Place2Be CORC report

Produced by CORC, March 2018

Reporting period: 01 May 2008 to 30 November 2017

Contents

About CORC

The Child Outcomes Research Consortium (CORC) is the UK's leading membership organisation that collects and uses evidence to improve children and young people's mental health and wellbeing.

Founded in 2002 by a group of mental health professionals determined to understand the impact of their work, today our members include mental health service providers, schools, professional bodies and research institutions from across Europe and beyond.

We analyse and interpret data relating to mental health and wellbeing outcomes of more than 400,000 children and young people in the UK, representing the largest data set of this kind worldwide.

The latest news and resources can be found on the CORC website:
www.corc.uk.net

Jordan House, 47 Brunswick Place, London, N1 6EB
020 7443 2225
CORC@annafreud.org

Key Contact

The key contacts within your service are:

Sarah Golden

Nikhil Naag

CORC
Child Outcomes
Research Consortium

Key Findings

Are children, young people and their families improving?

What do children young people and their families think of the service?

CORC
Child Outcomes
Research Consortium

Key Findings

Data quality:

The paired outcome measures completeness (the proportion of children and young people seen by your service who completed a measure at a first and last time point during their period of contact) was 67% (13000/19536). In the rest of CORC sample (filtered as specified on slide 14), this figure was 31% (2430/7918).

The follow up rates (the proportion of children and young people with a measure recorded at a second time point, out of those with the same measure completed at a first time point) ranged from 62% (10409/16816) for the Parent SDQ Impact to 78% (6189/7947) for the Child SDQ Total Difficulties. In the filtered rest of CORC sample, comparative follow up rates were 15% (921/6080) for the Parent SDQ Impact, and 16% (582/3688) for the Child SDQ Total Difficulties.

Change in raw scores:

The majority of children and young people for whom we have paired data improved by one point or more on child-reported SDQ Total Difficulties (70%) and parent-reported SDQ Total Difficulties and Total Impact (72% and 62%, respectively). Around half (54%) of children with parent data showed progress on self-reported SDQ Total Impact, while the rest reported no change or deterioration of at least one point (the lower proportion of children progressing on the Impact scales may be due to low time 1 scores).

Children in your service appear to progress similarly to children in the filtered rest of CORC, with the exception of Parent SDQ Total Difficulties where improvement was larger in your service. Time 1 scores on child- and parent-reported scales indicate children in your service have a lower starting level of difficulties and impact than children seen by (filtered) rest of CORC services.

Key Findings

Crossing cut-offs: Across the SDQ scales, the majority of children in your service either started above the cut-off (the top 10% of a community sample, in the range of "high"/"very high") and moved below (in the range of "close to average"/"slightly raised") by time 2, or remained below the cut-off at both time points. A higher proportion of children and young people in your service compared to the filtered rest of CORC sample moved from above cut-off at time 1 to below cut-off at time 2.

Added value: Based on analysis of Added Value, your service appears to have a positive effect on the children and young people for whom parent-reported SDQ data has been collected within a 4-8 month time frame. The effect size for your service is stronger than for the rest of CORC (filtered), which did not indicate positive progress. For both samples, we would need a higher follow-up rate within the time frame required for Added Value Score analysis to be able to generalise these results.

Key Findings

Reliable change and 'recovery': Analysis of reliable change was conducted using two different approaches.

On request by the service, we have calculated reliable change on individual scales using the sample of children for whom we have paired data on respective measures **with no filter applied in relation to clinical cut-off at outset or case closure**. The majority of children and young people in your service do not show reliable improvement on SDQ scales, on par with results for the rest of CORC sample (filtered as specified on slide 14). Between 19% and 26% of children reliably improve, while a small number fall in the range of reliable deterioration. For Total Impact, low starting scores may contribute to a smaller proportion of children seeing reliable improvement (average time 1 scores are lower than the reliable change threshold for Child and Parent SDQ Impact in your service, and for Child SDQ Impact in the filtered rest of CORC sample).

We have also calculated reliable change and 'recovery' across child- and parent-reported scales, following inclusion criteria used for analysis in a national sample of child and adolescent mental health services (Wolpert et al., 2016), **which restricts the sample to include only closed cases with three or more recorded events, and above a clinical threshold on at least one paired child- or parent- reported measure at a first time point** (the 'paired clinical sample').

