# Care Quality Commission (CQC)

## Technical details – patient survey information
## 2014 Accident and Emergency patient survey
## December 2014

**Contents**

## 1. Introduction

This document outlines the methods used by the Care Quality Commission to score and analyse the results for the 2014 A&E patient survey, as available on the Care Quality Commission website, and in the benchmark report for each trust.

The survey results are available for each trust on the CQC website. The survey data is shown in a simplified way, identifying whether a trust performed 'better' or 'worse' or 'about the same' as the majority of other trusts for each question. This analysis is done using a statistic called the **'expected range'** (see section 5.3). On publication of the survey, an A-to-Z list of trust names will be available at the link below, containing further links to the survey data for all NHS trusts that took part in the survey: **www.cqc.org.uk/accidentandemergency**

The CQC webpage also contains the national results for England, comparing against the results for the previous survey.

Results displayed in the benchmark report for each trust are a graphical representation of the results displayed for the public on the CQC website (see further information section). These have been provided to all trusts and will be available on the survey co-ordination centre website from 2$^{nd}$ December 2014, at: **www.nhssurveys.org**  The tables in the back of the benchmark report also highlight any statistically significant changes in trust scores between 2012 and 2014.


## 2. Selecting data for the reporting

The survey information used and published by the Care Quality Commission consists of a standard set of questions - i.e. questions where results are available from every trust.

Scores are assigned to responses to questions that are of an evaluative nature: in other words, those questions where results can be used to assess the performance of a trust (see section 5 "Scoring individual questions" for more detail). Questions that are not presented in this way tend to be those included solely for 'filtering' respondents past any questions that may not be relevant to them (such as: 'Were you taken to hospital in an ambulance?') or those used for descriptive or information purposes.

The scores for each question are grouped on the website and in the benchmark reports mostly according to the sections of the questionnaire as completed by respondents.  For example, the A&E survey includes sections on 'waiting times', 'doctors and nurses' and 'care and treatment' amongst others. The average score for each trust, for each section, was calculated and will be presented on the website and in the benchmark report for each trust.

Alongside both the question and the section scores on the website are one of three statements:

- Better
- About the same
- Worse

This analysis is done using a statistic called the **'expected range'** (see section 5.3)

### 3. The CQC organisation search tool

The organisation search tool was previously referred to as the Care Directory, and survey data has been displayed in it since 2007. It is intended for a public audience, and contains information from various areas within the Care Quality Commission's functions. The presentation of the survey data was designed using feedback from people who use the data, so that as well as meeting their needs, it presents the groupings of the trust results in a simple and fair way, to show where we are more confident that a trust's score is 'better' or 'worse' than we'd expect, when compared with most other trusts.

The survey data can be found from the A to Z link available at:
**www.cqc.org.uk/accidentandemergency**

Or by searching for a hospital from the CQC home page, then selecting the survey under the 'Surveys' tab.

### 4. The trust benchmark reports

Benchmark reports should be used by NHS trusts to identify how they are performing in relation to all other trusts that took part in the survey. They also show if a score has significantly increased or decreased compared with the last survey. From this, areas for improvement can be identified. The reports are available from the survey co-ordination centre website: **www.nhssurveys.org**

The graphs included in the reports display the scores for a trust, compared with the full range of results from all other trusts that took part in the survey. Each bar represents the range of results for each question across all trusts that took part in the survey. In the graphs, the bar is divided into three sections:

- If a trust score lies in the orange section of the graph, the trust result is 'about the same' as most other trusts in the survey
- If a trust scores lies in the red section of the graph, the trust result is 'worse' than expected when compared with most other trusts in the survey.
- If your a score lies in the green section of the graph, the trust result is 'better' than expected when compared with most other trusts in the survey

A black diamond represents the score for this trust. The black diamond (score) is not shown for questions answered by fewer than 30 people because the uncertainty around the result would be too great.

### 5. Interpreting the data

### 5.1 Scoring

The questions are scored on a scale from 0 to 10. Details of the scoring for this survey are available in Appendix A at the end of this document.

The scores represent the extent to which the patient's experience could be improved. A score of 0 was assigned to all responses that reflect considerable scope for improvement, whereas a response that was assigned a score of 10 referred to the most positive patient experience reported. Where a number of options lay between

the negative and positive responses, they are generally at equal intervals along the scale. Where options were provided that did not have any bearing on the trust's performance in terms of patient experience, the responses were classified as "not applicable" and a score was not given. Where respondents stated they could not remember or did not know the answer to a question, a score was not given.

