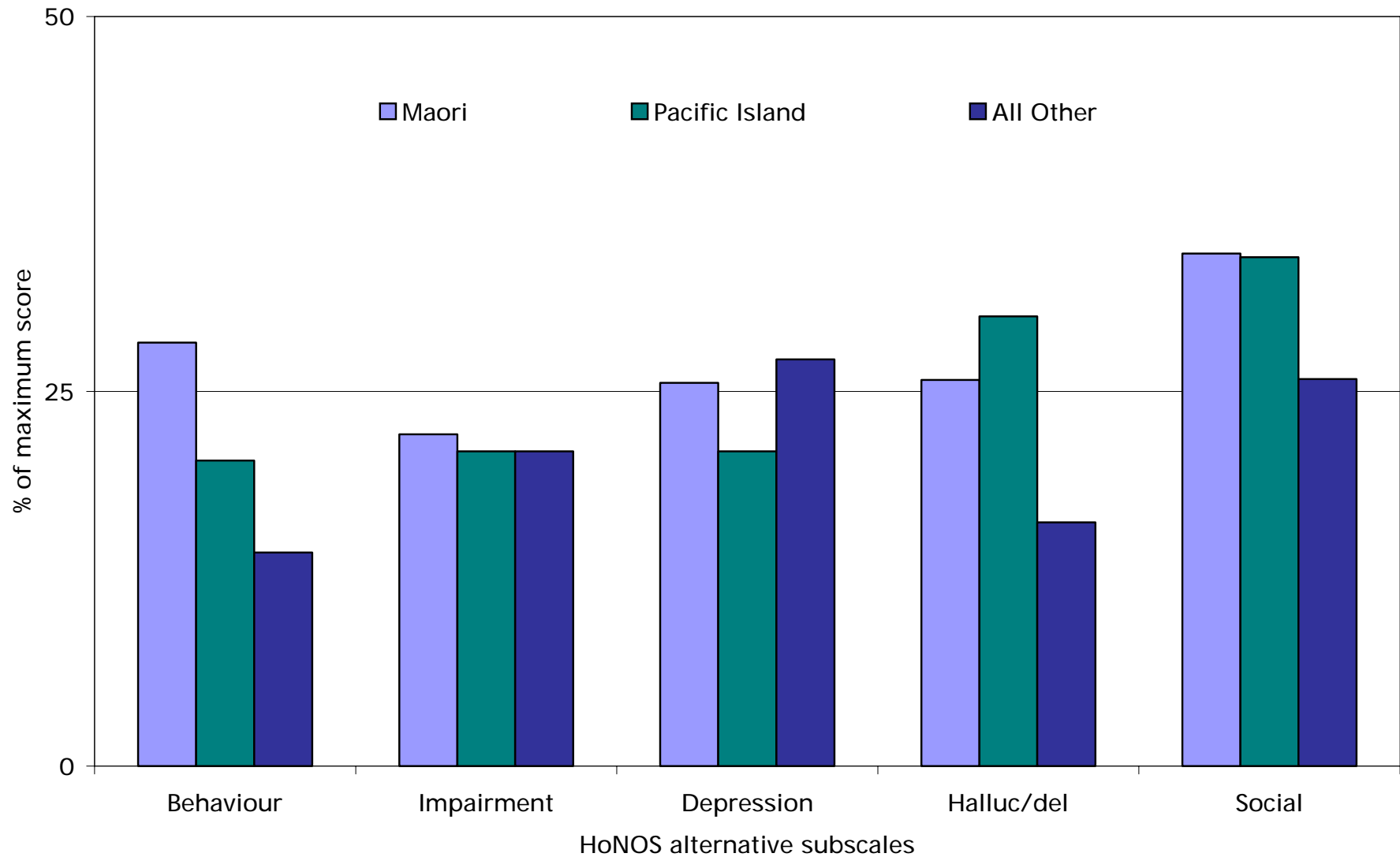
HoNOS subscales, inpatient



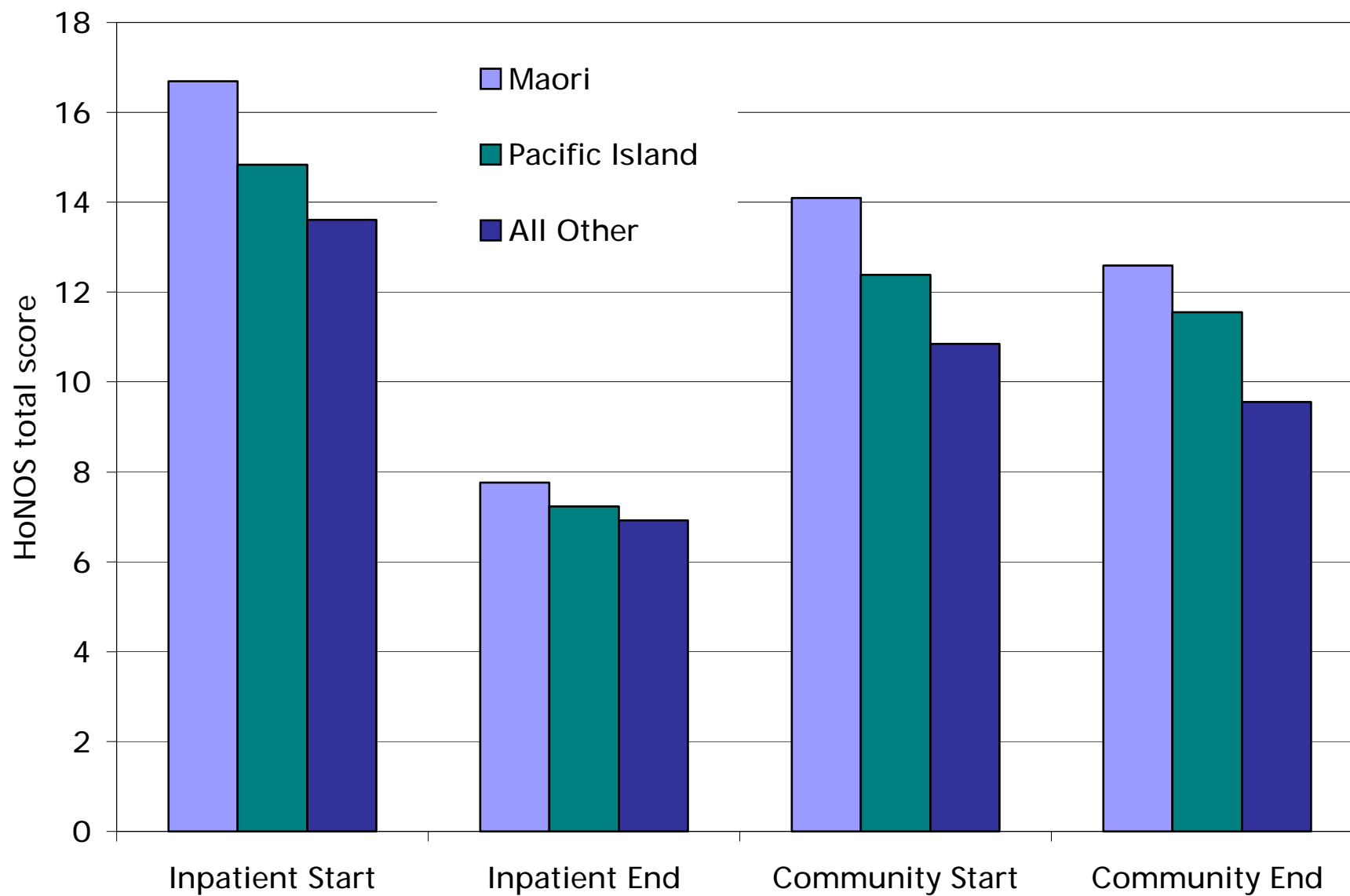
HoNOS, community



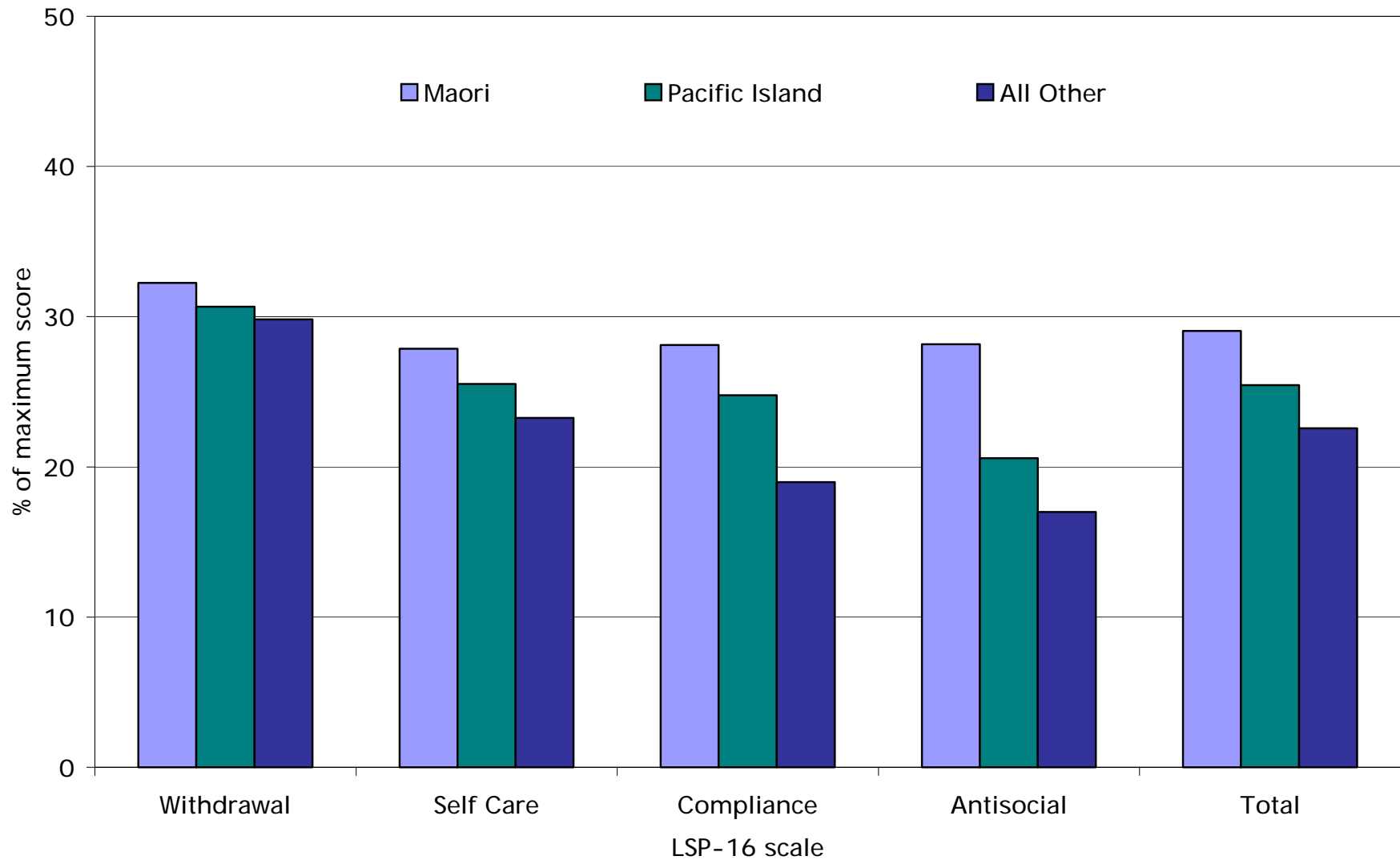
HoNOS subscales, community



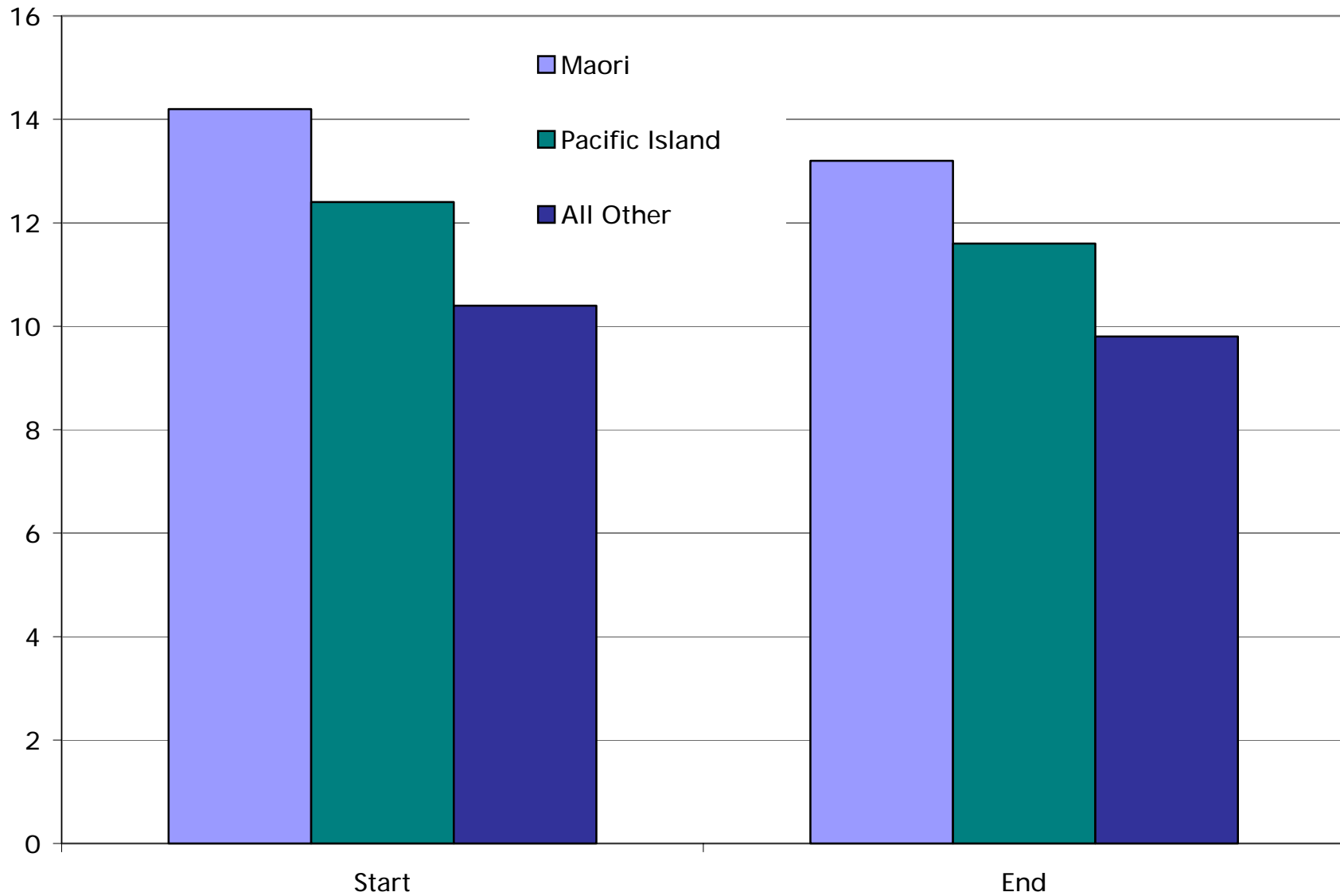
HoNOS start and end scores



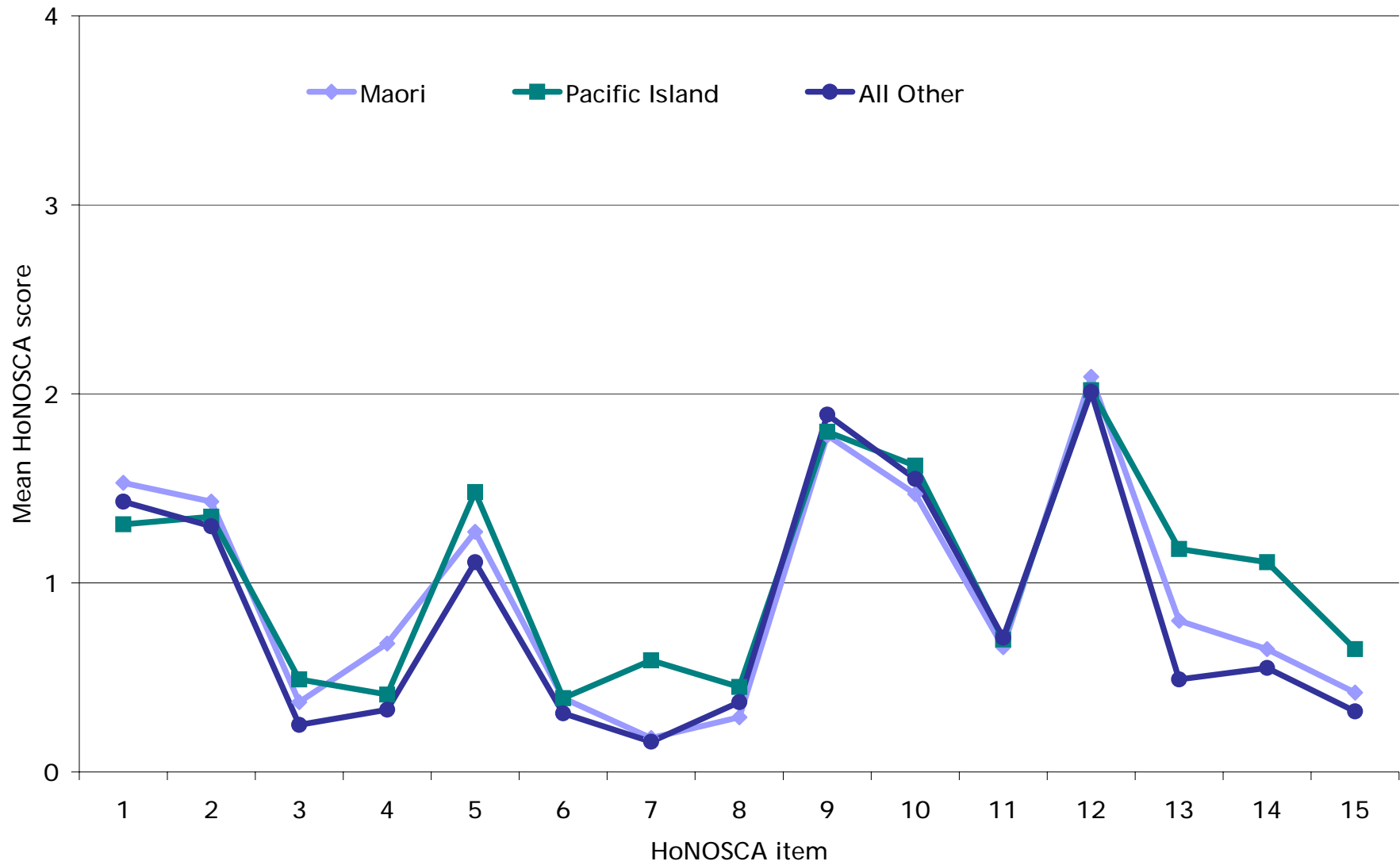
