NHS Patient Survey Programme

2016 Emergency Department Survey

Identifying outliers within trust-level results

Published October 2017

Contents

Summary

The 2016 Emergency Department survey involved 137 NHS trusts. People aged 16 and over were eligible to participate if they had attended a Type 1 or Type 3 emergency department run by an NHS acute trust[1] during September 2016.

Type 1: A major, consultant-led A&E department with full resuscitation facilities operating 24 hours a day, seven days a week.

Type 3: Other A&E/minor injury unit/urgent care centre treating minor injuries and illnesses. Can be doctor or nurse-led and accessed without appointment.

This report provides an analysis of the results from patients who had attended a Type 1 emergency department only. We received a smaller number of responses, from a smaller sample size, for people attending Type 3 settings.[2] We report on Type 3 settings in the statistical release of the overall results for England, published on our website.

Although 88 trusts provided a Type 1 sample only, this does not necessarily mean that there are no other alternative urgent care services available locally. For example, there may be services outside of the scope of the survey, such as walk-in centres (Type 4), a minor injury unit or urgent care centre run by another provider, or an out-of-hours GP service. If alternative local services are available for less serious cases, this would affect the type of patients (case-mix) seen at the Type 1 department. As it is not possible to account for this local variation in the analysis, it is important to take this into account when interpreting trust level results.

The analysis of the results for England and the trust-level benchmark results are published on CQC's website. In this report, we identify variation in results at trust level across all survey questions. This new methodology ensures that all 'scored' (evaluative) questions are analysed simultaneously and moves beyond assessing a trust's performance using mean scores, which can mask where experience is highly polarised. The new methodology is set out in appendix A and B, and the next section explains the difference between approaches.

We assigned each trust into one of five bands: 'much better than expected', 'better than expected', 'about the same, 'worse than expected' or 'much worse than expected'.

1. Only Type 3 departments run directly by NHS acute trusts were included, and not those run in collaboration with, or exclusively by others.

2. For Type 1 services, we received responses from more than 40,500 people, a response rate of 28%. For Type 3 services, we received responses from more than 3,500 people, a response rate of 25%.

For this survey, no trusts achieved the highest band, 'much better than expected'.

Encouragingly, patients from the following six trusts experienced care that was 'better than expected':

- Taunton and Somerset NHS Foundation Trust
- Harrogate and District NHS Foundation Trust
- Poole Hospital NHS Foundation Trust
- Salisbury NHS Foundation Trust
- Gateshead Health NHS Foundation Trust
- North Bristol NHS Trust

Eight trusts were identified as achieving 'worse than expected' results:

- Bradford Teaching Hospitals NHS Foundation Trust
- North Middlesex University Hospital NHS Trust
- Barking, Havering and Redbridge University Hospitals NHS Trust
- Royal Liverpool and Broadgreen University Hospitals NHS Trust
- Homerton University Hospital NHS Foundation Trust
- East Sussex Healthcare NHS Trust
- Blackpool Teaching Hospitals NHS Foundation Trust
- Imperial College Healthcare NHS Trust

In addition, the Hillingdon Hospitals NHS Foundation Trust was identified as achieving 'much worse than expected' results.

Our Chief Inspector of Hospitals, Professor Edward Baker, has written to all trusts that the survey identified as being better or worse than average and these letters have been shared with NHS Improvement[3]. We recognise that trusts may have been working locally to improve services since the survey took place, however, the nine trusts identified as worse, or much worse, have been asked to review their results and to outline what actions they will take to continue to address the areas of concern. CQC will continue to reflect each trust's performance on this survey within our Insight products as part of the wider information we have on how trusts are performing; and will review trust progress during their next planned inspection. As part of our inspections, our inspection teams will be focusing on the areas raised in the survey where results suggest that people's experiences were worse than we would expect and looking for reassurance that appropriate action is being taken.

[3] 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 the emergency department survey to inform quality and governance activities as part of its Oversight Model for NHS Trusts.

Outlier analysis and trust-level benchmark reports

To analyse the variation between trusts, we focus on identifying significantly higher levels of better or worse patient experience across the entire survey. This enables us to identify a trust as an ‘outlier’ in the survey.