3655 of the 19536 cases (children and young people's periods of contact) seen by your service met criteria for analysis of child-reported reliable change and 'recovery'. Taking into account uncertainty in the data due to random variation, we estimate that between 43% and 46% reliably improved, between 45% and 48% 'recovered', and between 26% and 29% reliably 'recovered'. This compares to reliable improvement of between 51% and 53%, 'recovery' of between 35% and 37%, and reliable 'recovery' of between 25% and 28% in national data (see Appendix D).

8029 of the 19536 cases (children and young people's periods of contact) seen by your service met criteria for analysis of parent-reported reliable change and 'recovery'. Taking into account uncertainty in the data due to random variation, we estimate that between 42% and 45% reliably improved, between 41% and 43% 'recovered', and between 23% and 25% reliably 'recovered'. This compares to reliable improvement of between 39% and 42%, 'recovery' of between 26% and 29%, and reliable 'recovery' of between 15% and 18% in national data (see Appendix D).

Wolpert, M., Jacob, J., Napoleone, E., Whale, A., Calderon, A., & Edbrooke-Childs, J. (2016). *Child- and Parent-reported Outcomes and Experience from Child and Young People's Mental Health Services 2011-2015*. London: CAMHS Press

CORC
Child Outcomes
Research Consortium

Implications

What are the main implications?

Implications

Data collection: Your service had higher follow up rates on individual scales and a higher overall paired outcome rate compared to services in the rest of CORC data set (filtered as specified on slide 14). You may wish to ensure that these rates are maintained and are continued to be reinforced in your staff group. You might also consider whether you would find it meaningful to capture any additional outcome or feedback measures to gain further insight into how service users are progressing and experiencing your service.

Examining progress: Children in your service showed positive progress in child- and parent- reported difficulties and impact in terms of the change in raw scores. Progress was similar in your service and in the filtered rest of CORC sample, except for on the Parent SDQ, where scores improved more on average in your service. You may wish to discuss whether results reflect the kind of impact your service targets, and what factors could be influencing the outcomes found (for example, completion rates, differences between young people in the two samples, appropriateness of outcome measures used).

Interpreting reliable change: Reliable change results calculated on the respective SDQ scales suggest that in a minority of cases with paired data in your service (between 19% and 26%) improvement is more than could be attributed to measurement error. A higher proportion of children who met criteria for the restricted 'paired clinical sample' showed reliable improvement when analysis was conducted across measures (between 43% and 45% on child-reported measures, and between 42% and 45% on parent-reported measures). It is worth considering how differences between samples used for these analyses may shape results, particularly in regards to case closure, number of sessions attended, and time 1 scores (the case characteristics used to define the analytical sample in Wolpert et al., 2016).

Your Report

A Note About the Data in this Report

Some tables and graphs in this report contain instances of small numbers (for example, information relating to one or two individuals).

We would encourage you to consider the level of risk to service-user confidentiality before sharing this information with anyone who does not already have access to the original data. For guidance, see for example the *Anonymisation Standard for Publishing Health and Social Care Data*, available from the NHS Digital website (https://digital.nhs.uk/home).

Please do get in touch with the CORC team (CORC@annafreud.org) should you require any support to suppress data or information relating to small numbers in this report.

