2018 Adult Inpatient Survey Benchmark Reports: Q&A

This document contains answers to questions you may have about the trust-level results for the 2018 adult inpatient survey, as provided in the benchmark reports, and on the CQC website. A technical document, with further detail on the statistical techniques used to categorise trust scores, is also available on the CQC website. In addition, a quality and methodology report outlines further details of the survey. Both documents can be found here: **www.cqc.org.uk/inpatientsurvey**

Questions and Answers

The Benchmark Reports 2
What are the orange, grey and green sections in the chart? 2
How do I know which category my trust's score is in if the diamond representing the score appears to be on the threshold in the benchmark charts? 2
How do I refer to these scores and categories when reporting on the results for my trust? 2
About the Scores 2
Why are the percentage results for all trusts not provided? 2
Why are the scores presented out of ten? 3
How are the scores calculated? 3
Why aren't all questions scored? 3
About the Analysis 3
What is the 'expected range'? 3
Why is the data standardised by the age, gender and route of admission of respondents? 4
Why are there no confidence intervals surrounding the score? 4
Understanding the Data 4
Why do most trusts appear to be performing 'about the same'? 4
Why does the number of trusts performing 'better' or 'worse' at each question vary? 5
Why has no trust come out as performing better or worse for a particular question? 5
Is the lowest scoring trust the worst trust in the country, for each question? And likewise the highest scoring trust the best? 5
The score for one of our questions has gone up but is categorised as 'about the same' yet in the 2017 survey we were 'better'? 5
We are categorised as 'about the same' for a question yet a trust with a slightly lower score than us is categorised as 'better'. Why is this? 6
Why is one of our sections categorised as 'worse' yet all of the questions that fall into that section are 'about the same'? 6
How do I calculate an overall score for my trust? 6
Why do the results and / or number of respondents provided by CQC differ from those provided to me by our approved contractor? 6
Comparing Results 7
Why is statistical significance relevant? 7
Which trusts are performing best / worst? 7
How can I make comparisons to previous years' survey data, or to other trusts? 8
Why can't I sort the scores for all trusts and rank the trusts in order of performance? 8
Further Questions 9

The Benchmark Reports

What are the orange, grey and green sections in the chart?

The coloured bars represent the full range of scores across all trusts, from the lowest score achieved by a trust to the highest. The grey section in the charts represents the **expected range** for a score for a trust. This is the range within which we would expect a particular trust to score if it performed 'about the same' as most other trusts in the survey. If a score falls above or below the expected range it will be in the 'better' or 'worse' category, represented by green and orange areas respectively.

The calculation of the expected range takes into account the number of respondents from each trust as well as the scores for all other trusts, and allows us to identify which scores we can confidently say are 'better' or 'worse' than the majority of other trusts (see the technical guidance for more details, available from: **www.cqc.org.uk/inpatientsurvey** and sent to survey trust leads prior to publication).

How do I know which category my trust's score is in if the diamond representing the score appears to be on the threshold in the benchmark charts?

Text to the right of the graphs clearly states if a trust score for a particular question, or section, is 'better' or 'worse' compared with most other trusts that took part in the survey. If there is no text present, the result is 'about the same'.

How do I refer to these scores and categories when reporting on the results for my trust?

We have produced a brief guide on how to refer to the findings when disseminating the scored data. This was provided to trust survey leads prior to publication, and is available on request from the surveys team at: **patient.survey@cqc.org.uk.**

About the Scores

Why are the percentage results for all trusts not provided?

The percentage data—the number and percentage of respondents who selected each response option for each question—are provided to trusts for their own information only and can be used to understand the results for individual trusts.

Percentage data are not suitable for making comparisons between trusts. This is because the results are not standardised, which means that differences in the profiles of respondents are not taken into account. Any differences across trusts in non-standardised data may be due to differences in the characteristics of respondents.
Instead, we publish scored benchmark data for all trust. This data is adjusted for the demographic differences to make comparisons fairer across trusts with differing population profiles. For the adult inpatient survey, we standardise by age, gender and way of admission (emergency or planned).

A further advantage of using scored data is that it allows for all response options to be taken into account. For example, if you were to look only at the 'yes definitely' responses displayed in the table below, you might conclude that trust A and trust B are performing similarly. However, considering the other responses, it is apparent that trust B has the more positive survey results overall.

8: In your opinion, had the specialist you saw in hospital been given all the necessary information about your condition or illness from the person who referred you?

