Healthy London Partnership

HoNOS analytical framework for mental health services

Table of Contents

1. What does HoNOS measure?

HoNOS[1] (Health of the Nation Outcome Scales) is a method of measuring the health and social functioning of people with severe mental illness. It is a Clinician Reported Outcome Measure (CROM). It is comprised of 12 scales (see Appendix 1) measuring behaviour, impairment, symptoms and social functioning[2], which were developed using stringent testing for acceptability, usability, sensitivity, reliability and validity. Each scale is rated based on severity (on a five point Likert scale):

Level	Severity description	Clinically significant?	Incorporated in patient care plan?
0	no problem	No	
1	minor problem requiring no action	No	
2	mild problem but definitely present	Yes	Maybe
3	moderately severe problem	Yes	Always
4	severe to very severe problem	Yes	Always

[1] Copyright in the Health of the Nation Outcome Scales (HoNOS) is owned by the Royal College of Psychiatrists. Commercial copying, renting, and adaptation are prohibited © Royal College of Psychiatrists 1996

[2] Wing, J. K., Curtis, R. H. & Beevor, A. S. (1996) *HoNOS. Health of the Nation Outcome Scales. Report on research and development.*

The scales are completed after routine clinical assessments in any setting and form part of the Mental Health Services Data Set. For each scale there is further guidance and context notes to aid scoring. These can be found here. HoNOS allows the interpretation of scores on every single item of interest. As far as possible, results should always be presented for all 12 scales/ questions.

2. Why should we analyse HoNOS?

2.1 At a patient level

Using HoNOS as a clinical rated outcome measure allows clinicians to build a picture of service users' needs across a range of areas, incorporating broad domains of mental health, safety, physical health, relationships, housing and functioning. Depending on the time point when the HoNOS rating is gathered, it offers an estimate of patient needs and changes to these over time.

Measuring change (particularly in specific domains) can offer proxy measure of responses to interventions / treatments. Measuring change in scores can potentially help infer the impact of clinical interventions and can be discussed within supervision sessions.

For HoNOS data to be used as a CROM it should be gathered at critical time points in the patient journey, and the rating needs to be reliable.

2.2 At a team and service level

At a team or service level pooled scores at entry can offer mapping of needs of the local population at referral or at baseline. A comparison between entry (Admission) and exit (Discharge) level scores offers an opportunity to measure change in patient need over time. A pooled reduction of need potentially indicates tendency towards recovery. This is particularly important for brief or acute interventions.

Interim scores or review of HoNOS scores can also offer potentially useful information. There might be evidence of improvement and recovery.

Understanding change in clinical outcome also requires understanding of the interventions applied and the context in which these occur (for example initial severity, gender, ethnicity, length of stay etc).

When reviewing team outcomes, it is important data collected and analysed is fed back to the team who are collecting the data. This serves the purpose of clinical engagement as well as stimulating improvement of data quality. Demonstrating improvement, reduction of patient need or a desired outcome can potentially improve team morale.

Training of staff is critical to ensure reliability of data. One needs to be mindful about potential sources of bias before embarking on inter-team comparison. However, when data quality is assured, different teams and interventions serving similar populations and offering similar interventions can be potentially compared. It would offer intelligence to inform service design using quality improvement (QI) methodologies. It provides the opportunity to learn from best practice, with the ultimate aim of improving patient outcomes.

It is worth noting that comparisons between teams within the same service (e.g. home treatment teams) are not always possible due to differences in demography, operational focus or diagnostic groups. Other factors to take into consideration would be staffing levels, case-mix, vacancies and locum use as well as access to interventions (psychological, occupational therapy (OT) etc).

It does however give the clinician an indication of teams that are, on balance, more effective at treating certain diagnoses / symptoms, and to start asking questions about why that service is more effective. Questions clinicians could ask about are about complexity of patients seen, numbers of patients seen, staffing to patient ratios, permanent vs locum staff, case mix of the staffing etc.

While assessing results for clinical significance and statistical significance it is important to note statistical significance is only relevant if the results are being used to extrapolate outcome beyond the team (or patient population) where the data is collected. Within the team whatever change is noted is true for the population on which the data was collected from.

2.3 At a Trust level

Many of the challenges of aggregating data and thereafter comparison between different pooled data sources are already covered in the service level pooled data section above. These considerations remain at Trust level data. However, Trust or provider level data offers larger numbers, i.e. bigger data sets. These can be used to assess local population level needs and outcomes, and whether there is variation in outcomes for different patient groups (e.g. diagnostic groups or different ethnic groups with same diagnosis).

Trust level data is also useful to assess organisational engagement with CROM or HoNOS gathering, timeliness and data quality and completeness.

Extrapolating large data sets to population level hypothesis would involve appropriate statistical modelling around these (see section 7).