## 5.2 Standardisation

Results are based on 'standardised' data.  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 sex. For example, older respondents tend to report more positive experiences than younger respondents, and women tend to report less positive experiences than men. Because the mix of patients 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 patients. To account for this we 'standardise' the data. Standardising data adjusts for these differences and enables the results for trusts to be compared more fairly than could be achieved using non-standardised data.

The A&E survey is standardised by age and gender only.

## 5.3 Expected range

The better / about the same / worse categories are based on the 'expected range' that 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 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 Appendix B for more details). 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.

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

## 5.4 Comparing scores across or within trusts, or across survey years

The expected range statistic is used to arrive at a judgement of how a trust is performing compared with all other trusts that took part in the survey. However, if you want to use the scored data in another way, to compare scores (either as trend data for an individual trust or between different trusts) you will need to undertake an appropriate statistical test to ensure that any changes are 'statistically significant'. 'Statistically significant' means that you can be very confident that any change between scores is real and not due to chance.

The benchmark report for each trust includes a comparison to the 2012 survey scores and indicates whether the change is statistically significant. However, to compare back to earlier surveys (where possible) you would need to undertake a similar significance test.

**5.5 Conclusions made on performance**

It should be noted that the data only show performance relative to other trusts: there are no absolute thresholds for 'good' or 'bad' performance. Thus, a trust may score lowly relative to others on a certain question whilst still performing very well on the whole. This is particularly true on questions where the majority of trusts score very highly.

The better / worse categories are intended to help trusts identify areas of good or poor performance. However, 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 following year, 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 accurate to look at actual changes in scores and to test for statistically significant differences.

It is also 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*. Adding up the number of 'better' and 'worse' categories to find out which trust did better or worse overall will be misleading if the limitations are not recognised . The number of questions on each aspect of patient experience in the survey varies, and often so will trusts' performance across these. So if you counted across all of the questions, some topics will have more influence on the overall average than others, when in fact some might not be so important. For more detailed scrutiny of performance, it is more comprehensive to look at trusts individual reports and assess performance by individual questions.


**6. Further information**

The full national results are on the CQC website, together with an A to Z list to view the results for each trust (alongside the technical document outlining the methodology and the scoring applied to each question):
**www.cqc.org.uk/accidentandemergency**

The results for previous A&E surveys can be found at:
**http://www.nhssurveys.org/surveys/296**

Full details of the methodology of the survey can be found at:
**http://www.nhssurveys.org/surveys/738**

More information on the programme of NHS patient surveys is available at:
**www.cqc.org.uk/public/reports-surveys-and-reviews/surveys**

More information on CQC's Intelligent Monitoring is available on the CQC website at:
**http://www.cqc.org.uk/public/hospital-intelligent-monitoring**

**Appendix A: Scoring for the 2014 A&E patient survey**

The following describes the scoring system applied to the evaluative questions in the survey. Taking question 13 as an example (Figure A1), it asks respondents whether doctors and nurses listened to patients. The option of "No" was allocated a score of 0, as this suggests that the experiences of the patient need to be improved. A score of 10 was assigned to the option 'Yes, completely', as it reflects a positive patient experience. The remaining option, 'Yes, to some extent', was assigned a score of 5 as the patient felt their fears were somewhat discussed – either sometimes but not always, or some aspects of their anxieties were discussed but not others. Hence it was placed on the midpoint of the scale.

If the patient did not have any anxieties or fears, this was classified as a 'not applicable' response, as this option was not a direct measure of the explanations that had been given.

**Figure A1 Scoring example:**
**Question 13 (2014 A&E Survey)**

| Q13. If you had any anxieties or fears about your condition or treatment, did a doctor or nurse discuss them with you? | |
|---|---|
| Yes, completely | 10 |
| Yes, to some extent | 5 |
| No | 0 |
| I did not have anxieties or fears | Not applicable |

Where a number of options lay between the negative and positive responses, they were placed at equal intervals along the scale. For example, question 31 asks respondents how clean the A&E Department was, in their opinion (Figure A2). The following response options were provided:

- Very clean
- Fairly clean
- Not very clean
- Not at all clean
- Can't say

A score of 10 was assigned to the option 'Very clean', as this represents best outcome in terms of patient experience. A response that the A&E was 'not at all clean' was given a score of 0.  The remaining two answers were assigned a score that reflected their position in terms of quality of experience, spread evenly across the scale. Hence the option 'fairly clean' was assigned a score of 6.7, and 'not very clean' was given a score of 3.3.