CORC
Child Outcomes
Research Consortium
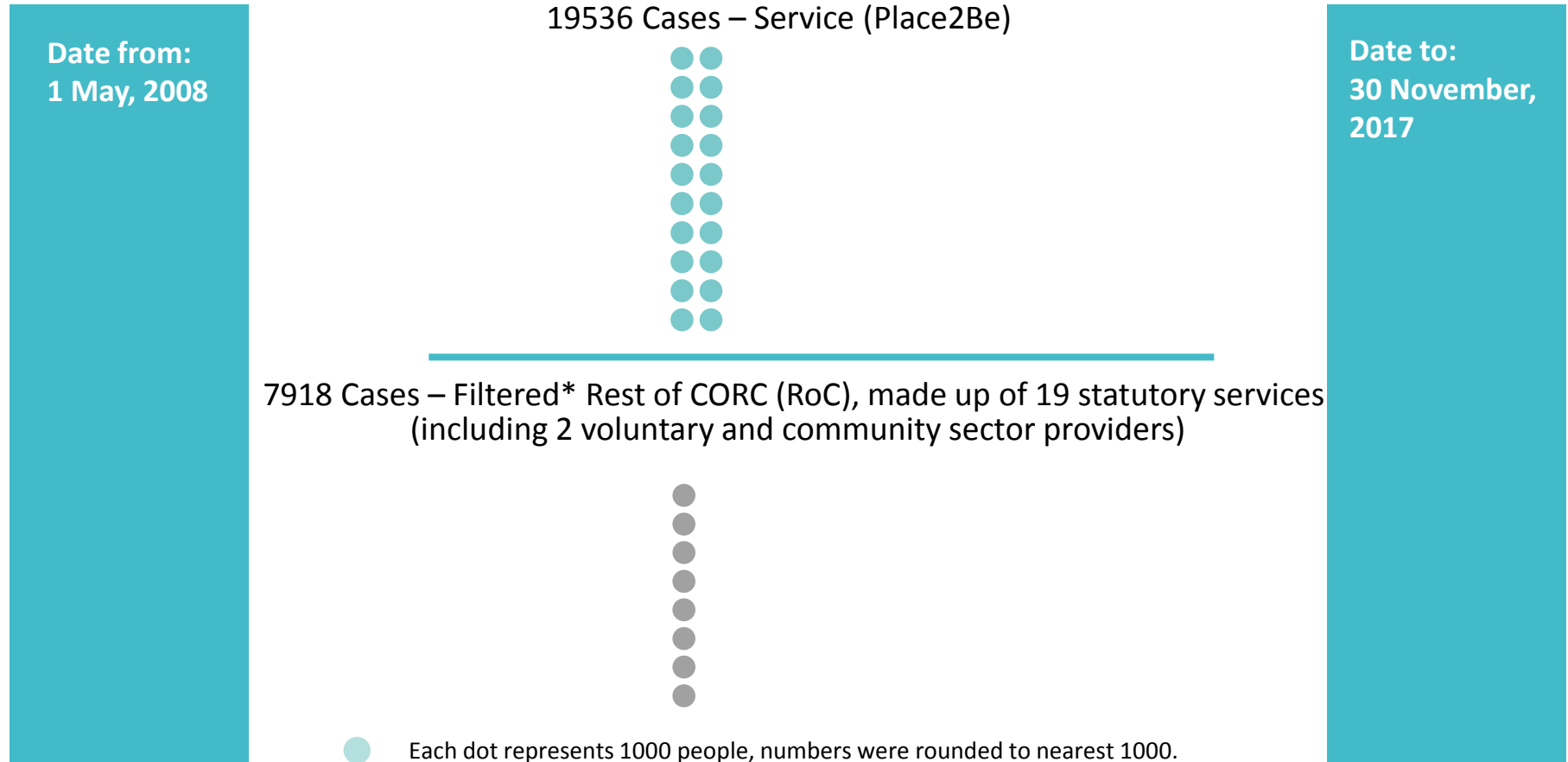
Data Completion: Outcome Measures

What is the sample size?

Paired outcome rates

Follow up rates

Who is in your sample?

CORC
Child Outcomes
Research Consortium

What is the sample size?

Date from:
1 May, 2008

Date to:
30 November, 2017

19536 Cases – Service (Place2Be)

7918 Cases – Filtered* Rest of CORC (RoC), made up of 19 statutory services (including 2 voluntary and community sector providers)

Each dot represents 1000 people, numbers were rounded to nearest 1000.

*At the request of the service, to make the 'Rest of CORC (RoC)' sample more comparable to the Place2Be sample, data has been filtered to include only cases from teams identifying as Tier 2 or Tier 2/3, between the ages of 6 and 14 and with a valid time 1 SDQ. Please note team type information is self-reported by services and is not an official service type designation. Tier 2 generally indicates a combination of some specialist CAMH services and community-based services such as primary mental health work.