3. When should ratings be obtained?

For each patient and treatment episode, there are entry and exit ratings, and for longer term treatments there are review ratings in between. In principle, ratings should be obtained at the beginning of any new treatment episode (e.g. admission to in-patient treatment, home treatment team (HTT) or other community service, referral to new provider organisation) and at the end of that treatment episode (e.g. discharge from that service or provider organisation). For acute treatment, the expectation is that initial ratings are obtained at the latest 48 hours after admission, this includes admission to the ward or to the HTT.

For treatments over longer periods of time, including within community based teams, review ratings are required (ideally at fixed time intervals). The fixed time intervals for acute treatment (in-patient and crisis teams) can be shorter (e.g. every four weeks), for on-going treatment in community services and out-patient clinics they should not be longer than every six months. If patients' needs significantly change whilst receiving treatment it might also be appropriate to re-rate HoNOS outside these timescales.

4. Importance of data quality and completion rates

Paired data is critical to measure change. For data to be meaningful, HoNOS scores need to be collected in completeness (as far as practicable) and at critical time points (Section 3).

In order to be more confident about data quality it is essential for clinicians who fill in HoNOS to be trained about how to complete HoNOS and the clinical utility of HoNOS via a Trust internal training programme or using the Healthy London Partnership HoNOS training programme.

4.1 Open and closed cases

Often with analysis of HoNOS there is the option to review cases that are open to the team, discharged (closed) or both open and closed.

It is important to keep in mind the differences between the patient groups for open and closed which will affect the analysis and clinical interpretation of the HoNOS (clinical) outcome scores.

There is a presumption that cases that are closed or discharged to the team have improved or got better and so would be different to those who are open, who are likely to be either early in their recovery journey, long term patients or patients who are more complex and have greater needs. Patients who are long term (chronic) with more complex needs are likely to show smaller improvements or may not show any improvements but show no deterioration, which in itself shows effective treatment in much the same way as patients with chronic complex diabetes will not show vast improvements in their clinical outcomes.

4.2 Missing scores

Mean scores can be calculated even when one item (of the 12 scales) is missing (that item is then ignored when calculating the mean). Yet, when more than one item is missing, mean scores may be substantially affected.

Unrepresentative and incomplete data prevents comparisons and limits the utility of feedback. Trusts are therefore encouraged to improve the data quality and collection of HoNOS paired scores, in order to enable more meaningful analysis and interpretation of the data.

4.3 Paired data

Paired data is when two sets of HoNOS scales have been recorded for a patient – for example at admission (T1) and discharge from team, or at six monthly intervals (T2). Paired data may represent admission to discharge, admission to review or review to discharge and (rarely) review to review.

When measuring team level data to assess efficacy or effectiveness of interventions offered, higher levels of completeness (proportion of patients who have two-point data reported) is warranted. Smaller datasets involving smaller proportions of the patient population are prone to greater *bias* and the results might not be truly representative of the team's entire work. However, at the first instance we would like to encourage regular data gathering and internal analysis of such data.

The greater the percentage of patients who have paired outcomes recorded, the more representative the data and analysis will be. We propose an aspiration of recording HoNOS pairs for 70% of a team's treatment episodes. Meaningful analysis is limited where paired scores represent less than 30% of activity.

For this reason, publication of outcomes data should specify the percentage of all closed episodes that are represented by the sample with HoNOS pairs.

5. Interpreting the data

Changes are the differences between ratings at two time points (e.g. admission and discharge) and can be displayed graphically through a variety of methods. In November 2018 it was agreed by the London Mental Health in Integrated Care Systems Technical and Clinical Group that the two preferred methods to use in London are HoNOS profiles and categorical change (described in more detail below).

We recommend that Trusts use both methods in order to promote understanding, discussion and reflective clinical practice.

Consideration of service type and case-mix data are important to contextualise HoNOS data, including age, gender, ethnicity, episode duration, diagnosis and Cluster (understanding clinical severity). We recommend the HoNOS scores are linked to the service inputting the data wherever possible and the data is linked with patient diagnosis and / or Cluster (if ICD 10 or SNOMED data is unreliable).

Depending on the condition being treated maintenance of needs (or scores), i.e. a lack of worsening may also be a desired outcome e.g. in early intervention treatment or neurodegenerative conditions such as dementia.

Initial severity is a critical predictor of outcome. The greater the initial severity, the greater the possibility of a significant reduction in severity following treatment. Therefore, greater reduction of needs may be observed across an acute episode compared to an episode of longer term community treatment.