LSP-16, community



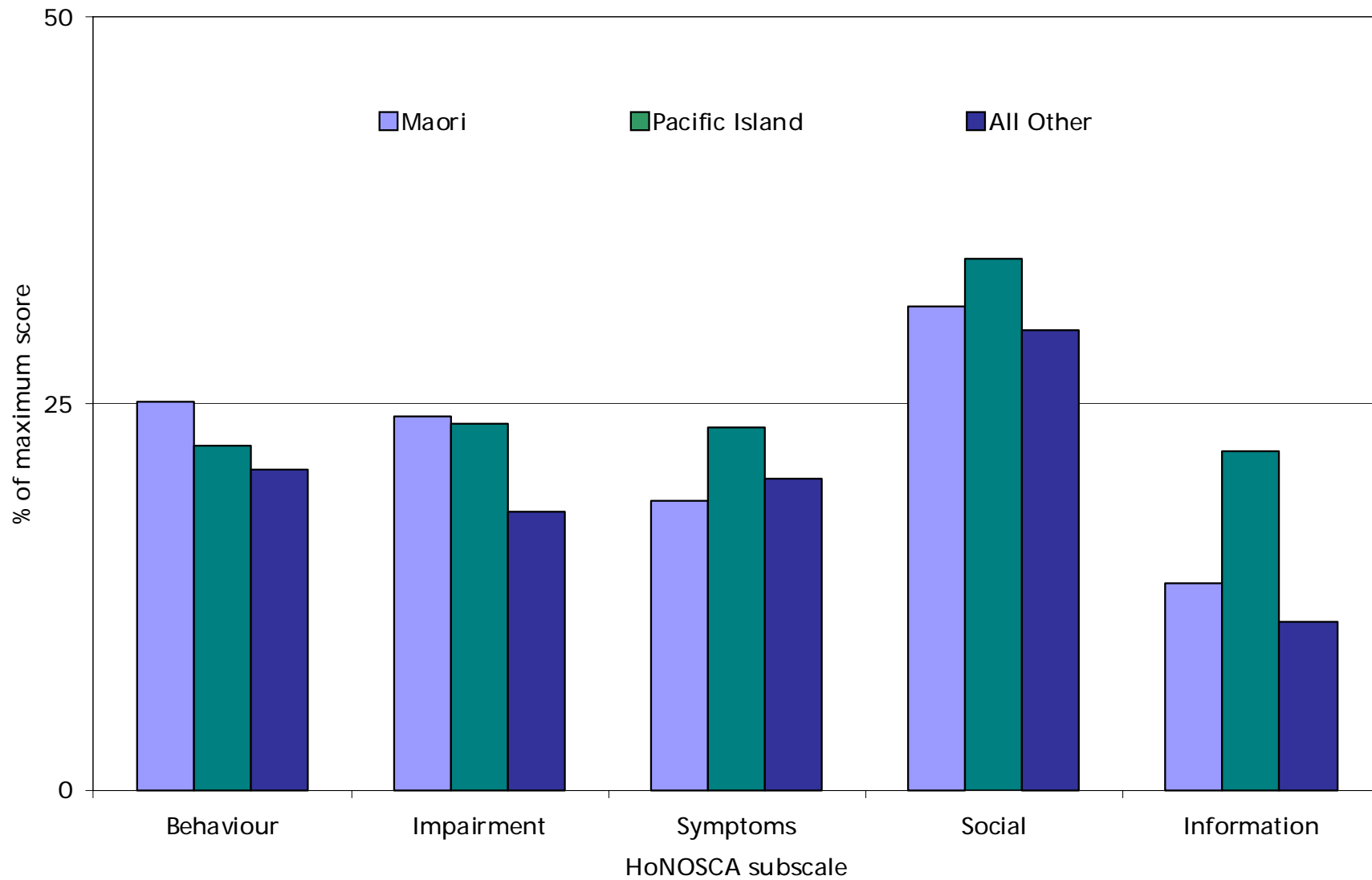
LSP-16 start and end scores



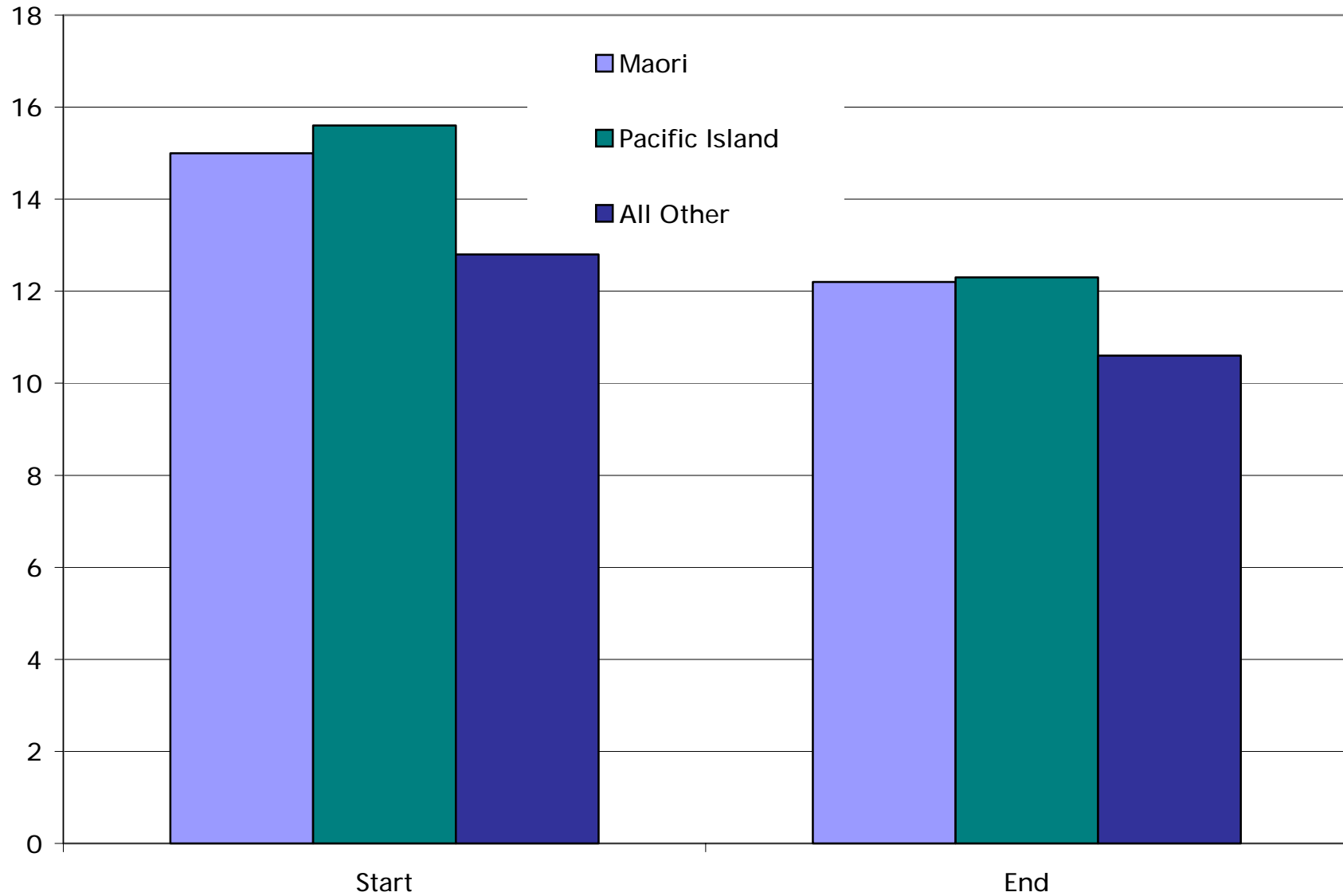
HoNOSCA, community



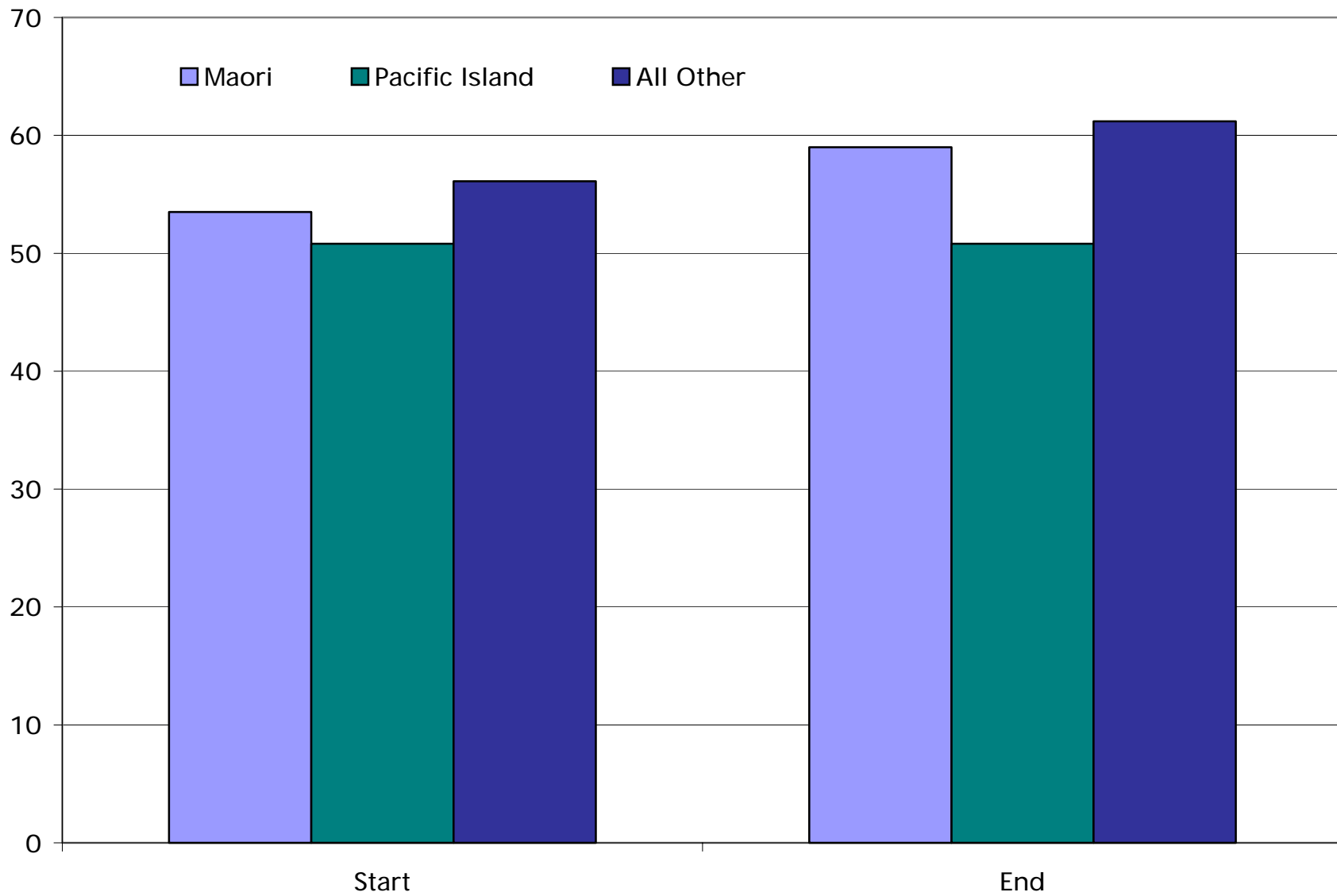
HoNOSCA sub-scales



HoNOSCA start and end scores



CGAS, start and end scores



Summary of differences - 1

- **Like the casemix analysis, differences between Maori, Pacific Island and other adult consumers.**
- **Unlike the casemix results, also differences between child and youth consumers.**
- **Overall, Maori & Pacific Island consumers are rated as having higher levels of symptom severity and lower levels of functioning than other consumers.**