	Trust A	Trust B
Yes, definitely	59%	59%
Yes, to some extent	10%	39%
No	31%	2%

Scored and standardised survey data is therefore considered to be the fairest data to include in CQC's regulatory activities, as well as the measures and assessments of other stakeholders, such as NHS England and the Department of Health and Social Care.

In the past, percentage results or scores were used to present data in a league table form, or to identify the 'better' or 'worse' trusts. However, such use and interpretations of survey data, without significance tests, would likely be misleading and inaccurate.

Why are the scores presented out of ten?

The scores are presented out of ten to emphasise that they are scores and not percentages.

How are the scores calculated?

For each question in the survey that can be scored (see section below), the **standardised** individual responses are converted into scores on a scale of 0 to 10. A score of 10 is assigned to the most positive response and a score of 0 to the least positive. The higher the score for each question, the better the trust is performing. For more detailed information on the methodology, including the scores assigned to each question, please see the technical document.

Why aren't all questions scored?

Some questions are not scored because they do not evaluate trust performance. For example, some are 'routing questions', which are designed to filter out respondents for whom subsequent questions do not apply. For example, Question 1 'Was your most recent hospital stay planned in advance or an emergency?', is a routing question. Those who report a planned admission are instructed not to answer Questions 2, 3 and 4. Other questions that are not scored are descriptive, for example, Question 11 'While in hospital, did you ever stay in a critical care area (e.g. Intensive Care Unit, High Dependency Unit or Coronary Care Unit)?' or demographic questions such as Question 79 'What is your religion?'.

About the Analysis

What is the 'expected range'?

The better / about the same / worse categories are based on a statistic called the 'expected range', which is calculated for each question for each trust. This is the range within which we would expect a particular trust to score if it performed about the same as most other trusts in the survey. The range takes into account the number of respondents from each trust as well as the distribution of scores for all other trusts, and allows us to identify which scores we can confidently say are 'better' or 'worse' than the majority of other trusts. Analysing the survey information in such a way allows for fairer conclusions

to be made in terms of each trust's performance. This approach presents the findings in a way that takes account of all necessary factors, yet is presented in a simple manner.

If a trust is categorised as 'better' or 'worse' than average then we can be very confident that it would continue to appear better or worse than average if the survey was repeated with a different sample. More detail on the calculation of the expected range is available in the technical document.

Why is the data standardised by the age, gender and route of admission of respondents?

The reason for 'standardising' data is that we know that the views of a respondent can reflect not only their experience of NHS services, but can also relate to certain demographic characteristics, such as their age. For example, older respondents tend to report more positive experiences than younger patients. Because the demographic profile of inpatients varies across trusts (for example one trust may serve a considerably older population than another), this could potentially lead to the result for a trust appearing better or worse than they would if they had a slightly different service-user profile. To account for this we 'standardise' the data. Standardising data adjusts for these differences and enables the results for trusts with different population profiles to be compared more fairly than could be achieved using non-standardised data.

Why are there no confidence intervals surrounding the score?

As the 'expected range' calculation takes into account the number of respondents at each trust who answer a question, as well as the scores for all other trusts, it is not necessary to present confidence intervals around each score.

Understanding the Data

Why do most trusts appear to be performing 'about the same'?

The expected range is a conservative statistic. It accounts for the possibility that there is variation across trusts for other reasons, aside from differences in trust performance. There may be significant variation between trusts due to certain factors that are not within the trusts' control. The technique used takes this into account, and so if a trust is found to be performing 'better' or 'worse' compared with most other trusts that took part in the survey, this is extremely unlikely to have occurred by chance.

Even though your trust may appear to be performing 'about the same' compared to most other trusts in England, the results should still be useful to you locally, for example you may want to:

- Make comparisons to the results from previous surveys to look for questions where results have improved or declined. Tables in the benchmark reports for your trust identify any statistically significant changes from the 2017 survey.
- Identify particular areas you may wish to improve on ahead of the next survey.
- Compare your results with those of other similar trusts.
- Look at your results by different service user groups to understand their different experiences, for example, by age, long term condition, ethnic group, etc. The data necessary to produce extra analysis can be obtained from CQC upon request.

- Undertake follow up activity with service users such as interviews, workshops or focus groups to get more in depth information into areas in which you would like to improve.

Please remember that for points 2-3 above, to do this accurately you should undertake an appropriate **significance test.**

Why does the number of trusts performing 'better' or 'worse' at each question vary?

It is important to be aware that the range of performance on questions varies. This variability influences how much a trust's score needs to differ from the average to be considered 'better' or 'worse' than the average. This means that the number of trusts performing 'better' or 'worse' on each question will vary.

Why has no trust come out as performing better or worse for a particular question?