This holistic approach is different to the technique we use to analyse results in the benchmarking reports for trusts. In the 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, benchmark reports do not attempt to look across all questions as a whole and therefore do not provide an overall assessment of the proportion of positive or negative patient experience reported across the entire survey.

Furthermore, being assigned to a band of ‘better’ for an overall experience question is not the same as being ‘better than expected’ across the entire survey. For comparison, appendix C lists all trusts that were assigned to a band of ‘worse’ or ‘better’ for the overall experience question.

Historically, any trust that received a banding of ‘worse/better’ for at least 20% of scored survey questions was considered as being ‘worse/better than expected’ across the entire survey. 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 hide variation in people’s experience as the scores are ‘averaged’ in that analysis. This new approach allows CQC to identify that variation and to highlight potential concerns that were raised by some people across the survey.

Interpreting the results

We 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.[4]

The example below shows how responses are categorised as either 'most negative', 'middle' or 'most positive'.

Q14. Did the doctors and nurses listen to what you had to say?

- Yes, definitely – most positive
- Yes, to some extent – middle
- No – most negative

For this outlier analysis, where the experience of a trust's patients is either better or worse than elsewhere, there will be a statistically significant difference between that trust's result and the average result 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 significant that variation is.

If we take a hypothetical trust as an example, its proportion of responses breaks down as: 'most positive' 63%, 'middle' 22% and 'most negative' 15%. This is then compared with the trust average of 'most positive' 68%, 'middle' 18% and 'most negative' 14%. The adjusted z-score for the difference between the hypothetical trust's 'most negative' proportion (15%) and the trust average 'most negative' proportion (14%) is -0.37. This means that, despite the hypothetical trust having 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'.

The tables on pages 6-8 provide the results. They show trusts' survey banding (under the '2016' column header) and their CQC overall rating and core service rating (urgent and emergency services). 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 trusts. So for example, Taunton and Somerset NHS Foundation Trust achieved a survey banding of 'better than expected', an overall CQC rating of good, and a core service rating of requires improvement. The trust achieved the most positive response options for 72% of all questions, the average for all trusts for this is 66%.

4. The analysis only includes questions that are able to be scored. Please see the scored questionnaire to see which questions these are.

Trusts achieving ‘better than expected’ results

Six trusts were classed as ‘better than expected’ across the entire survey. No trusts achieved ‘much better than expected’.

Three trusts that achieved ‘better than expected’ have received an overall CQC rating of good. Three of these trusts have received an overall CQC rating of requires improvement.

		Overall results (%)				
	2016	Most Positive (10/10)	Middle	Most Negative (0/10)	Overall CQC rating	Core service rating
Trust average		**66**	**23**	**11**		
Taunton and Somerset NHS Foundation Trust	(B)	72	20	8	(G)	(RI)
Harrogate and District NHS Foundation Trust	(B)	73	20	7	(G)	(RI)
Poole Hospital NHS Foundation Trust	(B)	72	20	8	(RI)	(G)
Salisbury NHS Foundation Trust	(B)	73	19	8	(RI)	(RI)
Gateshead Health NHS Foundation Trust	(B)	73	19	8	(G)	(G)
North Bristol NHS Trust	(B)	73	20	8	RI	(G)

Key:

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

Core service rating for: urgent and emergency services

Trusts achieving ‘worse than expected’ results

Eight trusts were classed as ‘worse than expected’. CQC has rated six of these trusts as requires improvement.

	Overall results (%)				Overall CQC rating	Core service rating
	2016	Most positive (10/10)	Middle	Most negative (0/10)		
Trust average		**66**	**23**	**11**		
Bradford Teaching Hospitals NHS Foundation Trust	(**W**)	59	25	16	(RI)	(RI)
North Middlesex University Hospital NHS Trust	(**W**)	57	27	16	(RI)	(RI)
Barking, Havering and Redbridge University Hospitals NHS Trust	(**W**)	58	26	16	(RI)	(RI)
Royal Liverpool and Broadgreen University Hospitals NHS Trust	(**W**)	62	24	15	(G)	(G)
Homerton University Hospital NHS Foundation Trust	(**W**)	60	26	14	(G)	(O)
East Sussex Healthcare NHS Trust	(**W**)	59	26	14	(RI)	(RI)
Blackpool Teaching Hospitals NHS Foundation Trust	(**W**)	59	26	15	(RI)	(RI)
Imperial College Healthcare NHS Trust	(**W**)	59	26	15	(RI)	(RI)

Key:

Trust performance	About the same (**S**)	Worse (**W**)	Much worse (**MW**)	
CQC rating	Inadequate (**I**)	Requires improvement (**RI**)	Good (**G**)	Outstanding (**O**)

Core service rating for: urgent and emergency services

Trusts achieving ‘much worse than expected’ results

One trust was classed as ‘much worse than expected’ and has also been rated by CQC as requires improvement.

