2018 Community Mental Health Survey
Benchmark Reports: Q&A

This document is provided to answer some of the questions you may have on the trust level results, as provided in the benchmark reports, and on the CQC website. A technical document is also available on the CQC website which has further detail on the statistical techniques used to categorise trust scores, and a Quality and Methodology report outlines further details on the survey. Both documents can be found here:
www.cqc.org.uk/cmhsurvey

Questions and Answers

The Benchmark Reports

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

The coloured bars represent the full range of all trust scores, 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 document for more details, available from: **www.cqc.org.uk/cmhsurvey** and sent to trust survey 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 is provided to trusts for their own information only and can be used to understand the results for individual trusts.

Percentage data is not suitable for making comparisons between trusts because the results are not **standardised**,meaning that differences in the profiles of respondents are not taken into account. Any differences across trusts that are shown in non-standardised data may be in part due to differences in the characteristics of respondents. We know that age and gender are two such characteristics that affect perceptions of care and so we adjust the data to account for this to make fairer comparisons across trusts with differing population profiles.

A further advantage of using scored data is that it allows for all response options to be taken into account, rather than looking at just a subset of responses to the question. For example, if you look at the table below, looking at the 'yes definitely' responses only, you would think that trust A and trust B are performing similarly. However, taking into account the other responses, it becomes apparent that trust B has the more positive result overall.

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

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

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

In the past the percentage results or scores have been used to present data in a league table form, or to identify the 'better' or 'worse' trusts. Such use would be misleading and inaccurate, as the differences have not been tested for significance.

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

For example, Q4 was scored as per the example below. The option of 'No' was allocated a score of 0, as this suggests that the respondents' experiences need to be improved. A score of 10 was assigned to the option 'Yes, definitely', as it reflects a positive experience. The remaining option, 'Yes, to some extent', reflects a middling experience, and was assigned a score of 5. The 'don't know / can't remember' response option was not scored as this does not evaluate trusts in any way.

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

Yes, definitely	10
Yes, to some extent	5
No	0
Don't know / can't remember	Not applicable

For more detailed information on the methodology, including the scores assigned to each question, please see the technical document referred to in the 'further information section'.

Why aren't all questions scored?

Not all questions are appropriate for scoring as they do not evaluate trust performance. For example, they may be 'routing questions' designed to filter out respondents for whom the following questions do not apply. For example, Q20 (In the last 12 months have you been receiving any medicines for your mental health needs?) is a routing question. Those who have indicated that they have not been receiving any medicines are instructed not to answer Q21 to Q25, which ask about medication. Other questions that can't be

scored are descriptive, for example, Q1 (When was the last time you saw someone from mental health services?) or demographic questions such as Q43 (Are you male or female?).

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 of service users.

More detail on this is available in this **technical document**.

Why is the data standardised by the age and gender 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 and gender. For example, older respondents tend to report more positive experiences than younger respondents, and women tend to report less positive experiences than do men. Because the mix of people who use services varies across trusts (for example, one trust may serve a considerably older population than another), this could potentially lead to the results for a trust appearing better or worse than they would if they had a slightly different profile of people who use services. 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, you can be confident that this is the case and it is 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 will want to:

- Make comparisons to the results from previous surveys to look for questions where you have improved or declined. Tables in the back of the benchmark report for your trust identify any statistically significant changes from the 2017 survey.
- Compare your results with those of other similar trusts.
- Look at your results by different respondent groups to understand their different experiences, for example, by age, gender, ethnic group, site/location etc.
- Identify particular areas you may wish to improve on ahead of the next survey.
- Undertake follow up activity with people who use your services such as interviews, workshops or focus groups to get more in-depth information into areas in which you would like to improve.
- Review the feedback provided in the 'other comments' section of the questionnaire. If you are using an approved contractor, they may analyse this for you, depending on what you have agreed with them.

Please remember that for points 1-3 above, to do this accurately you should undertake an appropriate **significance test**.

The survey instruction manual provides more information on making use of survey data. The instruction manual is available on the **NHS surveys** website.

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

It is important to be aware that the ranges of performance on different questions varies and this has an influence on how much a trust needs to differ from the average by, in order to be considered 'better' or 'worse' than the average. This means that the number of trusts performing 'better' or 'worse' for 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. This could be because there were few respondents and / or there was a great deal of variation in their answers.

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 means 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 test for statistical significance.

If you took the scores and ordered them by size, you would likely find that the highest and lowest scoring trusts would change if you ran the survey again. This is because the scores are estimates – we only received questionnaires from some of the people who used services during the sampling period (September – November 2017), not all of them. By analysing the data the way we have, we can say which trusts are likely to consistently be 'better' and which are likely to consistently be 'worse' if we were to repeat the survey. As such they should be looked at as a group of 'better' trusts and 'worse' trusts, rather than in order of scores. 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 be if all of the people who used services during the sampling period were surveyed.