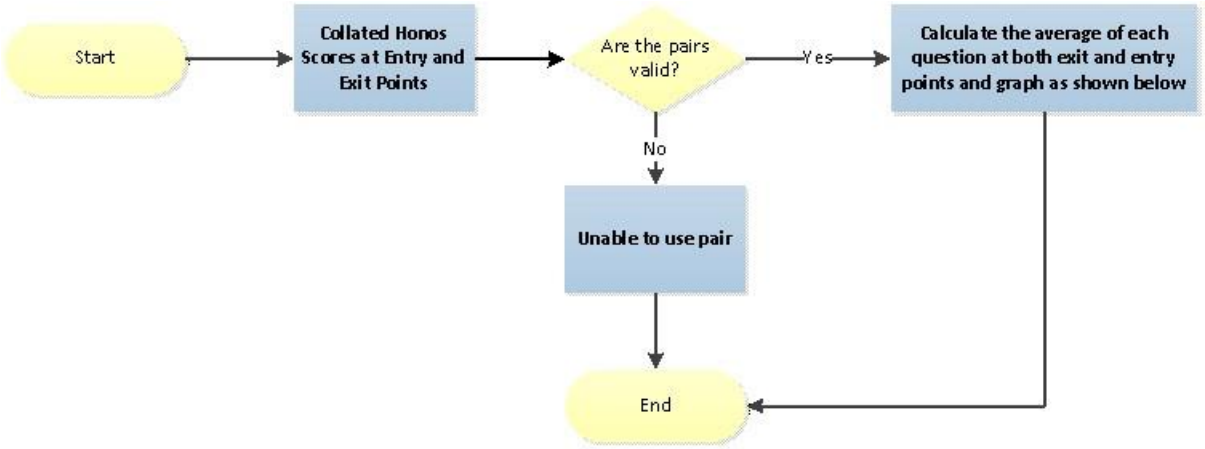
6. HoNOS profiles

HoNOS profiles measure change on each of the 12 HoNOS scales by comparing mean scores at the start and completion of treatment episodes, at patient and team / service level.

This approach is visually easily understood by clinicians, who can visualise where severity is greatest, and on which scales the greatest changes have occurred.

The figure below illustrates the process for calculating the HoNOS profiles.

Figure 1: Process for calculating HoNOS profiles, Central and North West London NHS Foundation Trust



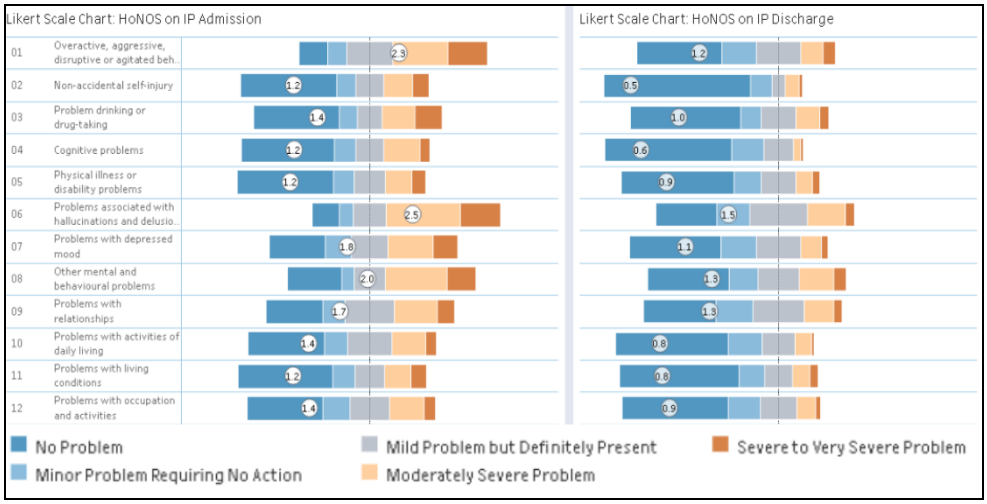
It is worth noting that within some services, there may be little change in the patient profile of the service – more complex or long term patients may show limited or no change in outcomes, which may dishearten clinicians. However, this is where a clinical narrative is vital as clinicians would be able to explain that within their service, for example, prevention of deterioration of symptoms/patient functioning is providing high quality care. Example HoNOS profiles for different mental health services are provided on the Healthy London Partnership HoNOS training programme website.

The data can be illustrated in a number of ways, for example using Likert scale charts or line charts (see examples below).

This approach can be supplemented with effect size statistics, to identify where the greatest changes in severity occurred and to estimate the magnitude of those changes (see Section 7).

Figure 2 below shows a HoNOS Likert scale chart, which illustrates the difference at admission and discharge, showing the spread of scores for each scale (0 being 'No problem' and 4 being 'Severe to very severe problem').