This can occur in the analysis of the data and is an acceptable consequence of the statistical technique that is used. The size of the expected range is constructed by considering how different all trust scores are across the range, as well as the confidence we can have in that particular trust's score (by looking at the number of respondents to that question). In some cases, this will lead to such a wide margin of error that the 'expected range' will be very wide, and hence will also cover the highest and / or lowest scoring trusts for that question.

Is the lowest scoring trust the worst trust in the country, for each question? And likewise the highest scoring trust the best?

If a trust is in the 'better' or 'worst' category this mean that they are performing either better or worse compared with *most other trusts* that took part in the survey. However, a trust is not necessarily *the best*, or *the wors*t, and this could not be determined without undertaking an appropriate significance test.

If another sample of people who use services were surveyed, and we were to order scores again, we would likely find that different trusts would have the highest and lowest scores. This is because the scores are estimates: we only received questionnaires from some of the patients who used services during the sampling period (July 2018). However, by analysing the data in the way that we have, we can say which trusts are likely to be consistently 'better' and those which are likely to be consistently 'worse'. These groups should be viewed as a group of 'better' trusts and 'worse' trusts. This is the fairest way to present the data. It means that individual trusts are not singled out as the very 'best' or very 'worst', when they might not have been if all patients who used acute services during the sampling period had been surveyed.

The score for one of our questions has gone up but is categorised as 'about the same' yet in the 2017 survey we were 'better'?

When looking at scores within a trust over time, it is important to be aware that they are relative to the performance of other trusts. If, for example, a trust was 'better' for one question, then 'about the same' the next time the survey was carried out, it may not indicate an actual decrease in the performance of the trust, but instead may be due to an improvement in many other trusts' scores, leaving the trust to appear more 'average'. Hence, when comparing your trust's result overtime, it is more useful to look at actual changes in scores over time rather than banding.

We are categorised as 'about the same' for a question yet a trust with a slightly lower score than us is categorised as 'better'. Why is this?

The 'expected range' calculation takes into account the number of respondents from each trust as well as the distribution of scores for all other trusts, and allows us to identify which scores we can confidently say are 'better' or 'worse' than the majority of other trusts. As set out above, the expected range is a conservative statistic: it accounts for the possibility that there is variation across trusts for other reasons, aside from differences in trust performance. There may be significant variation between trusts due to certain factors that are not within the trusts' control. The technique used takes this into account. It is likely that your trust came out as 'about the same' because your trust had fewer respondents to the question which creates a greater degree of uncertainty around the result. The trust with the lower score would likely have had more respondents to the question, and so their expected range would have been narrower.

Why is one of our sections categorised as 'worse' yet all of the questions that fall into that section are 'about the same'?

This can happen because the calculation of the section scores is a separate calculation and not an average of all questions that make up a particular section. If this has occurred, it is likely that your trust scored very lowly or even on the threshold for all or most of the questions that are in a section.

The thresholds for 'worse', 'about the same' and 'better' are based on the score variance. For sections, this is a composite of the separate question variances, but not a straightforward sum, because it also depends on the correlation between questions. It does not therefore follow that a trust that is above the threshold on separate questions will also be above the threshold when those questions are combined.

The 'expected range' is dependent on the (sampling) variance of the trust's results – with a more reliable score (as would normally be the case for section scores), it is easier to be significantly different from the 'average' group than for a less reliable score.

How do I calculate an overall score for my trust?

It is important to remember that there is no overall indicator or figure for 'patient experience', so it is not accurate to say that a trust is the 'best in the country' or 'best in the region' *overall*. Similarly, adding up the number of 'better' and 'worse' categories to find out which trust did better or worse overall would be misleading and simplistic.

The survey assesses a number of different aspects of patient's experience (such as the staff, care in hospital) and trust performance varies across these different aspects. This means that it is not recommended to compare trusts overall. It is better to look at the trusts that are similar to yours, or particular trusts against which you want to compare yourself, and see how they perform across the particular aspects that are of interest to you.

Why do the results and / or number of respondents provided by CQC differ from those provided to me by our approved contractor?