The dataset consists of demographics and mental health outcome information collected locally by members and submitted for collation by the CORC Team; the main purposes are service evaluation and to inform clinical practice.

CORC
Child Outcomes
Research Consortium

How many children and young people have a paired outcome measure?

Paired Outcome Rate: 'Paired Outcome Rate' shows the proportion of children and young people with **any** paired outcome measure.
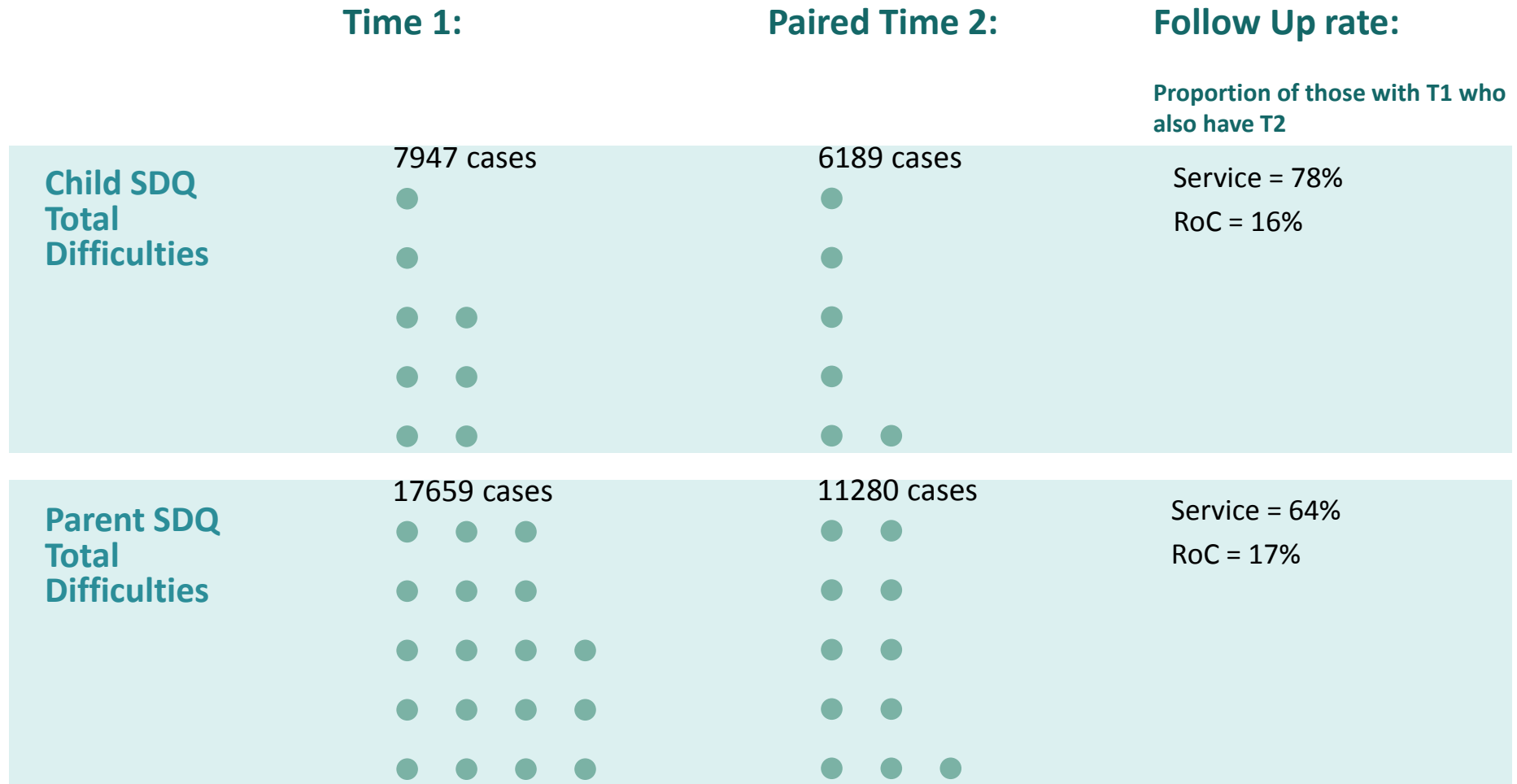
Service = 67%, 13000 cases