Figure 2: Example HoNOS Likert scale chart, Central and North West London NHS Foundation Trust



6.1 HoNOS profiles – minimum initial score

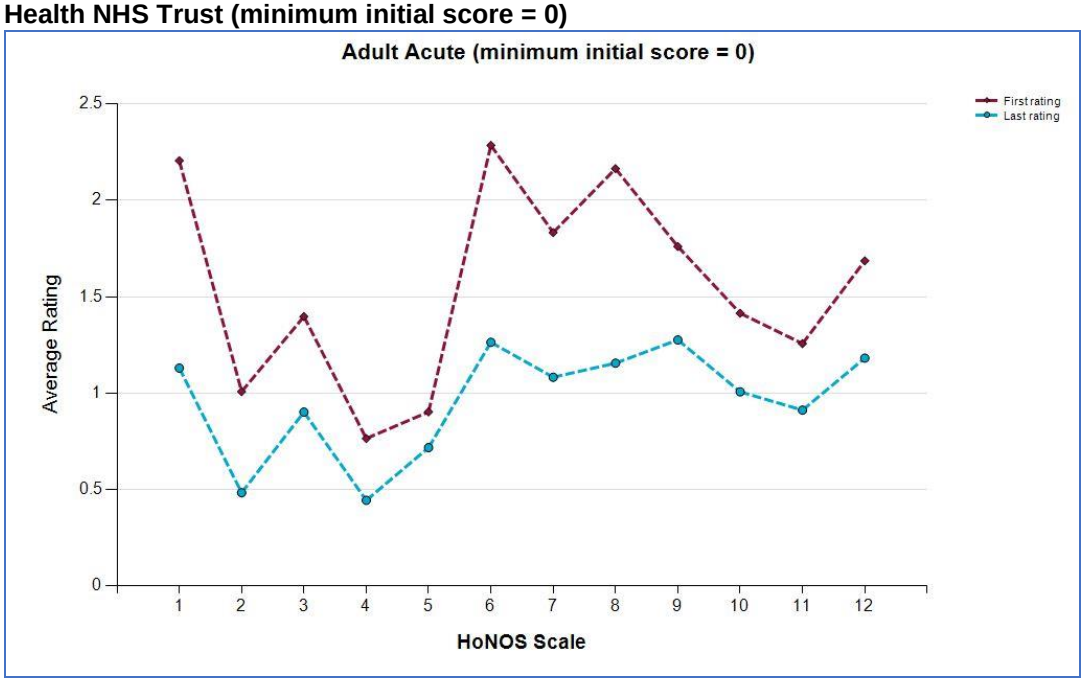
Another way of illustrating the data is through a line chart, showing a different line for the first and last ratings.

There are several line charts provided below, illustrating exactly the same data. The difference in profiles is due to the initial score chosen.

If the initial score for a scale is 0 – 'no problem' (common in community services), then the results in the last score for that scale are also likely to be 0. When looking at the mean team score for each scale, including results where the initial score is 0 can dilute the mean change for the team, when looking at the second HoNOS score. It has the effect of suppressing the mean change chart and masking the improvement made by patients who do score highly in a given domain.

The chart below shows the mean team scores, including all data where the initial score is 0.

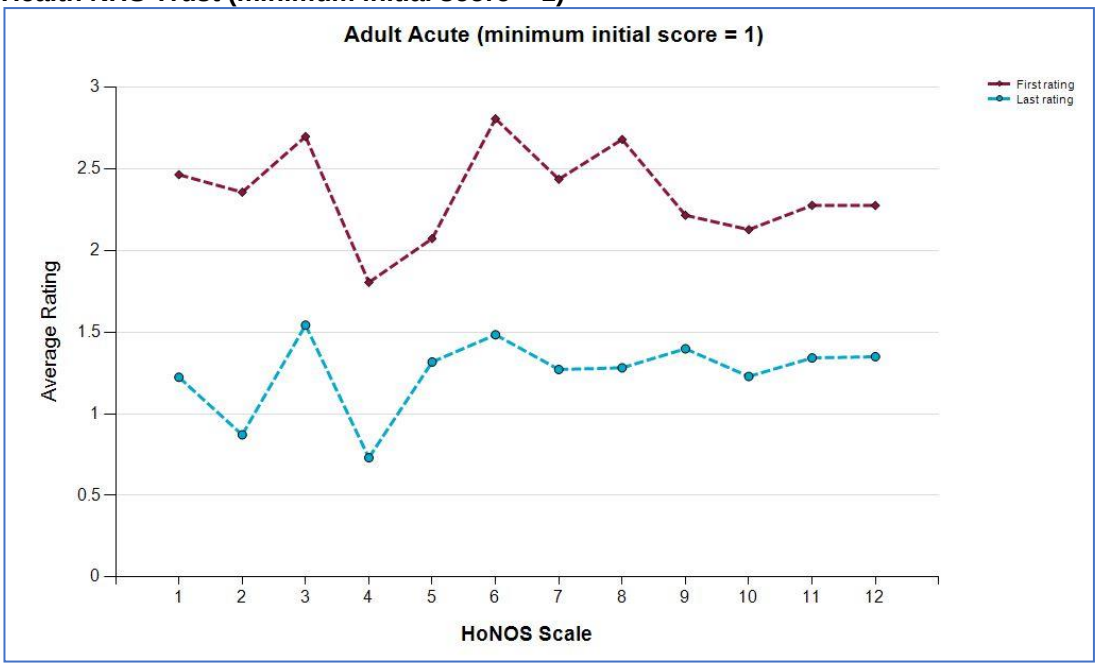
Figure 3: Example HoNOS profile chart, South West London and St George's Mental Health NHS Trust (minimum initial score = 0)