Summary of differences - 2

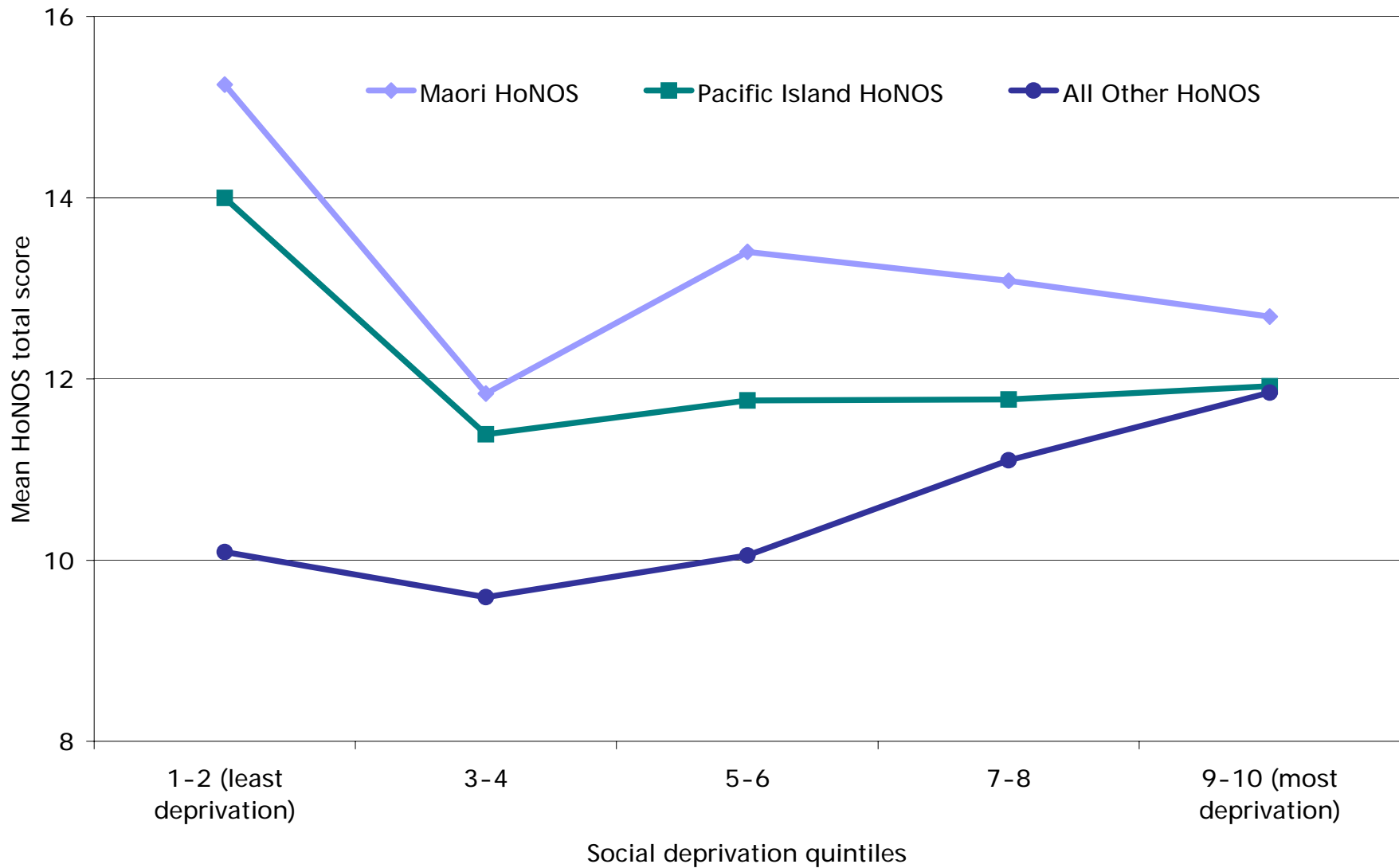
- **Overall, no ethnicity differences in the amount of change achieved during an episode of mental health care.**
- **Because any differences are due to differences in the ratings recorded at the start of an episode of care.**
- **Once these are taken into account, all three high-level ethnicity groupings - Maori, Pacific Island and All Other - achieve about the same level of improvement.**



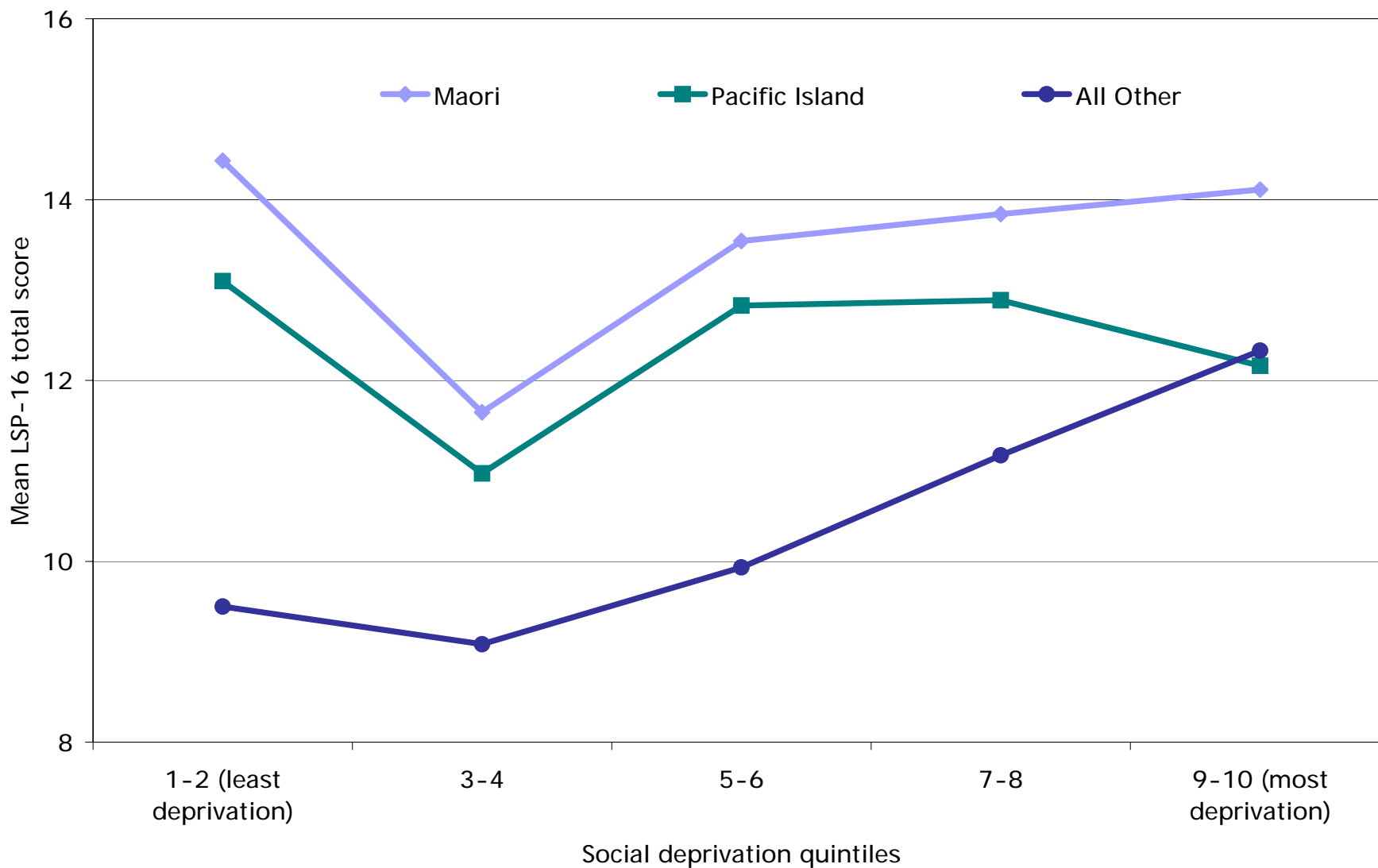
Outcomes, ethnicity and deprivation

- **Complicated!**
- **Meaningful?**
- **Not enough Pacific Island Child and Youth episodes, so analysis only on Adult episodes**

HoNOS total scores by deprivation quintile and ethnicity grouping



LSP-16 total scores by deprivation quintile and ethnicity grouping



Summary of findings

- **Overall, consumers living in areas of higher social deprivation have higher levels of severity and lower function than those living in areas of lower social deprivation.**
- **But reverse for Maori & Pacific Island.**
- **Ethnicity differences (both severity and functioning) are more marked when deprivation is least and less marked when deprivation is greatest.**

Possible reasons?