**Figure A2 Scoring example:**
**Question 31 (2014 A&E Survey)**

| Q31. In your opinion, how clean was the A&E Department? | |
|---|---|
| Very clean | 10 |
| Fairly clean | 6.7 |
| Not very clean | 3.3 |
| Not at all clean | 0 |
| Can't say | Not applicable |

Details of the method used to calculate the scores for each trust, for individual questions and each section of the questionnaire, are available in Appendix B. This also includes an explanation of the technique used to identify scores that are better, worse or about the same as most other trusts.

All analysis is carried out on a 'cleaned' data set. 'Cleaning' refers to the editing process that is undertaken on the survey data. A document describing this can be found at: **http://www.nhssurveys.org/surveys/810**

As part of the cleaning process, responses are removed from any trust that has fewer than 30 respondents to a question. This is because the uncertainty around the result is too high, and very low numbers would risk respondents being recognised from their responses.

The below details the scoring allocated to each scorable question.

## **Section 1: Arrival at A&E**

| Q4. Once you arrived at the hospital, how long did you wait with the ambulance crew before your care was handed over to the A&E staff? | |
| --- | --- |
| I did not have to wait | 10 |
| Up to 15 minutes | 10 |
| 16 - 30 minutes | 6.7 |
| 31 - 60 minutes | 3.3 |
| More than 1 hour but no more than 2 hours | 0 |
| More than 2 hours | 0 |
| Don't know/ can't remember | - |

Answered by those who arrived by ambulance

| Q5: Were you given enough privacy when discussing your condition with the receptionist? | |
| --- | --- |
| Yes, definitely | 10 |
| Yes, to some extent | 5 |
| No | 0 |
| I did not discuss my condition with a receptionist | - |

Answered by all

## **Section 2: Waiting times**

| Q6: How long did you wait before you first spoke to a nurse or doctor? | |
| --- | --- |
| 0 -15 minutes | 10 |
| 16-30 minutes | 6.7 |
| 31-60 minutes | 3.3 |
| More than 60 minutes | 0 |
| Don't know/ can't remember | - |

Answered by all

| Q7: From the time you first arrived at the A&E Department, how long did you wait before being examined by a doctor or nurse? | |
| --- | --- |
| I did not have to wait | 10 |
| 1 - 30 minutes | 8 |
| 31 - 60 minutes | 6 |
| More than 1 hour but no more than 2 hours | 4 |
| More than 2 hours but no more than 4 hours | 2 |
| More than 4 hours | 0 |
| Can't remember | - |
| I did not see a doctor or a nurse | - |

Answered by all

| Q8: Were you told how long you would have to wait to be examined? | |
| --- | --- |
| Yes, but the wait was shorter | 10 |
| Yes, and I had to wait about as long as I was told | 10 |
| Yes, but the wait was longer | 5 |
| No, I was not told | 0 |
| Don't know/ can't remember | - |

Answered by those who waited to see a doctor or nurse

| Q9: Overall, how long did your visit to the A&E Department last? | |
| --- | --- |
| Up to 1 hour | 10 |
| More than 1 hour but no more than 2 hours | 10 |
| More than 2 hours but no more than 4 hours | 8 |
| More than 4 hours but no more than 6 hours | 6 |
| More than 6 hours but no more than 8 hours | 4 |
| More than 8 hours but no more than 12 hours | 2 |
| More than 12 hours but no more than 24 hours | 0 |
| More than 24 hours | 0 |
| Can't remember | - |

Answered by all

## Section 3: Doctors and nurses

| Q10: Did you have enough time to discuss your health or medical problem with the doctor or nurse? | |
| --- | --- |
| Yes, definitely | 10 |
| Yes, to some extent | 5 |
| No | 0 |
| I did not see a doctor or nurse | - |

Answered by all

| Q11: While you were in the A&E Department, did a doctor or nurse explain your condition and treatment in a way you could understand? | |
|---|---|
| Yes, completely | 10 |
| Yes, to some extent | 5 |
| No | 0 |
| I did not need an explanation | - |