Starting the initial score at 1 (minor problem requiring no action) when calculating the team average would eliminate initial scores of 0 for that scale, so it only shows the effect on that scale for people who had a minor problem (score of 1) in that scale in the first place. Those with a score of 1 in a particular scale are unlikely to receive active medical or formal psychological treatment, so would not have an intervention from specialist (secondary) care mental health services for that scale.

The chart below shows the mean team scores, including all data where the initial score is 1.

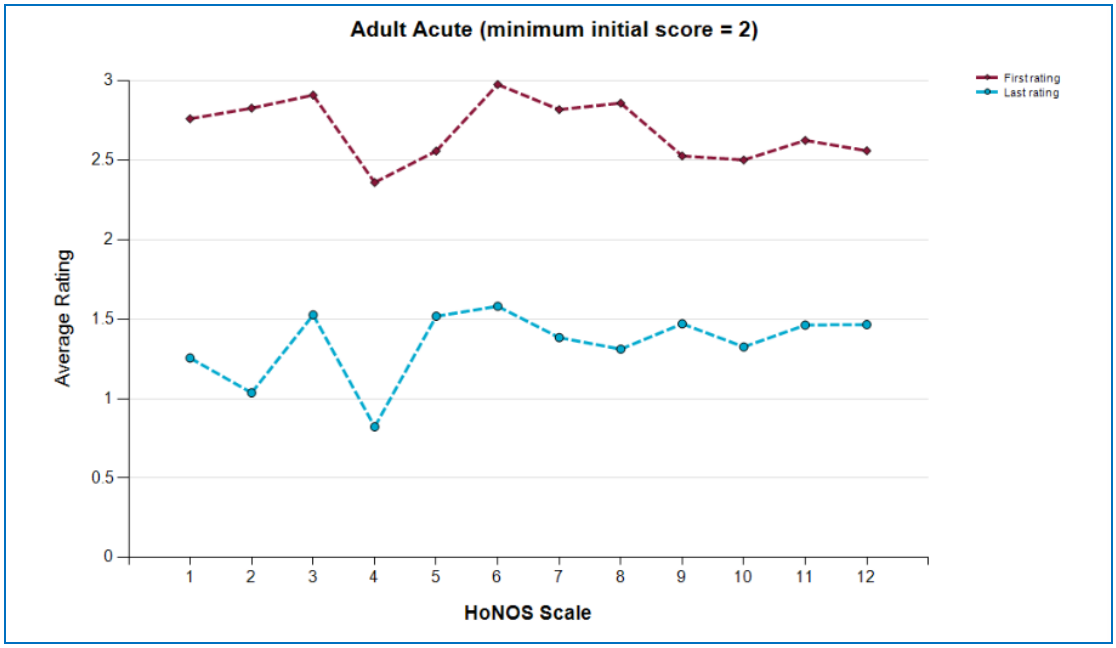
Figure 4: Example HoNOS profile chart, South West London and St George's Mental Health NHS Trust (minimum initial score = 1)



The initial score could also be set to 2 (mild problem but definitely present) depending on the service. Some London MH Trusts have decided to set the minimum threshold for calculating team average HoNOS scores, using an initial score of 2 (mild problem), as this reflects the severity of symptoms or difficulties that would normally be treated in secondary care.

The chart below shows the mean team scores, including all data where the initial score is 2.

Figure 5: Example HoNOS profile chart, South West London and St George's Mental Health NHS Trust (minimum initial score = 2)



It is worth considering setting a minimum initial score when calculating the team average when analysing HoNOS profiles within your Trusts.