CQC do not see the reports provided to you by your approved contractor and therefore cannot comment on these. You should raise any queries directly with your approved contractor. However, likely reasons for any discrepancies are:

- The approved contractor may have cleaned the data differently to CQC. In particular, CQC remove respondents from the base of a question that do not

analyse the performance of a trust - we refer to these as 'nonspecific responses', such as 'don't know or can't remember'. A guide to data cleaning is available here: **https://nhssurveys.org/surveys/1361**

- Trust level data published by CQC has been 'standardised' by age, gender and route of admission to enable fairer comparisons between the results of trusts which may have different population profiles. Approved Contractors may not have done this or may have applied a different standardisation. To be able to standardise the data, information is needed on both age and gender, if either of these pieces of information are missing, or not able to be determined, the respondent must be dropped from the analysis as it is not possible to apply a weight.
- CQC analyses trust level data by scoring (and standardising) the responses to each question. Each response option that evaluates performance is scored on a scale of 0-10. Approved Contractors may have analysed and / or scored the data in a different way.
- The Approved Contractor will not be able to make comparisons against all trusts that took part in the survey, only against those that commissioned them. Therefore any 'national' results they publish will not be based on all trusts and any thresholds they calculate may be different.

Comparing Results

Why is statistical significance relevant?

Survey scores are estimates – we have only received questionnaires from some patients who used services during the sampling period, not all patients, as the survey uses a sample of patients from a chosen month (July) and some choose not to respond. If another sample of service users were surveyed, you may find the results would change slightly. This is why it is important to test results for statistical significance.

A statistically significant difference means that the change in the results is very unlikely to have occurred by chance. Without significance testing you cannot be sure that a difference between two results would still be different if you repeated the survey again with a different sample. If a result is not significant then you cannot be sure of its accuracy. If a significant difference is present then it is likely that it is a true difference, and if the survey was repeated again that you would see the same outcome.

Which trusts are performing best / worst?

With the analysis technique used, all we can say is that a particular trust's results on a particular question or section of the questionnaire are 'significantly worse or 'significantly better' *than most other trusts*.

We cannot say, for example, that that a trust that has a score of 4.5 (Trust A) is any better than a trust with a score of 4.3 (Trust B). To do so we would need to carry out a statistical test to determine whether this difference is statistically significant. If a difference is not significant, to say one trust is better than another may be unfair and inaccurate.

The trust outliers report identifies which trusts were 'much worse', 'worse', 'about the same', 'better' or 'much better' when analysing all scored questions simultaneously. The results within this separate report use a different technique to the benchmarking analysis, focussing on identifying significantly higher levels of better or worse experience across the entire survey. Full details of the analytical methodology are provided within the report, which will be available at: **www.cqc.org.uk/inpatientsurvey.**

How can I make comparisons to previous years' survey data, or to other trusts?

The purpose of the expected range is to arrive at a judgement of a how a trust is performing compared with all other trusts that took part in the survey. To use the data in another way – for example, to make comparisons to scores achieved in previous surveys, or between trusts, you will need to undertake an appropriate statistical test to ensure that any change is statistically significant. A statistically significant change means that you can be very confident that the change is real and not due to chance.

The previously published results for the inpatient surveys are available here:
http://nhssurveys.org/surveys/425

For advice on making accurate comparisons you may like to speak to someone within your trust with statistical expertise, or your approved contractor (if used) should be able to advise on this.

Why can't I sort the scores for all trusts and rank the trusts in order of performance?

It is not appropriate to sort the scores by size for two reasons:

1) Firstly, due to the analysis technique applied, where the number of respondents is taken into account, it is possible that one trust may score higher than another - though the higher scoring trust is classed as 'about the same' and the second, lower scoring, trust is put into the 'better' category. This may occur if the second trust has a considerably larger number of respondents, as it will be assumed that their score is more reliable, and hence more likely always to be high if the survey was repeated.

2) Secondly, the statistical technique does not measure how different individual trust scores are from one another (whether statistically significant), and so it would be too simple to attempt to sort by scores alone, without running more analysis on the data. The banding technique used is helpful in identifying which trusts are likely always to be in the 'better', 'worse', or 'about the same' category, no matter how many surveys are sent out.

Further information

The full national results for the 2018 survey are on the CQC website, together with an A to Z list to view the results for each trusts inpatient questions, and the technical document outlining the methodology and the scoring applied to each question:
www.cqc.org.uk/inpatientsurvey

The results for the all the previous surveys can be found on the NHS surveys website at:
http://nhssurveys.org/surveys/425

Full details of the methodology for the survey can be found at:
www.nhssurveys.org/surveys/1203

More information on the programme of NHS patient surveys is available at:
www.cqc.org.uk/surveys

More information on how CQC use data to monitor trusts that provide acute services is available at:
www.cqc.org.uk/what-we-do/how-we-use-information/using-data-monitor-services

Further Questions

If you have any further questions please contact the surveys team at CQC:
patient.survey@cqc.org.uk

CQC Surveys team
June 2019