

NHS Patient Survey Programme

**2017 Maternity
Experience Survey**

Identifying outliers within
trust-level results

Published February 2018

Contents

Summary	2
Interpreting the results	4
Difference between outlier analysis and trust-level benchmark reports	5
Results	6
Trusts achieving better than expected results	6
Trusts achieving worse than expected results.....	7
Trusts achieving much worse than expected results	8
Further information	9
Appendix A: Analysis methodology	10
Appendix B: Analytical stages of the outlier model	12

Summary

The 2017 Maternity Survey included 130 NHS trusts. Feedback was received from 18,426 mothers, a response rate of 37%.

Women aged 16 and over at the time of delivery were eligible for the survey if they had a live birth during the month of February 2017. Only women who received care from an NHS trust were considered in the sampling process.

The maternity pathway includes different stages: antenatal care, labour and birth, and postnatal care. This report provides an analysis of mothers' experiences during labour and birth because we can confidently attribute the care provided during this stage to the respective trusts. However, we cannot yet guarantee the reliability of the attribution data for antenatal and postnatal care, since women can receive these types of care from different providers and not all trusts submitted the necessary attribution information.

We have published the national results analysis on our [website](#). In this separate analysis, we identify the trusts where patients' experiences are better, or worse, than expected when we compare the survey results across trusts.

A new analysis methodology has been used for the 2017 Maternity Survey (detailed in appendix [A](#) and [B](#)) to identify variation in results at trust-level. This methodology is more robust as all scored questions for the labour and birth stage are analysed simultaneously and trust performance is no longer assessed using mean scores, which can mask where experience is highly polarised. There is more information on the difference between approaches in the section on [outlier analysis and trust-level benchmark reports](#).

Each trust has been categorised into one of five bands: 'much worse than expected', 'worse than expected', 'about the same', 'better than expected' or 'much better than expected'. The methodology for this model is set out in appendices [A](#) and [B](#).

Positive outliers

There were no trusts flagged as 'much better than expected' within this outlier analysis. Nevertheless, there were four trusts whose overall performance was identified as 'better than expected':

- Dartford and Gravesham NHS Trust
- East Cheshire NHS Trust
- Maidstone and Tunbridge Wells NHS Trust
- North Cumbria University Hospitals NHS Trust.

Negative outliers

Results from the analysis showed that overall, patients from five trusts had ‘worse than expected’ experiences of maternity care:

- Barts Health NHS Trust
- Birmingham Women's and Children's NHS Foundation Trust
- Hull and East Yorkshire Hospitals NHS Trust
- London North West Healthcare NHS Trust
- North Tees and Hartlepool NHS Foundation Trust.

Furthermore, two trusts were identified as ‘much worse than expected’ when all questions were analysed simultaneously:

- Sandwell and West Birmingham Hospitals NHS Trust
- The Rotherham NHS Foundation Trust.

Our Chief Inspector of Hospitals, Professor Edward Baker, has written to all trusts flagged as much worse by the outliers analysis. These letters have been shared with [NHS Improvement](#).¹ CQC is aware that trusts may have been working locally to improve services since the survey was carried out. However, trusts categorised as performing ‘worse’ or ‘much worse’ than others were asked to review their results to identify where patients feel they are being let down.

CQC will continue to reflect each trust’s performance in this survey in our CQC Insight monitoring as part of the wider information we hold on their performance. We will also review trusts’ progress on the actions they take to improve during their next planned inspection. As part of this, our inspection teams will focus on the areas raised in the survey where results suggest that people’s experiences were worse than we would expect, and they will look for reassurance that trusts are taking appropriate action.

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

Interpreting the results

We have calculated the proportion of responses that each trust received for the 'most negative', 'middle' and 'most positive' answer option(s) across all of the scored questions in the survey.² More details about the stages of this analysis can be found in Appendix [B](#).

The following example presents how response options are categorised as either 'most negative', 'middle' or 'most positive'.

C20. Did you have confidence and trust in the staff caring for you during your **labour and birth**?