Rest of CORC = 31%, 2430 cases

Service=67%

RoC=31%

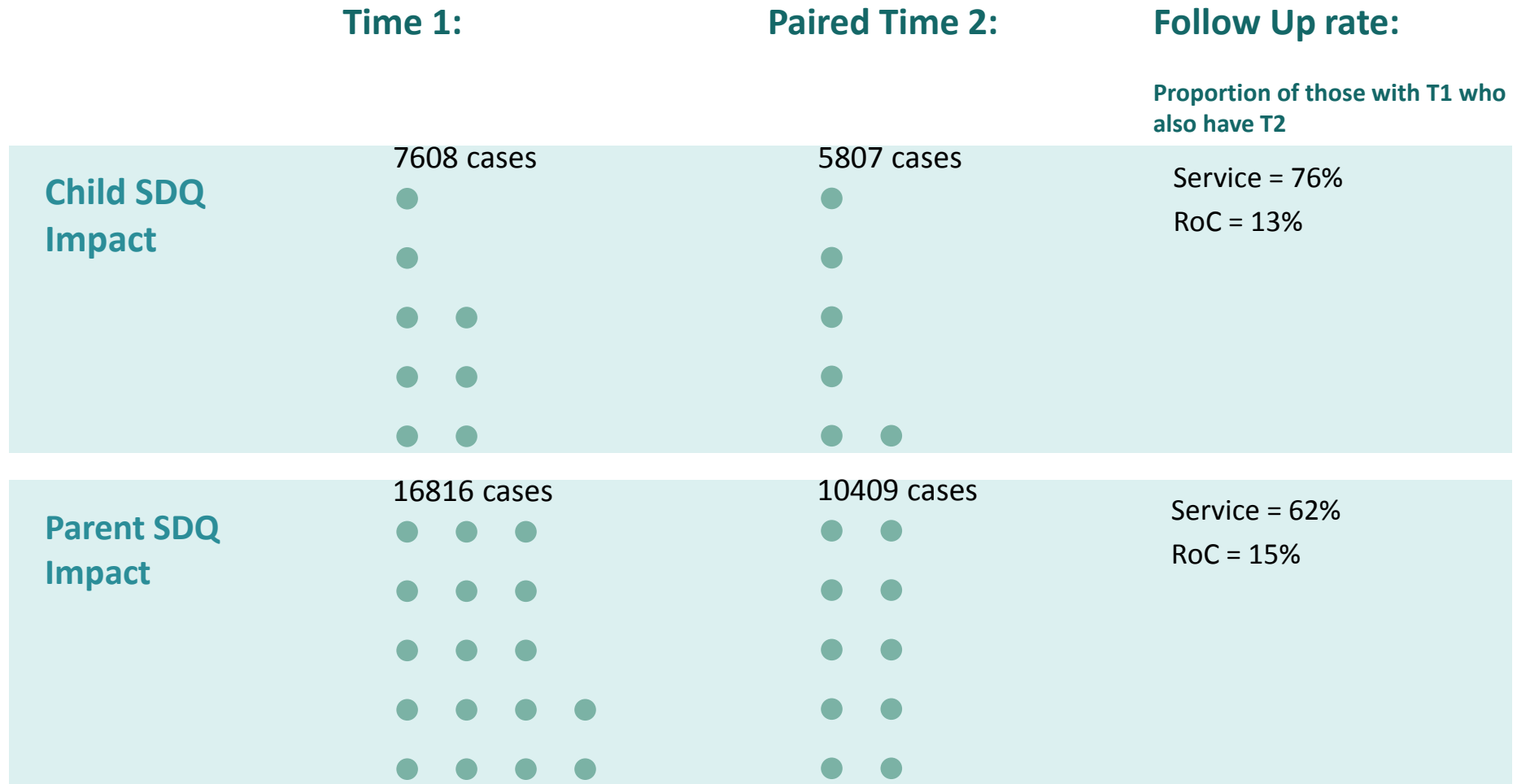How many children and young people have a paired outcome measure?

15

What are the Follow Up rates?