Answered by those who saw a doctor or nurse

| Q12: Did the doctors and nurses listen to what you had to say? | |
|---|---|
| Yes, definitely | 10 |
| Yes, to some extent | 5 |
| No | 0 |

Answered by those who saw a doctor or nurse

| Q13: If you had any anxieties or fears about your condition or treatment, did a doctor or nurse discuss them with you? | |
|---|---|
| Yes, completely | 10 |
| Yes, to some extent | 5 |
| No | 0 |
| I did not have anxieties or fears | - |

Answered by those who saw a doctor or nurse

| Q14: Did you have confidence and trust in the doctors and nurses examining and treating you? | |
|---|---|
| Yes, definitely | 10 |
| Yes, to some extent | 5 |
| No | 0 |

Answered by those who saw a doctor or nurse

| Q15: Did doctors or nurses talk to each other about you as if you weren't there? | |
|---|---|
| Yes, definitely | 0 |
| Yes, to some extent | 5 |
| No | 10 |

Answered by those who saw a doctor or nurse

| Q16: If your family or someone else close to you wanted to talk to a doctor, did they have enough opportunity to do so? | |
|---|---|
| Yes, definitely | 10 |
| Yes, to some extent | 5 |
| No | 0 |
| No family or friends were involved | - |
| My family or friends did not want or need information | - |
| I did not want my family or friends to talk to a doctor | - |

Answered by those who saw a doctor or nurse

## Section 4: Care and treatment

| Q17: While you were in the A&E Department, how much information about your condition or treatment was given to you? | |
|---|---|
| Not enough | 5 |
| Right amount | 10 |
| Too much | 5 |
| I was not given any information about my condition or treatment | 0 |

Answered by all

| Q18: Were you given enough privacy when being examined or treated? | |
|---|---|
| Yes, definitely | 0 |
| Yes, to some extent | 5 |
| No | 10 |

Answered by all

| Q19: If you needed attention, were you able to get a member of medical or nursing staff to help you? | |
|---|---|
| Yes, always | 10 |
| Yes, sometimes | 5 |
| No, I could not find a member of staff to help me | 0 |
| A member of staff was with me all the time | 10 |
| I did not need attention | - |

Answered by all

| Q20: Sometimes in a hospital, a member of staff will say one thing and another will say something quite different. Did this happen to you in the A&E Department? | |
|---|---|
| Yes, definitely | 0 |
| Yes, to some extent | 5 |
| No | 10 |

Answered by all

| Q21: Were you involved as much as you wanted to be in decisions about your care and treatment? | |
|---|---|
| Yes, definitely | 10 |
| Yes, to some extent | 5 |
| No | 0 |
| I was not well enough to be involved in decisions about my care | - |

Answered by all

| Q22: If you were feeling distressed while you were in the A&E department, did a member of staff help to reassure you? | |
|---|---|
| Yes, definitely | 10 |
| Yes, to some extent | 5 |
| No | 0 |
| I was not distressed | - |
| Not sure/ can't remember | - |

Answered by all

## Section 5: Tests

| Q24: Did a member of staff explain why you needed these test(s) in a way you could understand? | |
|---|---|
| Yes, completely | 10 |
| Yes, to some extent | 5 |
| No | 0 |

Answered by those who had tests

| Q26: Did a member of staff explain the results of the tests in a way you could understand? | |
|---|---|
| Yes, definitely | 10 |
| Yes, to some extent | 5 |
| No | 0 |
| Not sure/ can't remember | - |

Answered by those who received their test results before they left A&E

## Section 6: Pain

| Q29: How many minutes after you requested pain relief medication did it take before you got it? | |
|---|---|
| 0 minutes / right away | 10 |
| 1 - 5 minutes | 10 |
| 6 - 10 minutes | 7.5 |
| 11 - 15 minutes | 5 |
| 16 - 30 minutes | 2.5 |
| More than 30 minutes | 0 |
| I asked for pain relief medication but wasn't given any | 0 |

Answered by those who were in pain and requested pain relief

| Q30: Do you think the hospital staff did everything they could to help control your pain? | |
|---|---|
| Yes, definitely | 10 |
| Yes, to some extent | 5 |
| No | 0 |
| Can't say/ don't know | - |

Answered by those who were in pain

## Section 6: Hospital environment and facilities

| Q31: In your opinion, how clean was the A&E Department? | |
|---|---|
| Very clean | 10 |
| Fairly clean | 6.7 |
| Not very clean | 3.3 |
| Not at all clean | 0 |
| Can't say | - |