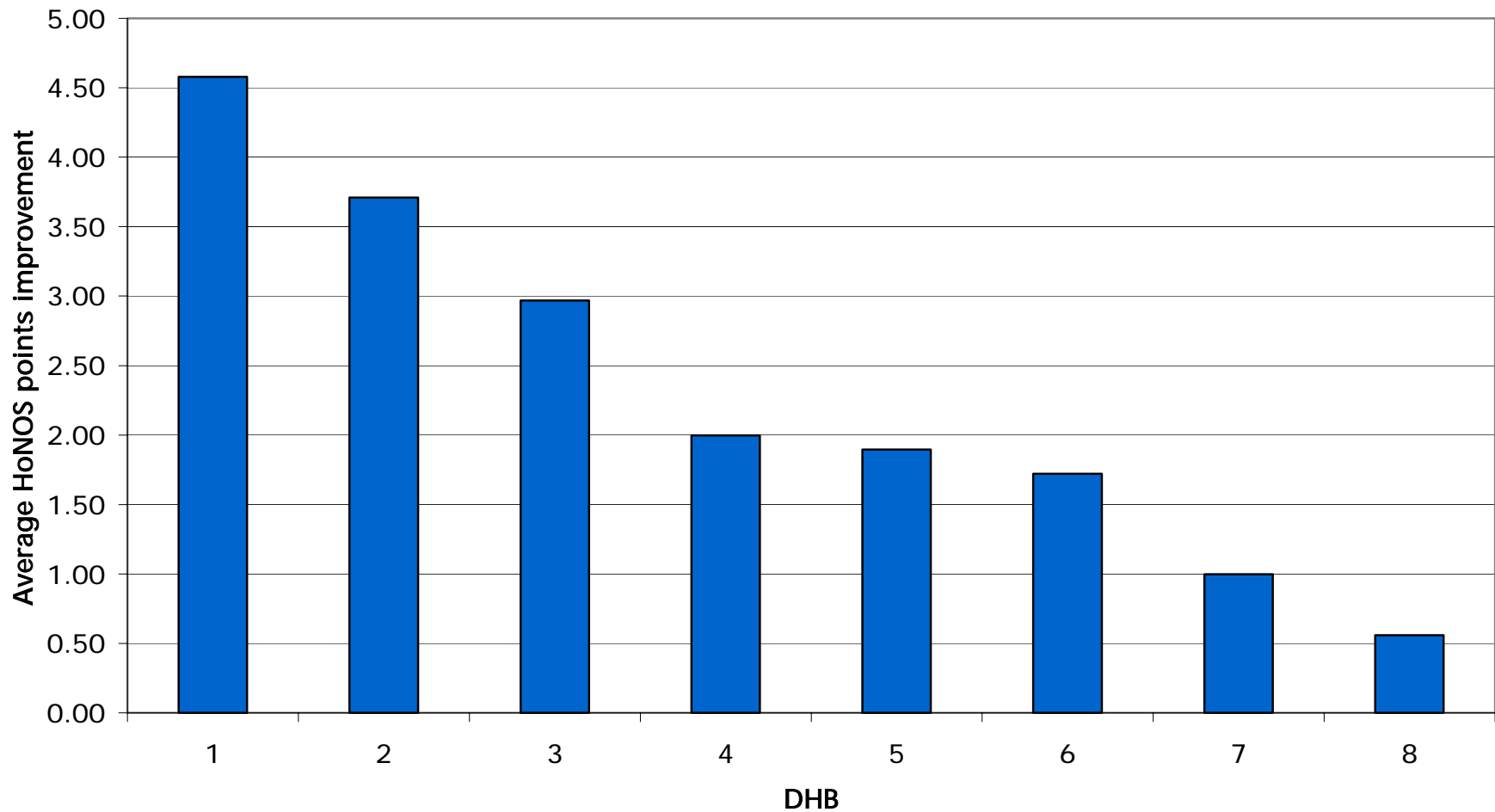
- **Ethnic/social class stereotypes**
- **Threshold for treatment differs by both ethnicity and area of residence.**
 - Differences in relation to when consumers get treatment
 - eg, Maori & Pacific Island consumers living in socially advantaged areas do not get into treatment until they are more unwell
- **A statistical aberration**
 - particularly because this is not about change, simply the mean score
- **Are these results meaningful & important?**



Casemix Adjusted Relative Mean Improvement (CARMI)

- A method to take account of the unique mix of consumers at each DHB

Average HoNOS improvement by DHB



Why the differences?

- **DHB 1 either:**
 - achieves the best outcomes or/and
 - has a mix of consumers that happen to be the most likely group to improve
- **Need to standardise for the consumer mix (ie, the casemix) to make the comparison meaningful**

Types of variation

- 1 variation due to differences in the ways that health services treat patients**
- 2 variation due to differences in the kinds of patients treated**

The challenge

- **We must be able to control for one type of cause of variation in order to understand the other**
- **If casemix classifications can be used to help to control for variations between patients...**
- **then we start to produce information which helps us to understand the differences between providers**

Controlling for differences in the mix of consumers between DHBs



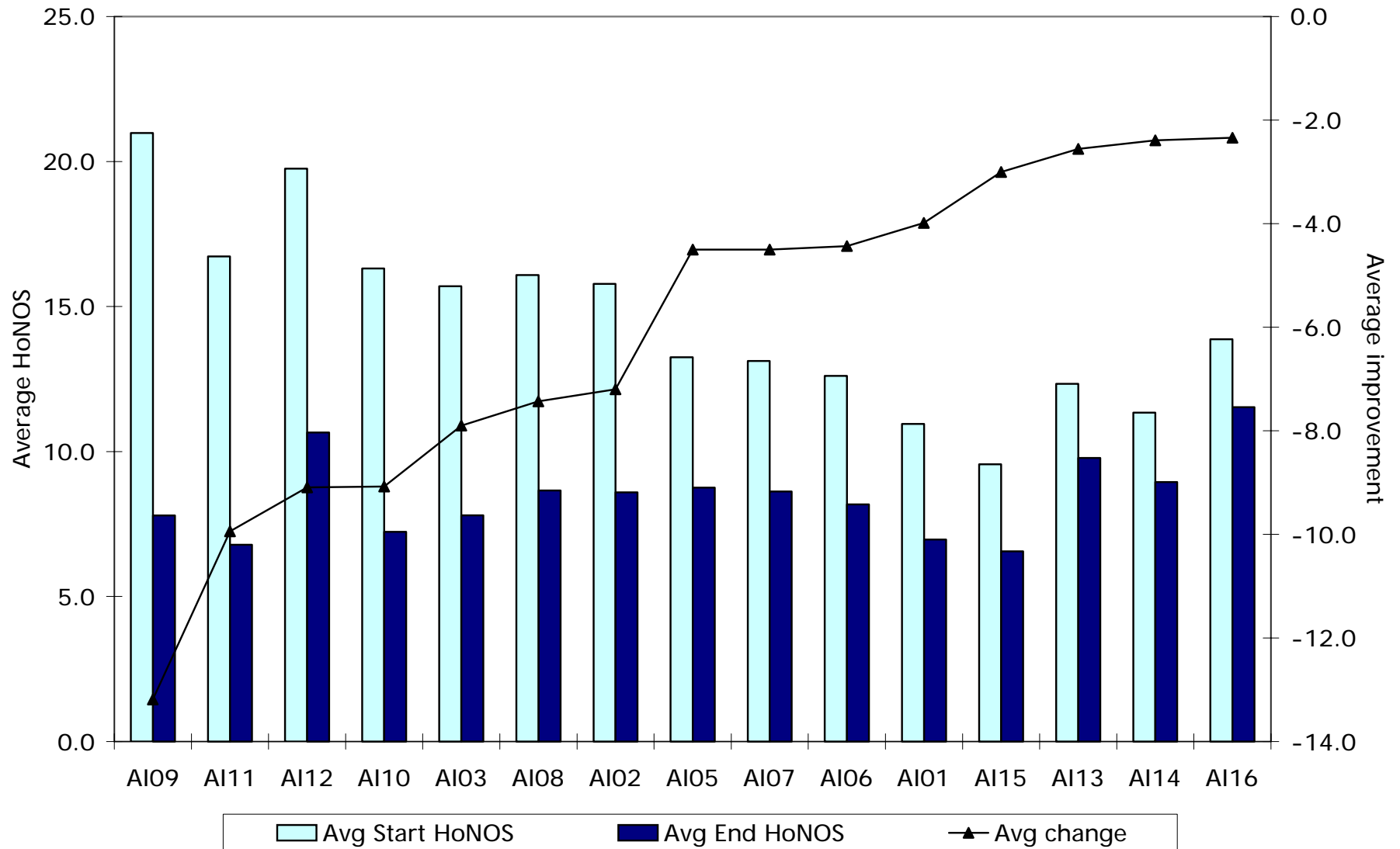
- **The casemix classification is the measurement tool**
- **Assigns episodes to a 'casemix class'**
 - **Similar consumers in the same class.**
 - **Different consumers in different classes.**
- **When outcomes results are standardised to take account the mix of consumers, any remaining differences can be attributed to differences between the DHBs.**
 - **Akin to 'risk-adjustment' by an insurer.**



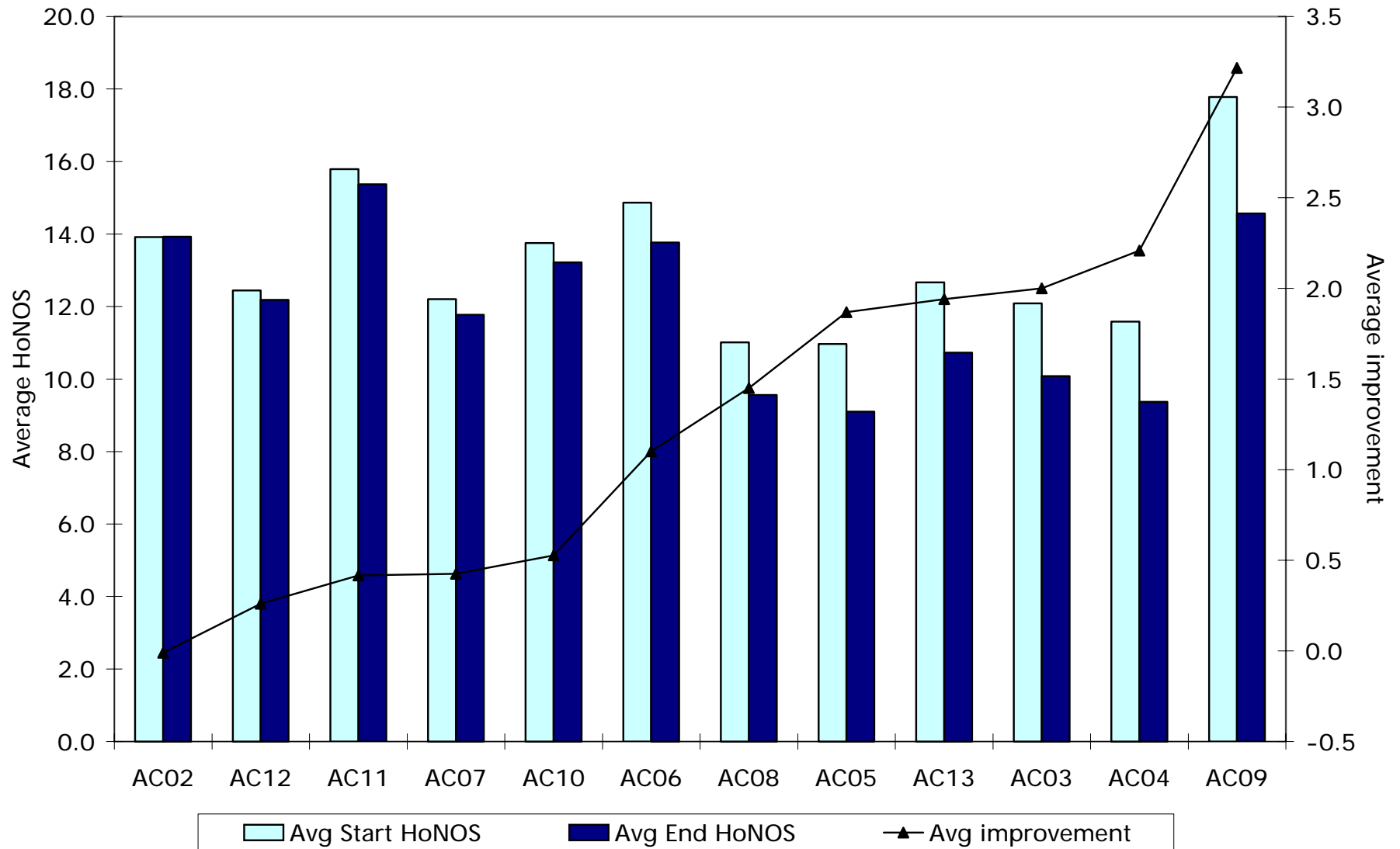
Variables used in the classification

- **Length of stay (Complete vs Ongoing inpatients)**
- **Age**
- **Ethnicity (adults)**
- **HoNOS start scores (adult inpatient)**
- **Diagnosis (child/youth inpatient)**
- **HoNOSCA start scores (child/youth)**
- **Legal status (adults)**
- **Focus of Care (adults)**