- Yes, definitely - **most positive**
- Yes, to some extent - **middle**
- No - **most negative**
- Don't know - not included

The outlier analysis shows where the experience of patients is significantly better or worse than the average across all trusts. Each trust is then assigned a banding of either 'much better than expected', 'better than expected', 'about the same', 'worse than expected' or 'much worse than expected' depending on how statistically significant the variation is.

Taking a hypothetical trust as an example, its proportion of responses breaks down as: 'most negative' 13%, 'middle' 8% and 'most positive' 79%. This is then compared with the trust average of 'most negative' 11%, 'middle' 11% and 'most positive' 78%. The adjusted z-score for the difference between the hypothetical trusts 'most negative' proportion (13%) and the trust average 'most negative' proportion (11%) is -1.21. This means that, despite a higher proportion of most negative responses than the trust average, this is not considered significant and the hypothetical trust is categorised as 'about the same'.

We present the results on pages [6-8](#). They show the trusts' survey bandings (under the 2017 column header), their CQC overall and maternity service ratings. The middle columns show the percentage of 'most positive' responses (scored 10/10), 'most negative' responses (scored 0/10) and 'middle' responses (scored on a scale between 0 and 10) achieved by the trust. The trust average is the average across all participating trusts in England.

2. The analysis only includes questions that are able to be scored. Please see the [scored questionnaire](#) to see which questions these are.

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 patient experience **across the entire survey**.

This approach is different to the one applied in the benchmarking reports for trusts. In benchmarking reports, trusts' results for each scored question are assigned to 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 patient experiences reported across the entire survey.

Historically any trust that received a banding of 'worse/better' for at least 20% of the scored survey questions was considered as being 'worse/better than expected' overall. The analysis methodology we use in this report has replaced the 20% better/worse rules-based method.

While both approaches are useful, analysing individual questions can disguise differences in people's experience as the scores are 'averaged' for each question. This new approach allows CQC to identify that variation and highlight potential concerns reflected in women's responses.

Results

Trusts achieving better than expected results

Four trusts were flagged as 'better than expected' across all labour and birth questions. All of these trusts had an overall CQC rating of requires improvement.

	Overall results				Overall CQC rating	Maternity service rating		
	2017	Most Negative (0/10)	Middle	Most Positive (10/10)		Site 1	Site 2	Site 3
Trust average		11	11	78				
Dartford and Gravesham NHS Trust	B	7	9	84	RI	G		
East Cheshire NHS Trust	B	7	8	85	RI	RI		
Maidstone and Tunbridge Wells NHS Trust	B	8	9	84	RI	G	RI	
North Cumbria University Hospitals NHS Trust	B	8	8	84	RI	RI	G	RI

Key:	Trust performance	Much worse (MW)	Worse (W)	About the same (S)	Better (B)	Much better (MB)
	CQC rating	Inadequate (I)	Requires improvement (RI)	Good (G)	Outstanding (O)	No rating (NR)

Trusts achieving worse than expected results

Five trusts were flagged as ‘worse than expected’ overall across all questions relating to experience of labour and birth.

Four trusts that achieved ‘worse than expected’ had an overall CQC rating of requires improvement and one trust had an overall CQC rating of outstanding. Birmingham Women’s and Children’s NHS Foundation Trust did not have a rating for maternity services at the time of this report.

	Overall results				Overall CQC rating	Maternity service rating		
	2017	Most Negative (0/10)	Middle	Most Positive (10/10)		Site 1	Site 2	Site 3
Trust average		11	11	78				
Barts Health NHS Trust	W	15	15	70	RI	G	RI	RI
Birmingham Women's and Children's NHS Foundation Trust	W	14	12	74	O	NR		
Hull and East Yorkshire Hospitals NHS Trust	W	14	12	74	RI	RI		
London North West Healthcare NHS Trust	W	16	13	71	RI	RI		
North Tees and Hartlepool NHS Foundation Trust	W	14	13	73	RI	RI	RI	

Key: Trust performance	Much worse (MW)	Worse (W)	About the same (S)	Better (B)	Much better (MB)
CQC rating	Inadequate (I)	Requires improvement (RI)	Good (G)	Outstanding (O)	No rating (NR)

Trusts achieving much worse than expected results

Two trusts were flagged as ‘much worse than expected’ across all labour and birth questions, both trusts had an overall CQC rating of requires improvement.