Answered by all

| Q32: While you were in the A&E Department, did you feel threatened by other patients or visitors? | |
|---|---|
| Yes, definitely | 0 |
| Yes, to some extent | 5 |
| No | 10 |

Answered by all

| Q33: Were you able to get suitable food or drinks when you were in the A&E Department? | |
|---|---|
| Yes | 10 |
| No | 0 |
| I was told not to eat or drink | 10 |
| I did not know if I was allowed to eat or drink | 0 |
| I did not want anything to eat or drink | - |

Answered by all

## Section 7: Leaving A&E

| Q36: Did a member of staff explain the purpose of the medications you were to take at home in a way you could understand? | |
|---|---|
| Yes, completely | 10 |
| Yes, to some extent | 5 |
| No | 0 |
| I did not need an explanation | - |

Answered by those who were not admitted to hospital or a nursing home and who were prescribed medication

| Q37: Did a member of staff tell you about medication side effects to watch for? | |
|---|---|
| Yes, completely | 10 |
| Yes, to some extent | 5 |
| No | 0 |
| I did not need this type of information | - |

Answered by those who were not admitted to hospital or a nursing home and who were prescribed medication

| Q38: Did a member of staff tell you when you could resume your usual activities, such as when to go back to work or drive a car? | |
|---|---|
| Yes, definitely | 10 |
| Yes, to some extent | 5 |
| No | 0 |
| I did not need this type of information | - |

Answered by those who were not admitted to hospital or a nursing home

| Q40: Did a member of staff tell you about what danger signals regarding your illness or treatment to watch for after you went home? | |
|---|---|
| Yes, completely | 10 |
| Yes, to some extent | 5 |
| No | 0 |
| I did not need this type of information | - |

Answered by those who were not admitted to hospital or a nursing home

| Q41: Did hospital staff tell you who to contact if you were worried about your condition or treatment after you left the A&E Department? | |
|---|---|
| Yes | 10 |
| No | 0 |
| Don't know/ can't remember | - |

Answered by those who were not admitted to hospital or a nursing home

## Section 8: Experience overall

| Q42: Overall, did you feel you were treated with respect and dignity while you were in the A&E Department? | |
|---|---|
| Yes, all of the time | 10 |
| Yes, some of the time | 5 |
| No | 0 |

Answered by all

| Q43: Overall... | |
|---|---|
| 0 (I had a very poor experience) | 0 |
| 1 | 1 |
| 2 | 2 |
| 3 | 3 |
| 4 | 4 |
| 5 | 5 |
| 6 | 6 |
| 7 | 7 |
| 8 | 8 |
| 9 | 9 |
| 10 (I had a very good experience) | 10 |

Answered by all

**Appendix B: Calculating the trust score and category**

**Calculating trust scores**

The scores for each question and section in each trust were calculated using the method described below.

Weights were calculated to adjust for any variation between trusts that resulted from differences in the age and sex of respondents.  A weight was calculated for each respondent by dividing the national proportion of respondents in their age/sex group by the corresponding trust proportion. The reason for weighting the data was that younger people and women tend to be more critical in their responses than older people and men. If a trust had a large population of young people or women, their performance might be judged more harshly than if there was a more consistent distribution of respondents by age and sex.

**Weighting survey responses**

The first stage of the analysis involved calculating national age/ sex proportions. It must be noted that the term "national proportion" is used loosely here as it was obtained from pooling the survey data from all trusts, and was therefore based on the respondent population rather than the entire population of England.

The questionnaire asked respondents to state their year of birth. The approximate age of each patient was then calculated by subtracting the figure given from 2014. The respondents were then grouped according to the categories shown in Figure B1.

If a patient did not fill in their year of birth or sex on the questionnaire, this information was inputted from the sample file. If information on a respondent's age and/or sex was missing from both the questionnaire and the sample file, the patient was excluded from the analysis.

The national age/sex proportions relate to the proportion of men and women of different age groups. As shown in Figure B1, the proportion of respondents who were male, and aged 51 to 65 years is 0.116; the proportion who were women and aged 51 to 65 years is 0.126, etc.

**Figure B1 National Proportions**

| Sex | Age Group | National proportion 2014 |
|---|---|---|
| Men | ≤35 | 0.056 |
| | 36-50 | 0.073 |
| | **51-65** | **0.116** |
| | 66+ | 0.207 |
| Women | ≤35 | 0.091 |
| | 36-50 | 0.092 |
| | **51-65** | **0.126** |
| | 66+ | 0.239 |

These proportions were calculated for each trust, using the same procedure.