Improvement on the HoNOS by inpatient class



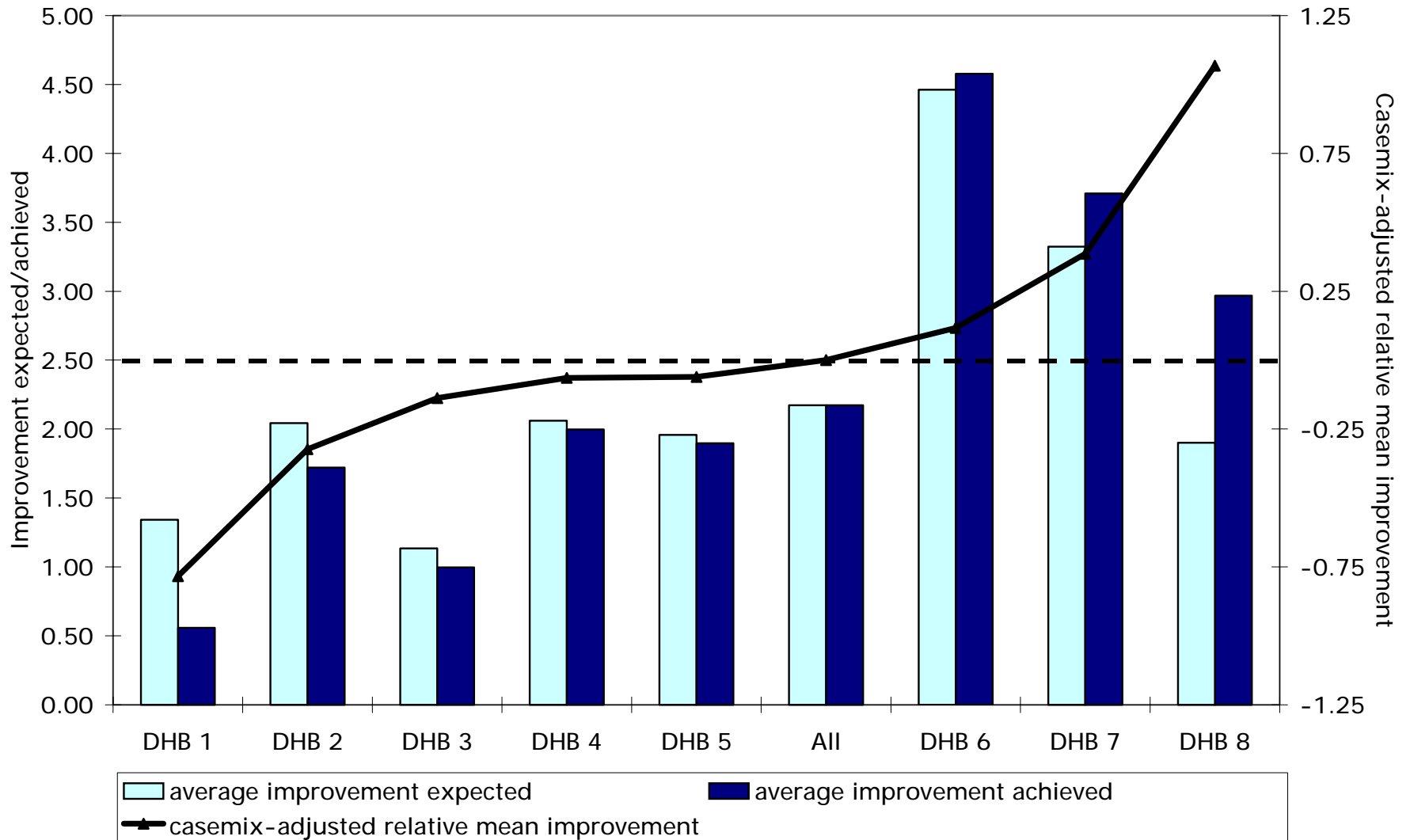
Improvement on the HoNOS by community class



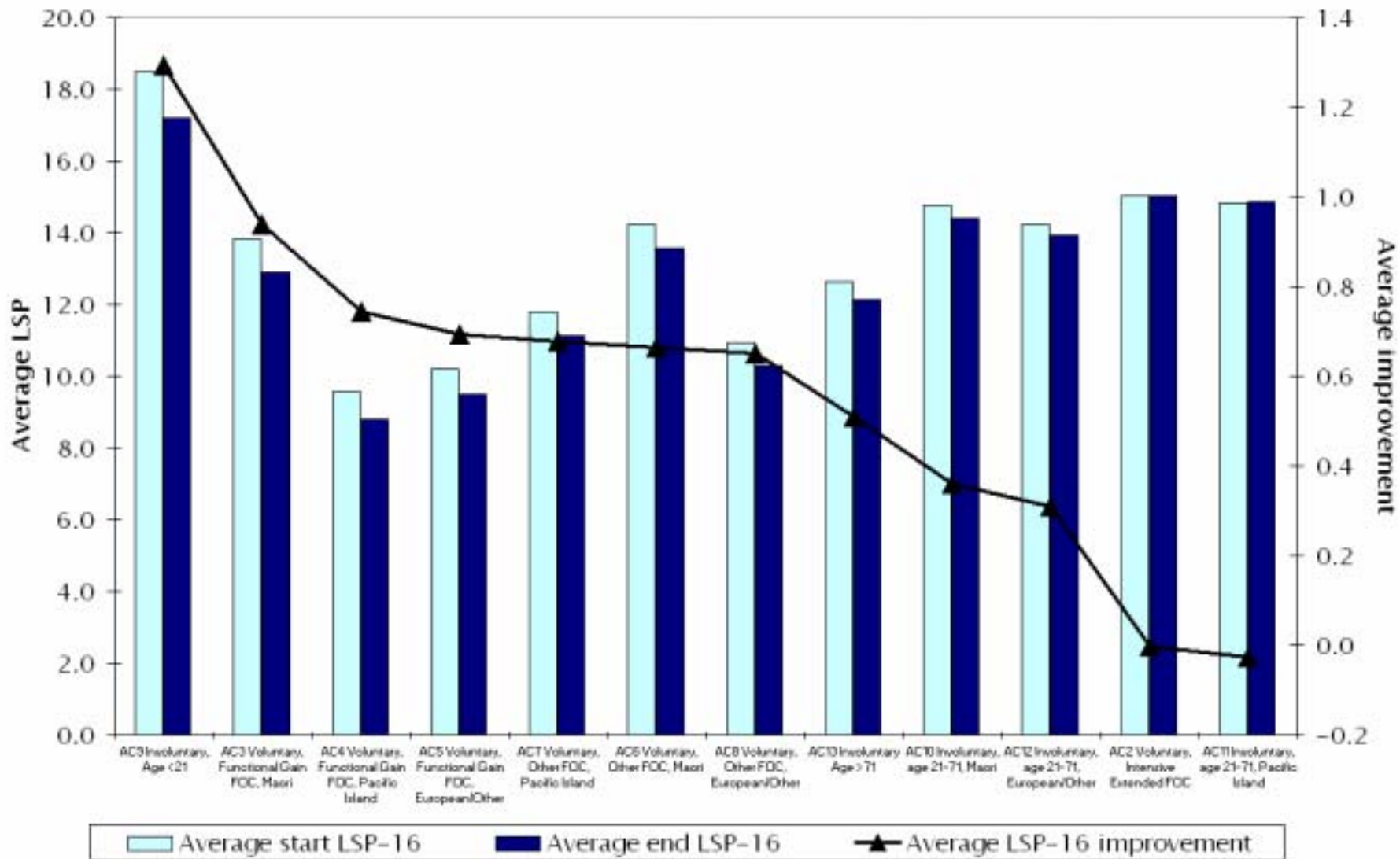
The CARMi Score

- **For each episode, calculate the difference between their change and the average change for their casemix class.**
- **Average across the DHB.**
- **CARMi of 1 point**
 - on average, consumers improved by 1 point more than expected.
- **Negative CARMi**
 - consumers achieved less improvement than similar consumers in the study.