		Overall results (%)				
	2016	Most Positive (10/10)	Middle	Most Negative (0/10)	Overall CQC rating	Core service rating
Trust average		**66**	**23**	**11**		
The Hillingdon Hospitals NHS Foundation Trust	(**MW**)	56	27	17	**RI**	**RI**

Key:

Trust performance	About the same (**S**)	Worse (**W**)	Much worse (**MW**)	
CQC rating	Inadequate (**I**)	Requires improvement (**RI**)	Good (**G**)	Outstanding (**O**)

Core service rating for: urgent and emergency services

Further information

The results for England and trust level results are available on CQC's website. There is also a technical document that describes the methodology for analysing the trust level benchmark results, and a Quality and Methodology report that discusses methodological issues.

www.cqc.org.uk/emergencydepartmentsurvey

Trusts' results from previous emergency department surveys are available at the link below. However, please note that results from the 2016 survey are **not comparable** with previous surveys. For more information on this please see the statistical release or the Quality and Methodology report.

www.nhssurveys.org/surveys/296

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:

www.nhssurveys.org/surveys/957

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/content/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 patient experience was 'worse than expected' uses responses for all scored questions (except the overall experience question asked to parents).[5]

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 indicates a poor patient experience.

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

We use two thresholds to flag trusts that have a concerning level 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

In order to identify 'better than expected' patient experience, a count of the number of responses scored as '10' (the most positive option) is calculated 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 also 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

5. Overall experience is not included in the analysis because of the ambiguity around what should be classed as the 'most negative' (and 'most positive') option(s).

Standardisation

As in the benchmark results for each trust, 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 the profile of its respondents.

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 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 Q33.

Q33. In your opinion, how clean was the emergency department?

Very clean	10
Fairly clean	6.7
Not very clean	3.3
Not at all clean	0
Can't say	Not applicable

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 the trusts. For example, they may be descriptive questions such as asking respondents if they had a test while in the emergency department.

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 respondent[6]

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 within 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, 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.

6. Compute the z-score for the proportion

The Z-score formula used is:

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

6. The analytical approach used to identify positive patient experience 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 patient experience.

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:

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

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

9. Adjust for overdispersion

If I $\hat{\phi}$ 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:

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

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

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

Appendix C: overall experience questions – 'better' and 'worse' trusts

The trusts listed in this appendix were identified using the benchmark data and are trusts that were categorised as 'better' or 'worse' than expected in the benchmark reports for one 'Overall experience' question.[7]

Six trusts were identified as being **'better than expected'** for the overall experience question (Q45):[8]

- Harrogate and District NHS Foundation Trust
- The Queen Elizabeth Hospital, King's Lynn, NHS Foundation Trust
- Poole Hospital NHS Foundation Trust
- Salford Royal NHS Foundation Trust
- Salisbury NHS Foundation Trust
- Gateshead Health NHS Foundation Trust.

Five trusts were identified as being **'worse than expected'** for the overall experience question (Q45):

- Bradford Teaching Hospitals NHS Foundation Trust
- The Hillingdon Hospitals NHS Foundation Trust
- Barking, Havering and Redbridge University Hospitals NHS Trust
- The Princess Alexandra Hospital NHS Trust
- East Sussex Healthcare NHS Trust.

7. We were unable to include the overall experience in the analysis, as used elsewhere in this report, because of the ambiguity around what should be classed as the 'most negative' and 'most positive' option(s).

8. Q45 asks respondents to rate their overall experience on a scale of 0-10 where zero is 'I had a very poor experience' and 10 is 'I had a very good experience'.

How to contact us

Call us on: 03000 616161

Email us at: enquiries@cqc.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