The next step was to calculate the weighting for each individual. Age/sex weightings were calculated for each respondent by dividing the national proportion of respondents in their age/sex group by the corresponding trust proportion.

If, for example, a lower proportion of men aged between 51 and 65 years within Trust A responded to the survey, in comparison with the national proportion, then this group would be under-represented in the final scores. Dividing the national proportion by the trust proportion results in a weighting greater than "1" for members of this group (Figure B2). This increases the influence of responses made by respondents within that group in the final score, thus counteracting the low representation.

**Figure B2 Proportion and Weighting for Trust A**

| Sex | Age Group | National Proportion | Trust A Proportion | Trust A Weight (National/Trust A) |
|---|---|---|---|---|
| Men | ≤35 | 0.056 | 0.036 | 1.556 |
| | 36-50 | 0.073 | 0.071 | 1.028 |
| | **51-65** | **0.116** | **0.094** | **1.234** |
| | 66+ | 0.207 | 0.189 | 1.095 |
| Women | ≤35 | 0.091 | 0.092 | 0.989 |
| | 36-50 | 0.092 | 0.114 | 0.807 |
| | 51-65 | **0.126** | 0.168 | 0.750 |
| | 66+ | 0.239 | 0.236 | 1.013 |
| | 36-50 | 0.056 | 0.036 | 1.556 |
| | 51-65 | 0.073 | 0.071 | 1.028 |
| | 66+ | **0.116** | 0.094 | 1.234 |

Likewise, if a considerably higher proportion of women aged between 36 and 50 years from Trust B responded to the survey (Figure B3), then this group would be over-represented within the sample, compared with national representation of this group. Subsequently this group would have a greater influence over the final score. To counteract this, dividing the national proportion by the proportion for Trust B, results in a weighting of less than one for this group.

**Figure B3 Proportion and Weighting for Trust B**

| Sex | Age Group | National Proportion | Trust B Proportion | Trust B Weight (National/Trust B) |
|---|---|---|---|---|
| Men | ≤35 | 0.056 | 0.032 | 1.750 |
| | 36-50 | 0.073 | 0.058 | 1.259 |
| | 51-65 | **0.116** | 0.124 | 0.935 |
| | 66+ | 0.207 | 0.188 | 1.101 |
| Women | ≤35 | 0.091 | 0.068 | 1.338 |
| | 36-50 | 0.092 | 0.207 | 0.444 |
| | 51-65 | **0.126** | 0.112 | 1.125 |
| | 66+ | 0.239 | 0.211 | 1.133 |

To prevent the possibility of excessive weight being given to respondents in an extremely underrepresented group, the maximum value for any weight was set at five.

**Calculating question scores**

The trust score for each question displayed on the website was calculated by applying the weighting for each respondent to the scores allocated to each response.

The responses given by each respondent were entered into a dataset using the 0-10 scale described in section 3. Each row corresponded to an individual respondent, and each column related to a survey question. For those questions that the respondent did not answer (or received a "not applicable" score for), the relevant cell remained empty. Alongside these were the weightings allocated to each respondent (Figure B6).

**Figure B4 Scoring for the 'Arrival at A&E' section, 2014 A&E survey, Trust B**

| Respondent | Scores Q4 | Scores Q5 | Weight |
|---|---|---|---|
| 1 | 10 | 0 | 1.750 |
| 2 | 5 | 10 | 0.935 |
| 3 | . | 5 | 0.444 |

Respondents' scores for each question were then multiplied individually by the relevant weighting, in order to obtain the numerators for the trust scores (Figure B5).

**Figure B5 Numerators for the 'Arrival at A&E' section, 2014 A&E survey, Trust B**

| Respondent | Scores Q4 | Scores Q5 | Weight |
|---|---|---|---|
| 1 | 17.5 | 0 | 1.75 |
| 2 | 4.675 | 9.35 | 0.935 |
| 3 | - | 2.22 | 0.444 |

**the denominators for each domain score**

A second dataset was then created. This contained a column for each question, grouped into domains, and again with each row corresponding to an individual respondent. A value of one was entered for the questions where a response had been given by the respondent, and all questions that had been left unanswered or allocated a scoring of "not applicable" were set to missing (Figure B8).