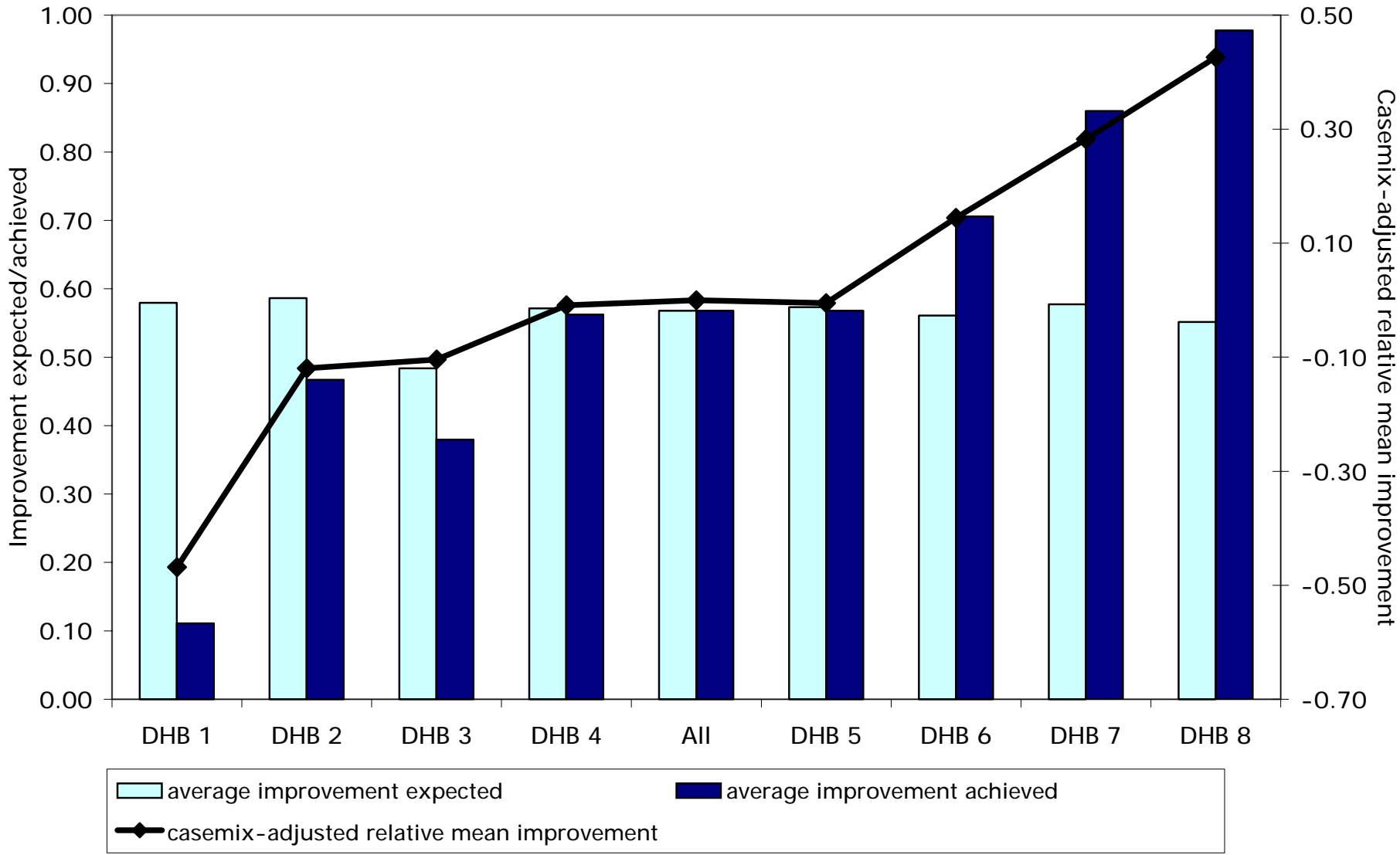
HoNOS CARMi



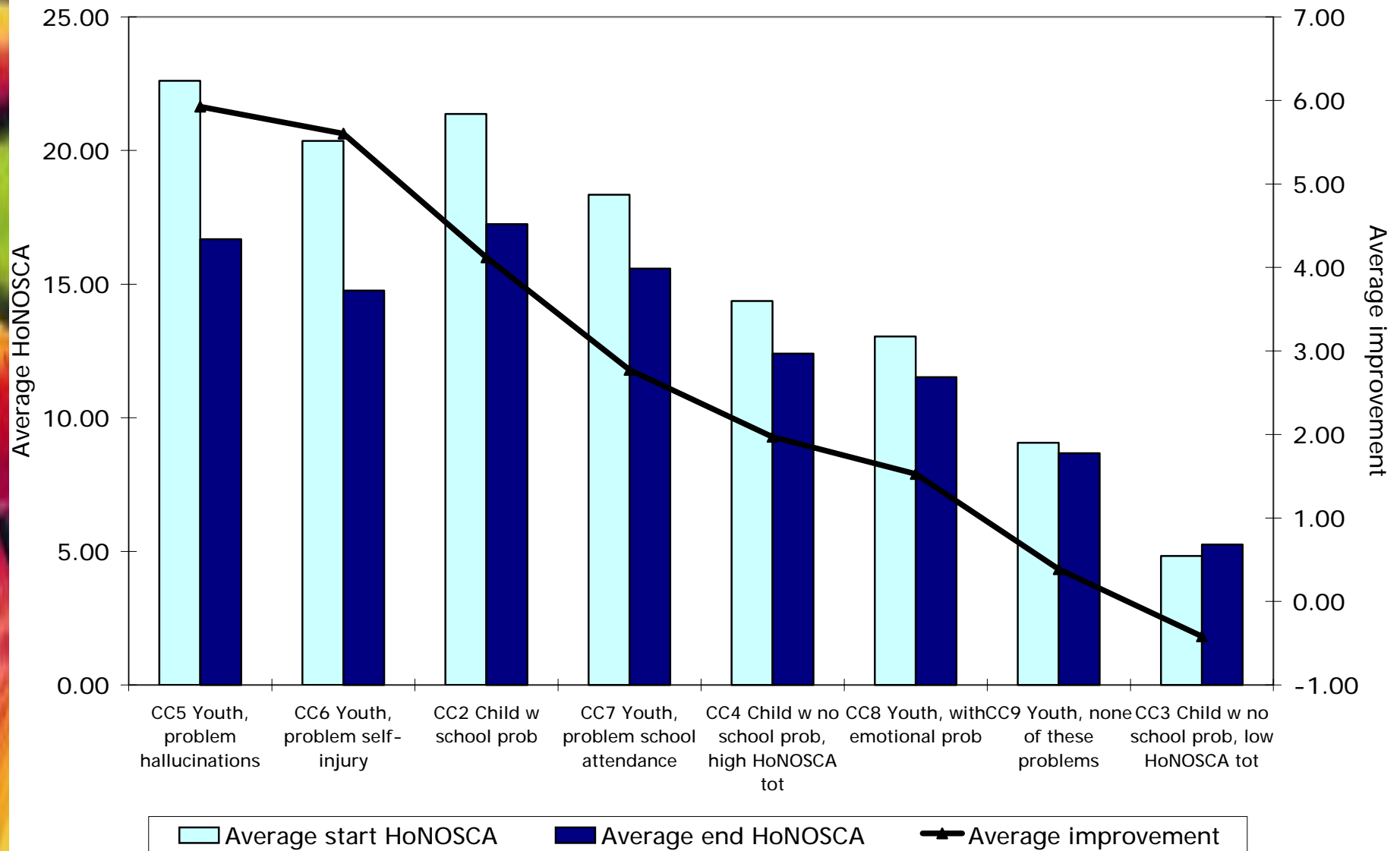
Improvement on the LSP-16 by community class