The score for one of my 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, it is more useful to look at actual changes in scores over time.

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 the badge category for one of my sections '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 low 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 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?

A separate **Trust outliers report** is available, which details how overall results between trusts vary across the country. This report focuses on identifying significantly higher levels of better or worse patient experience **across the entire survey**, rather than considering performance on individual questions.

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 them. You should raise any queries directly with your approved contractor. However, likely reasons for 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 'non-specific responses' and include options such as such as 'don't know or can't remember'. A guide to data cleaning is available at: **http://www.nhssurveys.org/surveys/1166**
- Trust level data published by CQC has been 'standardised' by age and gender 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 is missing, or not able to be determined, the respondent must be excluded 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 overall 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 only received questionnaires from some of the people who used services during the sampling period (September – November 2017), not everyone. We survey a random sample of 850 service users and some choose not to respond. If another sample of people were surveyed, we might find that the results would be slightly different. It is important to test results for statistical significance for this reason.

A statistically significant difference means that the change in results is very unlikely to have occurred by chance. Without significance testing you cannot be sure that a difference between two results would still occur if you repeated the survey again. If a result is not significant then you cannot be sure of its accuracy.

Which trusts are performing best / worst?

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, available at:**www.cqc.org.uk/cmhsurvey.**

How can I make comparisons to other trusts or previous surveys?

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

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 advice on this.

Which surveys years are comparable?

The results for most questions from the 2018 survey are comparable with the 2014, 2015, 2016 and 2017 survey.

For more details regarding the questionnaire changes please see the 'Development Report for the Community Mental Health Survey 2018' available here:
http://www.nhssurveys.org/surveys/1138

For more information on which questions from 2018 are able to be compared with 2014, 2015, 2016 and 2017, please see appendix A of the statistical release available here:
www.cqc.org.uk/cmhsurvey

Over time there have been a number of changes made to the survey including revisions to the eligible age range and major developments to revise the methodology and the questionnaire content which affect historical comparability:

- The 2004 and 2005 surveys included people aged 16-65 years. In 2006 the age range for the survey was extended to include people aged over 65. In 2012 the minimum age for inclusion was changed to 18. This means that to compare results to earlier surveys it is necessary to exclude older respondents from the analyses.
- The survey underwent two major redevelopments ahead of the 2010 and 2014 surveys to reflect changes in policy, best practice and patterns of service. This means that the 2018 survey is only comparable with the 2017, 2016, 2015 and 2014 surveys. Surveys carried out between 2010 and 2013 are comparable with each other but not with any previous surveys.

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 classed as 'better'. 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 to remain high if the survey was repeated.

2) Secondly, the statistical technique does not measure how different individual trust scores are from one another (i.e. whether statistically significant), and so it would be too simplistic to attempt to sort by scores alone, without running more analyses on the data. The banding technique used is helpful in identifying which trusts are likely to consistently be in the 'better', 'worse', or 'about the same' category if the survey was repeated.

Can I see results for my local clinic / site etc.?

The survey data is presented at trust level only. At present, we are unable to provide data at a more granular level. Some sites may have too few respondents to achieve sufficient numbers of respondents (we set the cut off limit as 30 respondents per organisation). Given that the survey is used by other stakeholders such as NHS England, the Department of Health and Social Care and others to measure trends over time, we are currently unable to change the sampling to accommodate this, without affecting the comparability across years. However, trusts are able to increase their sample size to enable this at a local level. Advice on how to do this is in the **Sampling instructions**.

Further information

The results for England, and trust level results, can be found on the CQC website. You can also find a 'technical document' here which describes the methodology for analysing the trust level results:
www.cqc.org.uk/cmhsurvey

The results from previous community mental health surveys that took place between 2004 and 2008,[1] and between 2010 and 2013 are available at the link below. Please note that due to redevelopment work, results from the 2018 survey are only comparable with 2014, 2015, 2016 and 2017[2]:
www.nhssurveys.org/surveys/290

Full details of the methodology for the survey, including questionnaires, letters sent to people who use services, instructions for trusts and contractors to carry out the survey, and the survey development report, are available at:
http://www.nhssurveys.org/surveys/1114

More information on the NHS Patient Survey Programme, including results from other surveys and a schedule of current and forthcoming surveys can be found at:
www.cqc.org.uk/ surveys

More information on how CQC monitor trusts that provide mental health services is available at:
www.cqc.org.uk/content/monitoring-trusts-provide-mental-health-services

[1] In 2009 a survey of mental health inpatient services took place
[2] Please note that the survey was also substantially redeveloped in 2010. This means that surveys carried out between 2010 and 2013 are comparable with each other but not with any other surveys

Further Questions

We welcome all feedback and questions. Please contact the surveys team at CQC: **patient.survey@cqc.org.uk**.

CQC Surveys team
November 2018