

NHS Patient Survey Programme

2018 community mental health survey

Identifying outliers in trust-level results

Published November 2018
CQC publication

Contents

Summary	2
Interpreting the results	4
Results	5
Trusts achieving ‘better than expected’ results	5
Trusts achieving ‘worse than expected’ results	6
Trusts achieving ‘much worse than expected’ results	7
Appendix A: Analysis methodology	8
Identifying worse than expected experience	8
Identifying better than expected experience	8
Weighting	9
Scoring	9
Appendix B: Analytical stages of the outlier model	10
Appendix C: Difference between outlier analysis and trust-level benchmark reports	13

Summary

The 2018 community mental health survey received feedback from 12,796 people who were treated for a mental health condition between 1 September and 30 November 2017. This is a response rate of 28%,^a which is an increase from last year's response rate of 26%.

We have published an analysis of the national results on our [website](#). In this separate analysis, we identify the trusts where people reported care experiences that were either better or worse than expected when compared with the survey results across the other trusts. The methodology for the analysis of the 2018 community mental health survey (detailed in [Appendix A](#) and [B](#)) identifies variation in results at the trust level. It differs from the approach used in trust-level benchmark reporting, which provides mean scores for individual questions only.

[Appendix C](#) provides more information on our different approaches to using survey data to explore variation in experience between trusts.

Each trust is categorised into one of five bands: 'much worse than expected', 'worse than expected', 'about the same', 'better than expected' or 'much better than expected'.

No trusts are in the highest band of 'much better than expected' this year.

2gether NHS Foundation Trust is in the 'better than expected' category for the second year running. Three other trusts have been identified as providing 'better than expected' experiences for people using their services this year: Cheshire and Wirral Partnership NHS Foundation Trust, Dorset Healthcare University NHS Foundation Trust and NAVIGO Health and Social Care CIC.

Northamptonshire Healthcare NHS Foundation Trust has been identified as 'worse than expected' again in 2018, with Lincolnshire Partnership NHS Foundation Trust in this category for the first time.

Leicestershire Partnership NHS Trust has been categorised in the lowest band of 'much worse than expected'. Isle of Wight NHS Trust remains in this category, unchanged from 2017.

a. We report the 'adjusted' response rate. The adjusted base is calculated by subtracting the number of questionnaires returned as undeliverable, or if someone had died, from the total number of questionnaires sent out. The adjusted response rate is then calculated by dividing the number of returned useable questionnaires by the adjusted base.

CQC's Deputy Chief Inspector of Hospitals (and lead for mental health), Dr Paul Lelliott, has written to all trusts identified as better or worse than expected and these letters have been shared with [NHS Improvement](#).^b

We will continue to reflect each trust's performance on this survey using our Insight products as part of the wider information we have on how trusts are performing. We recognise that trusts may have been working locally to improve services since the survey took place. However, we have asked the trusts that were identified as much worse to review their results and to outline what actions they will take to continue to address the areas of concern.

b. NHS Improvement oversees NHS trusts and independent organisations that provide NHS-funded care. It supports providers to give patients consistently safe, high-quality, compassionate care in local health systems. NHS Improvement will use the results of this survey to inform quality and governance activities as part of its Oversight Model for NHS Trusts.

Interpreting the results

We have calculated the overall proportion of responses that each trust received for the 'most negative', 'middle' and 'most positive' answer option(s) across all scored questions in the survey.^c

We use the following question from the 2018 community mental health survey to show how responses are categorised as either 'most negative', 'middle' and 'most positive'.

Q4. Were you given **enough time** to discuss your needs and treatment?

- Yes, definitely – **most positive**
- Yes, to some extent – **middle**
- No – **most negative**
- Don't know/can't remember – excluded from the analysis

Where people's experience of using a trust's services are either better or worse than elsewhere, there will be a significant difference between that trust's results and the average results across all trusts. Each trust is then assigned a banding of either 'much worse than expected', 'worse than expected', 'about the same', 'better than expected' or 'much better than expected' depending on how significant that variation is.

For example, a trust's proportion of responses breaks down as: 'most negative' 15%, 'middle' 24% and 'most positive' 61%. This is then compared with the trust average of 'most negative' 20%, 'middle' 25% and 'most positive' 55%. An 'adjusted z-score'^d is calculated for the difference between 'most positive' trust proportions which, in this example, is -2.12. This means that this trust has a higher proportion of 'positive' responses than the trust average. This is considered significant with a p-value of less than 0.25 (z-score lower than -1.96) but not less than 0.01 (z-score -3.09). As a result, the trust is classed as 'better'.