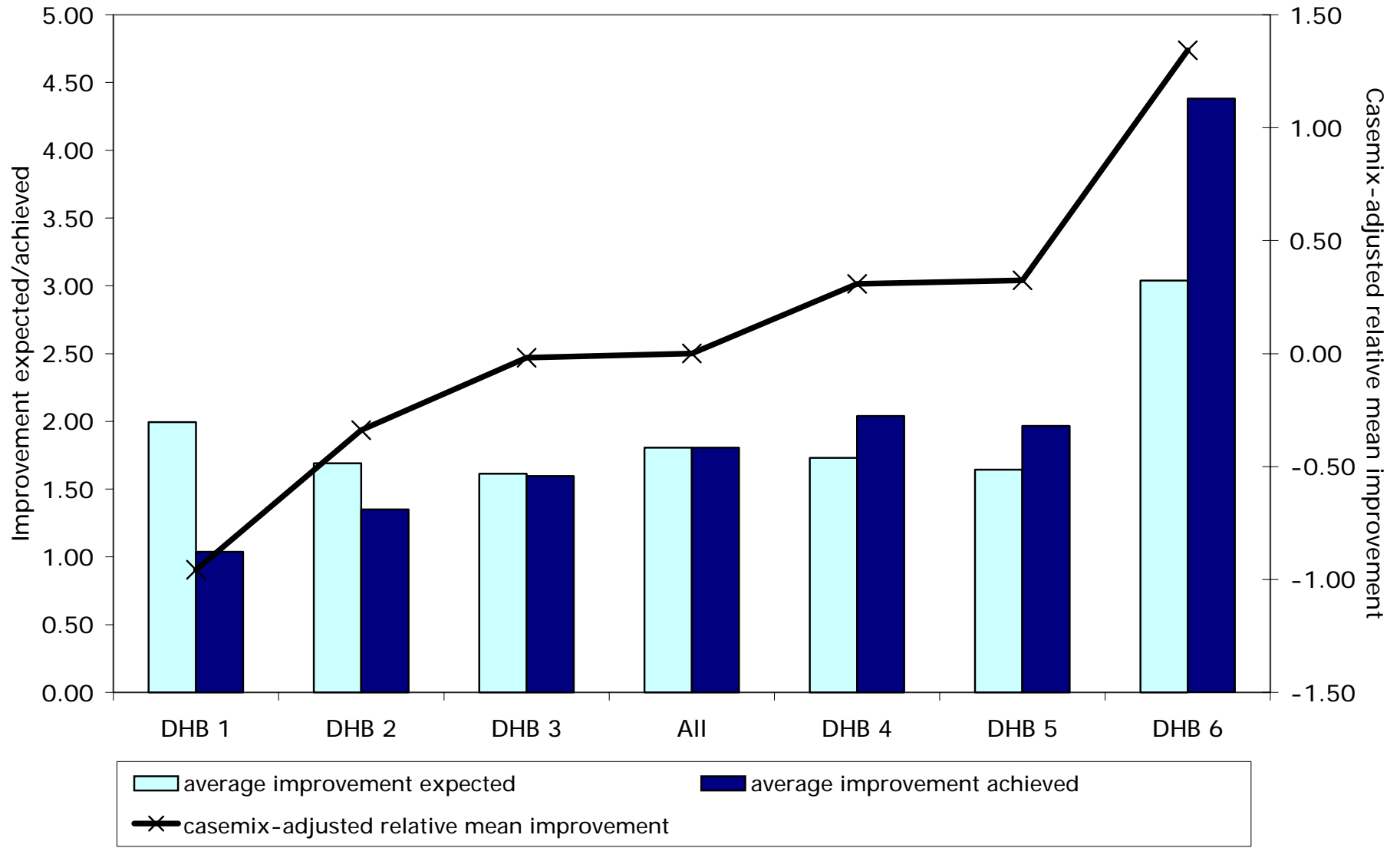
LSP-16 CARMi



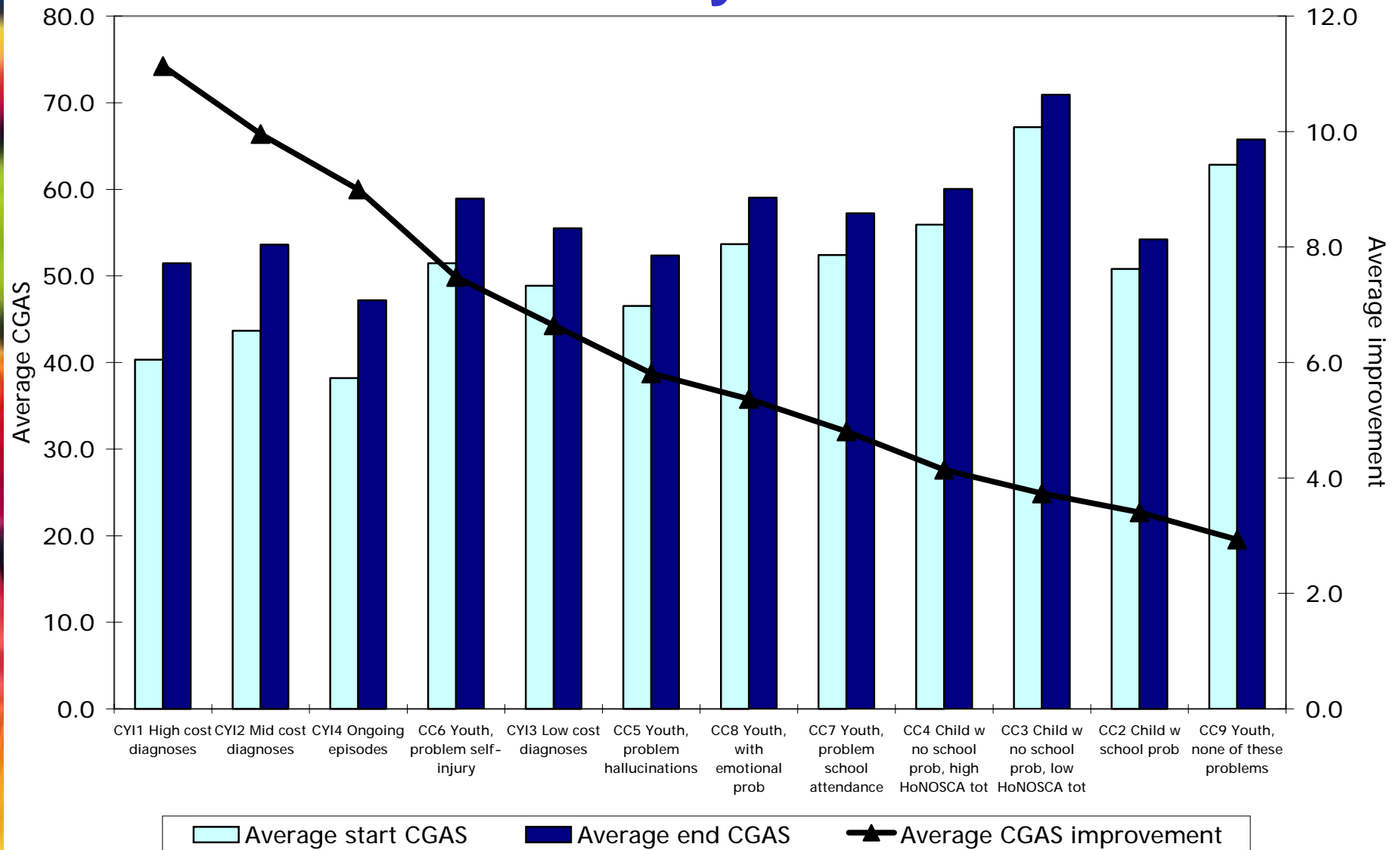
Improvement on the HoNOSCA by community class



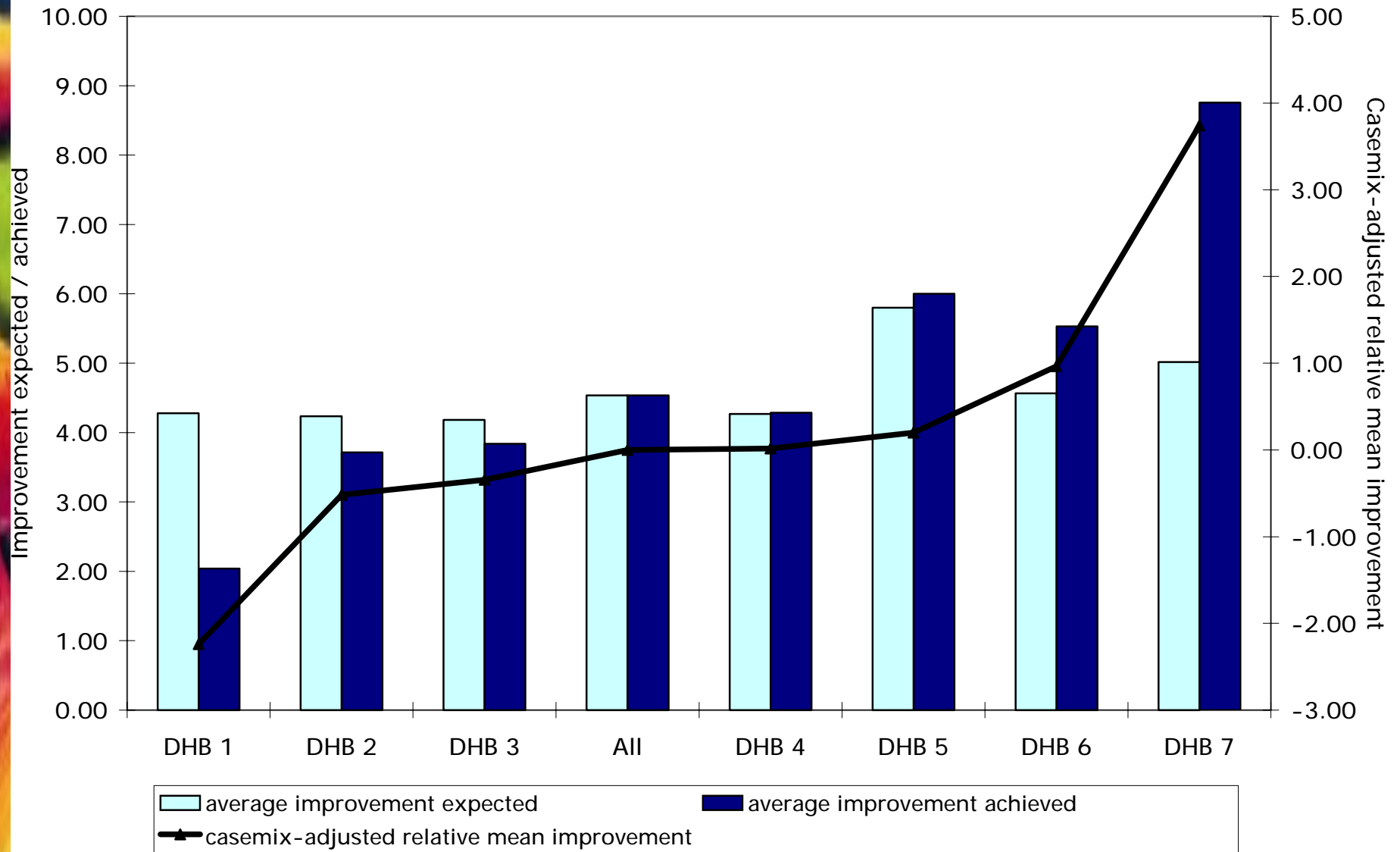
HoNOSCA CARMi



Improvement on the CGAS by community class



CGAS CARMi





Painting by Serena Young



When outcomes meet casemix

- Drawing the 2 studies together

Ethnicity results (from the casemix analysis)



Ethnicity Grouping	Cost per episode (raw)	Average case complexity	CW=1	Diff.	% diff.
Maori	\$7,032	1.49	\$4,719	\$39	0.8%
Pacific Island	\$9,235	1.93	\$4,792	\$111	2.4%
Euro	\$3,776	0.81	\$4,662	-\$19	-0.4%
All	\$4,681	1.00	\$4,681	\$0	0.0%

Diff. = difference from NZ casemix-adjusted average (the average cost per episode that is unexplained by the classification)

Setting-specific cost weights

Ethnicity Grouping	Average weight – inpatient-only weights	Average weight – community-only weights
Pacific Island	1.35	1.44
Maori	1.22	1.05
European/ Other	0.86	0.96
All	1.00	1.00

- **The casemix analysis found:**
 - **significant differences in cost between Maori, Pacific Island and other episodes**
 - **Relative to the national study average:**
 - **Maori - 49% more costly**
 - **Pacific Island - 93% more costly**
 - **Other - 19% less costly**
- **The outcomes analysis found:**
 - **after adjusting for different starting points, no differences in the amount of change recorded**



Painting by Serena Young



Conclusions and implications

- What does it all mean?



Conclusions - 1

- **Outcome measurement in routine clinical practice is possible.**
- **But training and support is required.**

Conclusions - 2

- **The measures are OK technically**
- **An alternate measure of function for adults and older people is required**
 - **consider both mainstream and mental health specific measures of function**
- **Work needed on the Focus Of Care**
 - **refine definitions – Acute, Functional Gain, Intensive Extended and Maintenance**
- **Complementary consumer-rated and family/carer-rated measures are needed**

Conclusions - 3

- **The outcomes achieved by different mental health services need to be casemix-adjusted to take into account the unique mix of consumers at each DHB.**
 - Crude numbers are misleading
 - The CARMI is useful

Conclusions - 5

- **It costs different amounts to achieve the same outcomes**
 - **Maori - 49% more**
 - **Pacific Island - 93% more**
 - **Other - 19% less**

Conclusions - 6

- **Cultural competency remains a challenge**
 - requires ongoing attention as part of the workforce outcomes development and training program.
- **Further work required to understand the reasons for ethnicity differences and the relationship between ethnicity, deprivation and consumer outcomes.**
 - National epidemiological study, further research and consultation with consumers, carers and clinicians.

Conclusions - 7

- **Routine outcome measurement is:**
 - **necessary, but insufficient, in developing and improving NZ mental health services**
 - **an important component of a bigger strategy**