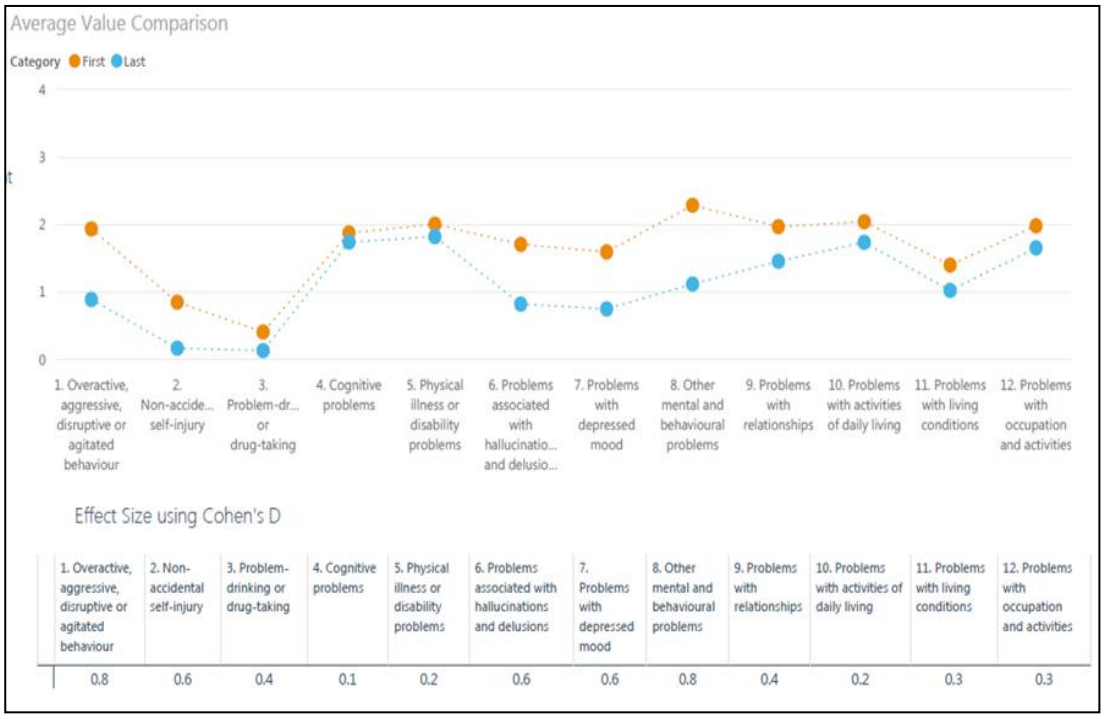
7. HoNOS profiles using statistical analysis

A provider may wish to incorporate statistical analysis into their HoNOS analysis, to calculate the significance of the change using effect size statistics. To assess whether the effect size might be significant, several London Trusts use Cohen's D (population standard deviation) to calculate effect sizes – see examples below.

Figure 6 below shows paired HoNOS profiles, with the effect size for each scale beneath it. The effect size aids interpretation as it can sometimes be difficult to interpret the size of the effect just by looking at the graph. In the graph below, there are large effect sizes (0.8) for *scale 1: Overactive, aggressive, disruptive or agitated behaviour* and *scale 8: Other mental and behavioural problems*.

Figure 6: Example HoNOS profile using Cohen's D to assess effect size, South London and Maudsley NHS Foundation Trust

Key:
0.2 – small effect size
0.5 – medium effect size
0.8 – large effect size



In Figure 7 below, the data shows that all scales have seen changes of moderate-critical clinical significance; the most significant change is in *scale 2: Non-accidental self-injury*, with an effect size of -1.46. The least significant change is in *scale 5: Physical illness or disability or disability problems*, with an effect size of -0.59.

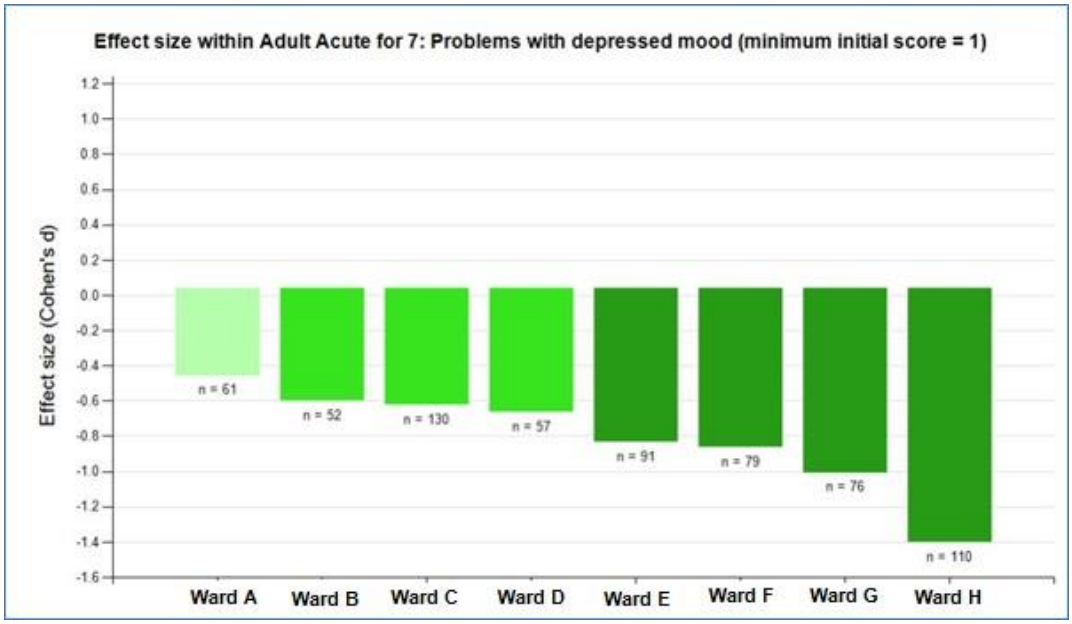
Figure 7: Example HoNOS Adult acute effect size using Cohen's D, South West London and St George's Mental Health NHS Trust