Finally, each table in the report includes the most recent trust-wide CQC rating. See [Appendix B](#) for details of the analytical method used to calculate these results.

c. Filter questions were not included in this analysis.

d. Z scores give an indication of how different a trust's proportion is from the average.

Results

Trusts achieving 'better than expected' results

Four trusts were classed as 'better than expected' in 2018. The banding for 2gether NHS Foundation Trust is unchanged from 2017.

	Historic results	Overall results			Overall CQC rating	
	2017	2018	Most Negative (0/10)	Middle		Most Positive (10/10)
Trust average			20	25	55	
2gether NHS Foundation Trust	B	B	15	24	61	G
Cheshire and Wirral Partnership NHS Foundation Trust	S	B	16	23	61	G
Dorset Healthcare University NHS Foundation Trust	S	B	16	23	61	G
NAVIGO Health and Social Care CIC	S	B	14	23	63	G

Key:	Trust performance	About the same (S)	Better (B)	Much better (MB)	
	CQC rating	Inadequate (I)	Requires Improvement (RI)	Good (G)	Outstanding (O)

Trusts achieving 'worse than expected' results

Two trusts were identified as 'worse than expected' this year across the entire survey. Northamptonshire Healthcare NHS Foundation Trust was also 'worse than expected' in 2017.

	Historic results	Overall results			Overall CQC rating	
	2017	2018	Most Negative (0/10)	Middle		Most Positive (10/10)
Trust average			20	25	55	
Lincolnshire Partnership NHS Foundation Trust	S	W	24	23	52	G
Northamptonshire Healthcare NHS Foundation Trust	W	W	25	26	49	O

Key:	Trust performance	About the same (S)	Worse (W)	Much worse (MW)
	CQC rating	Inadequate (I)	Requires Improvement (RI)	Good (G)

Trusts achieving 'much worse than expected' results

Two trusts have been classed as 'much worse than expected' this year across the entire survey. Isle of Wight NHS Trust, rated as inadequate by CQC, was also categorised as 'much worse than expected' in 2017.

	Historic results	Overall results			Overall CQC rating	
	2017	2018	Most Negative (0/10)	Middle		Most Positive (10/10)
Trust average			20	25	55	
Isle of Wight NHS Trust	MW	MW	28	28	44	I
Leicestershire Partnership NHS Trust	S	MW	28	25	47	RI

Key:	Trust performance	About the same (S)	Worse (W)	Much worse (MW)
	CQC rating	Inadequate (I)	Requires Improvement (RI)	Good (G)

Appendix A: Analysis methodology

Identifying worse than expected experience

The analytical approach to identifying those trusts where people's experiences of services was 'much/worse than expected' uses responses for all scored questions (except the overall experience question).^e

For each trust, we calculate the number of responses scored as '0' (the most negative option). This is then divided by the total number of responses scored as 0-10 to calculate the trust-level proportion of poor experience. A higher percentage of negative responses indicates poorer experience.

The analysis uses z-scores, which indicate how different a trust's poor experience proportion is from the average.

There are two thresholds for flagging trusts with concerning levels of poor experience:

- **Worse than expected:** z-score lower than -1.96
- **Much worse than expected:** z-score lower than -3.09

Appendix B provides full technical detail of the analytical process used.

Identifying better than expected experience

To identify 'much/better than expected' experience, we calculate a count of the number of responses scored as '10' (the most positive option) for each trust.

This is then divided by the total number of responses scored as 0-10 to calculate the trust-level proportion of positive experience. A higher percentage of positive responses indicates better experience.

Our analysis has found that those trusts with the highest proportion of positive responses also tend to have the lowest proportion of negative responses.

There are two thresholds for identifying trusts with high levels of good experience reported by people using services:

- **Better than expected:** z-score lower than -1.96
- **Much better than expected:** z-score lower than -3.09

e. Overall experience is excluded from the analysis due to the ambiguity around what should be classed as the 'most negative' (and 'most positive') option(s).

Weighting

Results have been standardised by the age and gender of respondents to ensure that no trust will appear better or worse than another because of its respondent profile.

Standardisation enables a more accurate comparison of results from trusts with different population profiles. In most cases this will not have a large impact on trust results, but it does make comparisons between trusts as fair as possible.

Scoring

For each question in the survey, the individual (standardised) responses are converted into scores on a scale from 0 to 10. A score of 10 represents the best possible response and a score of 0 the worst. The higher the score for each question, the better the trust is performing.

It is not appropriate to score all questions in the questionnaire as not all questions assess trusts' performance. For example, they may be descriptive questions such as Q1 asking when was the last time they saw someone from NHS mental health services.