**Figure B6 Values for non-missing responses, 'Arrival at A&E' section, 2014 A&E survey, Trust B**

| Respondent | Scores | | Weight |
|:---:|:---:|:---:|:---:|
| | **Q4** | **Q5** | |
| 1 | 1 | 1 | 1.75 |
| 2 | 1 | 1 | 0.935 |
| 3 | - | 1 | 0.444 |

The denominators were calculated by multiplying each of the cells within the second dataset by the weighting allocated to each respondent. This resulted in a figure for each question that the respondent had answered (Figure B9). Again, the cells relating to the questions that the respondent did not answer (or received a 'not applicable' score for) remained set to missing (Figure B8).

**Figure B7 Denominators for the 'Arrival at A&E' section, 2014 A&E survey, Trust B**

| Respondent | Score | | Weight |
|:---:|:---:|:---:|:---:|
| | **Q4** | **Q5** | |
| 1 | 1.75 | 1.75 | 1.75 |
| 2 | 0.935 | 0.935 | 0.935 |
| 3 | - | 0.444 | 0.444 |

The weighted mean score for each trust, for each question, was calculated by dividing the sum of the weighted scores for a question (i.e. numerators), by the weighted sum of all eligible respondents to the question (i.e. denominators) for each trust.

Using the example data for Trust B, we first calculated weighted mean scores for the two questions that contributed to the 'arrival at A&E' section of the questionnaire.

Q4:      $\dfrac{17.5 + 4.675}{1.75 + 0.935}$      =      8.259

Q5:      $\dfrac{0.000 + 9.35 + 2.22}{1.75 + 0.935 + 0.444}$   =   3.698

**Calculating section scores**

A simple arithmetic mean of each trust's question scores was then taken to give the score for each section.  Continuing the example from above, then, Trust B's score for the 'Arrival at A&E' section of the A&E survey would be calculated as:

(8.259 + 3.698) / 2 = 5.978

**Calculation of the expected ranges**

Z statistics (or Z scores) are standardized scores derived from normally distributed data, where the value of the Z score translates directly to a p-value. That p-value then translates to what level of confidence you have in saying that a value is significantly different from the mean of your data (or your 'target' value).

A standard Z score for a given item is calculated as:

$$z_i = \frac{y_i - \theta_0}{s_i} \quad (1)$$

where:      $s_i$ is the standard error of the trust score[1],
            $y_i$ is the trust score
            $\theta_0$ is the mean score for all trusts

Under this banding scheme, a trust with a Z score of < -1.96 is labeled as "Worse" (significantly below average; p<0.025 that the trust score is below the national average), -1.96 < Z < 1.96 as "About the same", and Z > 1.96 as "Better" (significantly above average; p<0.025 that the trust score is above the national average) than what would be expected based on the national distribution of trust scores.

However, for measures where there is a high level of precision (the survey indicators sample sizes average around 400 to 500 per trust) in the estimates, the standard Z score may give a disproportionately high number of trusts in the significantly above/ below average bands (because $s_i$ is generally so small). This is compounded by the fact that all the factors that may affect a trust's score cannot be controlled. For example, if trust scores are closely related to economic deprivation then there may be significant variation between trusts due to this factor, not necessarily due to factors within the trusts' control. In this situation, the data are said to be 'over dispersed'. That problem can be partially overcome by the use of an 'additive random effects model' to calculate the Z score (we refer to this modified Z score as the $Z_D$ score). Under that model, we accept that there is natural variation between trust scores, and this variation is then taken into account by adding this to the trust's local

---

[1] Calculated using the method in Appendix C.

standard error in the denominator of (1). In effect, rather than comparing each trust simply to one national target value, we are comparing them to a national distribution.

The $Z_D$ score for each question and section was calculated as the trust score minus the national mean score, divided by the standard error of the trust score plus the variance of the scores between trusts. This method of calculating a $Z_D$ score differs from the standard method of calculating a Z score in that it recognizes that there is likely to be natural variation between trusts which one should expect, and accept. Rather than comparing each trust to one point only (i.e. the national mean score), it compares each trust to a distribution of acceptable scores. This is achieved by adding some of the variance of the scores between trusts to the denominator.

The steps taken to calculate $Z_D$ scores are outlined below.