Effect size (Cohen's d)	Interpretation
-0.8 or lower	Improvement of critical clinical importance
-0.5 to -0.8	Improvement of moderate clinical significance
-0.2 to -0.5	Small or clinically negligible improvement
-0.2 to 0.2	No change
0.2 to 0.5	Small or clinically negligible deterioration
0.5 to 0.8	Deterioration of moderate clinical significance
0.8 or higher	Deterioration of critical clinical importance

Item	Paired HoNOS	Avg. first rating	Avg. last rating	Effect size
1: Overactive, aggressive, disruptive or agitated behaviour	791	2.46	1.22	-1.06
2: Non-accidental self-injury	373	2.36	0.87	-1.46
3: Problem-drinking or drug taking	433	2.70	1.54	-1.05
4: Cognitive problems	368	1.80	0.73	-0.86
5: Physical illness or disability or disability problems	372	2.07	1.32	-0.59
6: Problems associated with hallucinations and delusions	713	2.81	1.48	-1.02
7: Problems with depressed mood	656	2.44	1.27	-0.88
8: Other mental and behavioural problems	696	2.68	1.28	-1.01
9: Problems with relationships	668	2.22	1.40	-0.67
10: Problems with activites of daily living	573	2.13	1.23	-0.75
11: Problems with living conditions	471	2.28	1.34	-0.88
12: Problems with occupation and activities	632	2.28	1.35	-0.76

Effect size can also be used to look at the difference across teams or wards in improving outcomes for patients in particular scales. Figure 8 below shows the difference that similar wards make in improving patients' outcomes within *scale 7: Problems with depressed mood*. Ward H has three times the effect of ward A at improving outcomes for patients on that scale (-1.10 compared to -0.45). This analysis can be used to learn from what ward H is doing that has such a different impact on patient outcomes for this scale.

Figure 8: Example HoNOS Adult acute effect size using Cohen's D, within scale 7 (Problems with depressed mood), South West London and St George's Mental Health NHS Trust



8. Categorical Change

The Categorical Change method uses a scoring method which groups the answers scored for each question into two categories:

Low (L) severity [0-2] and High (H) severity [3-4].

It shows the percentage of patients that improve ('high to low'), deteriorate ('low to high') or remain unchanged ('high to high' or 'low to low') on each scale comparing first and last ratings.

This can be useful as it helps to evidence the nature of change in an easy to visualise format, and shows the dynamic nature of change.

Figure 9: Worked example of how to calculate HoNOS Categorical Change, Central and North West London NHS Foundation Trust

This example shows the string of HoNOS scores collected at two points for a patient. Each question is individually assessed for improvement.

HoNOS On Admission (Entry)

String	1	2	3	4	5	6	7	8	9	10	11	12
311331214332	3	1	1	3	3	1	2	1	4	3	3	2
	H	L	L	H	H	L	L	L	H	H	H	L

HoNOS On Discharge (Exit)