Appendix B: Analytical stages of the outlier model

The analytical approach to identifying outliers is based on all evaluative items in the survey. These are the questions that are scored for benchmarking purposes. The scored variables are the source data, and are required at respondent level. These variables take values between 0 (representing the worst rating of experience) and 10 (representing the best rating). The approach also makes use of the standardisation weight for the survey.

1. Count the poor-care ratings made by each respondent^f

Count of the '0' responses across the scored questions answered by each respondent (excluding the "Overall..." question).

2. Count the questions given specific (scored) answers by each respondent

Count of all '0-10' responses across the scored questions answered by each respondent (excluding the "Overall..." question).

3. Weight the data

Apply the standardisation weight for respondents. The weight adjusts the population of respondents in each trust to the national average proportions for age and gender.

4. Aggregate to trust-level and compute proportion of poor ratings

Obtain a weighted numerator and denominator for each trust. Divide the numerator by the denominator to obtain the trust-level proportion of poor care ratings, for example the overall percentage of responses which were scored as 0.

5. Compute the mean of the trust-level proportions

Sum all proportions and divide by the number of trusts to obtain the average trust-level proportion of poor care ratings.

f. The analytical approach used to identify positive experience of people using services uses a numerator count of the '10' responses across all scored questions (excluding the "overall..." question) to calculate the 'good-care ratings'. There are no other differences between the analytical approaches for identifying poor and good experience.

6. Compute the z-score for the proportion

The Z-score formula used is:

$$z_i = -2\sqrt{n_i} \{ \sin^{-1}(\sqrt{p_i}) - \sin^{-1}(\sqrt{p_0}) \} \quad (1)$$

where: n_i is the denominator for the trust

p_i is the trust proportion of poor care ratings

p_0 is the mean proportion for all trusts

7. Winsorize the z-scores

Winsorizing consists of shrinking in the extreme Z-scores to some selected percentile, using the following method:

1. Rank cases according to their naive Z-scores.
2. Identify Z_q and $Z_{(1-q)}$, the 100q% most extreme top and bottom naive Z-scores. For this work, we used a value of $q=0.1$
3. Set the lowest 10% of Z-scores to Z_q , and the highest 10% of Z-scores to $Z_{(1-q)}$. These are the Winsorized statistics.

This retains the same number of Z-scores but discounts the influence of outliers.

8. Calculate dispersion using Winsorized z-scores

An over dispersion factor $\hat{\phi}$ is estimated which allows us to say if the data are over dispersed or not:

$$\hat{\phi} = \frac{1}{I} \sum_{i=1}^I z_i^2 \quad (2)$$

Where I is the sample size (number of trusts) and z_i is the Z score for the i th trust given by (1). The Winsorized Z scores are used in estimating $\hat{\phi}$.

9. Adjust for overdispersion

If $\hat{\phi}$ is greater than $(I - 1)$ then we need to estimate the expected variance between trusts. We take this as the standard deviation of the distribution of p_i (trust proportions) for trusts, which are on target, we give this value the symbol $\hat{\tau}$, which is estimated using the following formula:

$$\hat{\tau}^2 = \frac{I\hat{\phi} - (I-1)}{\sum_i w_i - \sum_i w_i^2 / \sum_i w_i} \quad (3)$$

where $s_i = (p_i - p_0)/z_i$, $w_i = 1/s_i^2$ and $\hat{\phi}$ is from (2). Once $\hat{\tau}$ has been estimated, the Z_D score is calculated as:

$$Z_i^D = \frac{p_0 - p_i}{\sqrt{s_i^2 + \hat{\tau}^2}} \quad (4)$$

Appendix C: Difference between outlier analysis and trust-level benchmark reports

The approach used to analyse trust variation in this report is focused on identifying significantly higher levels of better or worse experience **across the entire survey**.

This holistic approach is different to the technique used to analyse results in trust benchmarking reports. In these reports, trust results for each scored question are assigned bands of either 'better', 'worse' or 'about the same' when compared with the findings for all other trusts. However, trust benchmark reports do not attempt to look across all questions concurrently and therefore do not provide an overall assessment of the proportion of positive or negative experiences reported across the entire survey.

While both approaches are useful, analysis of individual questions can hide variation in people's experience as the scores are 'averaged' in that analysis. The approach used in this report allows CQC to identify that variation and highlight potential concerns raised by people across the survey in its entirety.

How to contact us

Call us on: 03000 616161

Email us at: enquiries@ccq.org.uk

Look at our website: www.cqc.org.uk

Write to us at: Care Quality Commission
Citygate
Gallowgate
Newcastle upon Tyne
NE1 4PA



Follow us on Twitter: [@CareQualityComm](https://twitter.com/CareQualityComm)

Please contact us if you would like a summary of this document in another language or format.

Corporate member of
Plain English Campaign
Committed to clearer
communication

459