Winsorising Z-scores
The first step when calculating $Z_D$ is to 'Winsorise' the standard Z scores (from (1)). Winsorising 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=0.2

3. Set the lowest 100q% of Z-scores to $Z_q$, and the highest 100q% of Z-scores to $_{(1-q)}$. These are the Winsorised statistics.

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

**Estimation of over-dispersion**

An over dispersion factor $\hat{\phi}$ is estimated for each indicator which allows us to say if the data for that indicator 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 Winsorised Z scores are used in estimating $\hat{\phi}$.

**An additive random effects model**

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 $\theta_i$ (trust means) 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 $w_i = 1 / s_i^2$ and $\hat{\phi}$ is from (2). Once $\hat{\tau}$ has been estimated, the $Z_D$ score is calculated as:

$$z_i^D = \frac{y_i - \theta_0}{\sqrt{s_i^2 + \hat{\tau}^2}} \qquad (4)$$

**Appendix C: Calculation of standard errors**

**Calculation of standard errors**

In order to calculate statistical bandings from the data, it is necessary for CQC to have both trusts' scores for each question and section and the associated standard error. Since each section is based on an aggregation of question mean scores that are based on question responses, a standard error needs to be calculated using an appropriate methodology.

For the patient experience surveys, the z-scores are scores calculated for section and question scores, which combines relevant questions making up each section into one overall score, and uses the pooled variance of the question scores

**Assumptions and notation**

The following notation will be used in formulae:

$X_{ijk}$      is the score for respondent $j$ in trust $i$ to question $k$

$Q$      is the number of questions within section $d$

$w_{ij}$      is the standardization weight calculated for respondent $j$ in trust $i$

$Y_{ik}$      is the overall trust $i$ score for question $k$

$Y_{id}$      is the overall score for section $d$ for trust $i$

Associated with the subject or respondent is a weight $w_{ij}$ corresponding to how well the respondent's age/sex is represented in the survey compared with the population of interest.

**Calculating mean scores**

Given the notation described above, it follows that the overall score for trust $i$ on question $k$ is given as:

$$Y_{ik} = \frac{\sum_j w_{ij} X_{ijk}}{\sum_j w_{ij}}$$

The overall score for section $d$ for trust $i$ is then the average of the trust-level question means within section $d$. This is given as:

$$Y_{id} = \frac{\sum_{k=1}^{Q} Y_{ikd}}{Q}$$

**Calculating standard errors**

Standard errors are calculated for both sections and questions.

The variance within trust i on question k is given by:

$$\hat{\sigma}_{ik}^2 = \frac{\sum\limits_j w_{ij}\left(X_{ijk} - Y_{ik}\right)^2}{\sum\limits_j w_{ij}}$$

This assumes independence between respondents.

For ease of calculation, and as the sample size is large, we have used the biased estimate for variance.

The variance of the trust level average question score, is then given by:

$$V_{ik} = Var(Y_{ik}) = Var\left(\frac{\sum\limits_j w_{ij} X_{ijk}}{\sum\limits_j w_{ij}}\right)$$

$$= \frac{Var\left(\sum\limits_j w_{ij} X_{ijk}\right)}{\left(\sum\limits_j w_{ij}\right)^2}$$

$$= \frac{\hat{\sigma}_{ik}^2 \sum\limits_j w_{ij}^2}{\left(\sum\limits_j w_{ij}\right)^2}$$

Covariances between pairs of questions (here, *k* and *m*) can be calculated in a similar way:

$$COV_{ik.im.} = Cov(Y_{ik}, Y_{im}) = \frac{\hat{\sigma}_{ikm} \sum\limits_j w_{ij}^2}{\left(\sum\limits_j w_{ij}\right)^2}$$

Where $\hat{\sigma}_{ikm} = \dfrac{\sum\limits_j w_{ij}(X_{ijk} - Y_{ik})(X_{ijm} - Y_{im})}{\sum\limits_j w_{ij}}$

Note: $w_{ij}$ is set to zero in cases where patient *j* in trust *i* did not answer both questions *k* and *m*.

The trust level variance for the section score *d* for trust *i* is given by:

$$V_{id} = Var(Y_{id}) = \frac{1}{Q^2} \left\{ \sum_{k=1}^{Q} V_{ik} + 2 \sum_{k=2}^{Q} \sum_{m=1}^{k-1} COV_{ik,im} \right\}$$

The standard error of the section score is then:

$$SE_{id} = \sqrt{V_{id}}$$