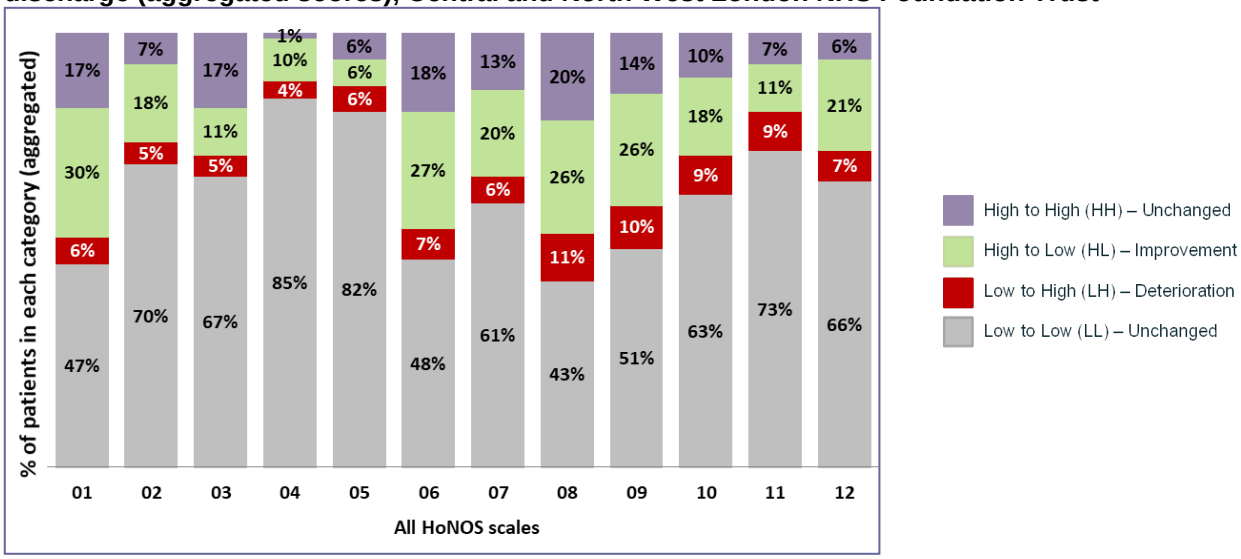
String	1	2	3	4	5	6	7	8	9	10	11	12
101310112113	1	0	1	3	1	0	1	1	2	1	1	3
	L	L	L	H	L	L	L	L	L	L	L	H

1	2	3	4	5	6	7	8	9	10	11	12
HL	LL	LL	HH	HL	LL	LL	LL	HL	HL	HL	LH

The score for *Scale 1: Overactive, aggressive, disruptive or agitated behaviour* shows an Entry score of **High** and exit score of **Low.** When put together there has been an improvement in that area for this patient.

This data can be visualised in a chart, as shown below:

Figure 10: Categorical change model by questions: Inpatient admission to inpatient discharge (aggregated scores), Central and North West London NHS Foundation Trust



To Note:
Determining which two points to take is critical and challenging. Where patients are transferred between wards the severity might remain **High** at the time of the transfer hence providing a false negative. Similarly determining the final exit HoNOS for a community patient is a challenge where patients have a multidisciplinary approach to their care.

9. Acknowledgements

We would like to thank everyone involved for their support and input into this analytical framework, in particular:

- Tricia Aveling, Clinical Lead for Mental Health Outcomes, North East London NHS Foundation Trust
- Asif Bachlani, Consultant Psychiatrist and Chief Clinical Information Officer, South West London and St George's Mental Health NHS Trust
- Rahul Bhattacharya, Consultant Psychiatrist and Associate Clinical Director, East London NHS Foundation Trust
- Nicola Byrne, Deputy Medical Director and Chief Clinical Information Officer, South London and Maudsley NHS Foundation Trust
- Martin Black, Quality Improvement Team, South London and Maudsley NHS Foundation Trust
- Jonathan Comfort, Information Manager, South West London and St George's Mental Health NHS Trust
- Andrew Crombie, Consultant Psychiatrist, Camden and Islington NHS Foundation Trust
- Mick James, HoNOS Advisor, Royal College of Psychiatrists
- Alastair MacDonald, Emeritus Professor of Clinical Outcomes Research at the Institute of Psychiatry, Psychology and Neuroscience at King's College London.
- Ayo Ogunkolade, Head of Information, Central and North West London NHS Foundation Trust
- Kevin Smith, Marisol Mental Health Consultants
- Debbie Taylor, Tariff Development Manager, North East London NHS Foundation Trust
- Bill Tiplady, Consultant Clinical Psychologist, Central and North West London NHS Foundation Trust
- Royal College of Psychiatry

Healthy London Partnership

Appendix 1: HoNOS Scales

The 12 HoNOS Scales

HoNOS comprises 12 scales that measure behaviour, impairment, symptoms and social functioning. The scales are completed after routine clinical assessments in any setting. The 12 HoNOS scales refer to what has happened recently and should be scored **within the last 2 weeks**.

SCALES	0	1	2	3	4
1. Overactive, aggressive, disruptive or agitated behaviour					
2. Non-accidental self-injury					
3. Problem drinking or drug-taking					
4. Cognitive problems					
5. Physical illness or disability problems					
6. Problems associated with hallucinations and / or delusions					
7. Problems with depressed mood					
8. Other mental and behavioural problems					
9. Problems with relationships					
10. Problems with activities of daily living					
11. Problems with housing and living conditions					
12. Problems with occupation and activities					

How is severity measured?

Level	Severity description
0	no problem
1	minor problem requiring no action
2	mild problem but definitely present
3	moderately severe problem
4	severe to very severe problem