	Overall results			Overall CQC rating	Maternity service rating Site 1
	2017	Most Negative (0/10)	Middle		
Trust average		11	11	78	
Sandwell and West Birmingham Hospitals NHS Trust	MW	18	13	69	G
The Rotherham NHS Foundation Trust	MW	18	12	70	RI

Key:	Trust performance	Much worse (MW)	Worse (W)	About the same (S)	Better (B)	Much better (MB)
	CQC rating	Inadequate (I)	Requires improvement (RI)	Good (G)	Outstanding (O)	No rating (NR)

Further information

The results for England and trust-level results are available on CQC's website. There is also a technical document that sets out the methodology used in the trust-level benchmark results analysis. On CQC's website there is also a Quality and Methodology report that provides additional methodological details.

<http://www.cqc.org.uk/maternitysurvey>

Results at a trust level for previous iterations of the Maternity Survey are available at the link below.

<http://www.nhssurveys.org/surveys/299>

Full details of the methodology for the survey, including questionnaires, scored questionnaire, letters sent to patients, instructions on how to carry out the survey and the survey development report are available at:

<http://www.nhssurveys.org/surveys/1055>

More information on the patient survey programme, including results from other surveys and a programme of current and forthcoming surveys is available at:

www.cqc.org.uk/surveys

More information on how CQC monitors hospitals is available on CQC's website at:

www.cqc.org.uk/content/monitoring-nhs-acute-hospitals

Appendix A: Analysis methodology

Identifying worse than expected patient experience

The analytical approach to identifying those trusts where women's experiences were 'worse than expected' uses responses for all scored questions asking about labour and birth.

For each trust, we count 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 indicate a poor patient experience.

The analysis uses z-scores to indicate the difference between the proportion of poor experience in a trust and the average.

There are two thresholds for flagging trusts with concerning levels of poor patient 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.

Identifying better than expected patient experience

To identify 'better than expected' patient 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 poor experience.

A higher percentage of positive responses is indicative of good patient experience.

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

There are two thresholds for identifying trusts with high levels of good patient experience:

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

Weighting

As in the benchmark results for each trust, results have been standardised by women's age and parity (whether women have given birth previously or not) to ensure that no trust will appear better or worse than another because of the profile of its respondents.

Standardisation allows a more accurate comparison of results from trusts that have different population profiles. In most cases, this will not have a large impact on a trust's results. However, 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.

Where a number of options lay between the negative and positive responses, they were placed at equal intervals along the scale. The example below shows the scoring for question C20.

C20. Did you have confidence and trust in the staff caring for you during your **labour and birth**?

- Yes, definitely - **most positive**
- Yes, to some extent - **middle**
- No - **most negative**
- Don't know - not included

For more detail, please see either the scored questionnaire or the technical document (see [further information](#) section).

It is not appropriate to score all questions in the questionnaire, as not all of them assess trusts' performance. For example, they may be descriptive questions such as asking women if they had a home birth. Questions were only used in this analysis if they asked about labour and birth.

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 case 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 respondent³

Count of the '0' responses across the scored labour and birth questions answered by each respondent.

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

Count of all '0-10' responses across the scored labour and birth questions answered by each respondent.

3. Weight the data

Apply the standardisation weight for respondents. The weight adjusts the population of respondents within each trust to the national average proportions for age and parity (whether women previously had a baby).

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, i.e. 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.

³ The analytical approach used to identify positive patient experience uses a numerator count of the '10' responses across all scored labour and birth questions to calculate the 'good-care ratings'. There are no other differences between the analytical approaches for identifying poor and good patient 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 over dispersion

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:

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

How to contact us

Call us on: 03000 616161

Email us at: enquiries@cqcc.org.uk

Look at our website: www.cqcc.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)

Corporate member of
Plain English Campaign

Committed to clearer
communication

459