	Time 1:	**Paired Time 2:**	**Follow Up rate:**
			Proportion of those with T1 who also have T2
Child SDQ Total Difficulties	7947 cases	6189 cases	Service = 78% RoC = 16%
Parent SDQ Total Difficulties	17659 cases	11280 cases	Service = 64% RoC = 17%

Each dot represents 1000 people, numbers were rounded to nearest 1000.

CORC
Child Outcomes
Research Consortium

What are the Follow Up rates?

	Time 1:	Paired Time 2:	Follow Up rate:
			Proportion of those with T1 who also have T2
Child SDQ Impact	7608 cases	5807 cases	Service = 76% RoC = 13%
Parent SDQ Impact	16816 cases	10409 cases	Service = 62% RoC = 15%

● Each dot represents 1000 people, numbers were rounded to nearest 1000.

CORC
Child Outcomes
Research Consortium

Who is in your sample?

19536 people were seen by the service

This slide displays a summary of age, gender, ethnicity and number of sessions attended for cases shown in further detail in the report. This sample may not be representative of every child seen by your service as data may not have been submitted for all children, and completeness may vary across variables.

## If the service saw 100 people:

Age
People would be:
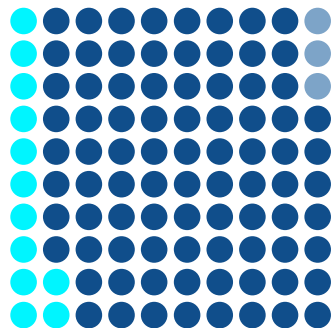
- 0-5 years
- 6-12 years
- 13-18 years

Gender
People would be:
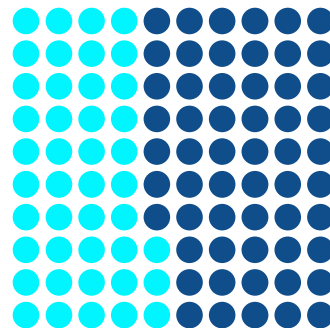
- Female
- Male

Ethnicity
People would be:
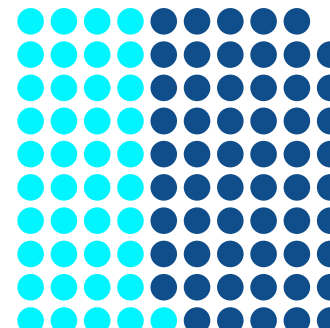
- Other Ethnic Group
- White British

Sessions
People would attend:
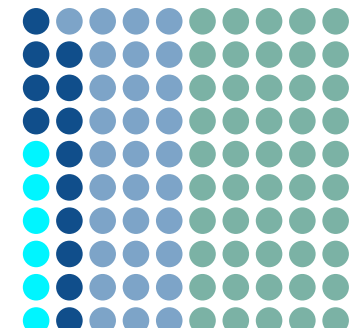
- 1 to 5
- 6 to 10
- 11 to 20
- More than 21

Due to rounding, percentages may not add up to 100.

Who in your sample has *paired data*?

13000 people were seen by the service

This slide displays a summary of age, gender, ethnicity and number of sessions attended for cases with paired data on any measure analysed in this report. This sample may not be representative of every child seen by your service as data may not have been submitted for all children, and completeness may vary across variables.

If the service saw 100 people:

Age
People would be:
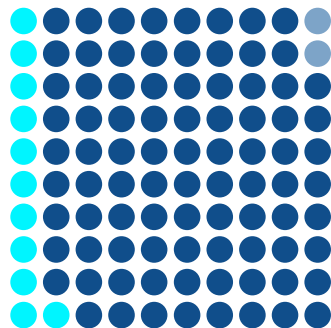
0-5 years
6-12 years
13-18 years

Gender
People would be:
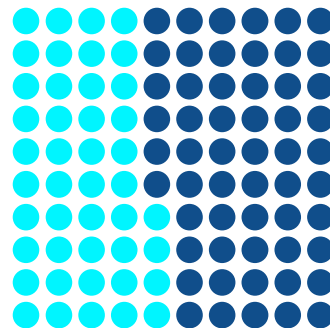
Female
Male

Ethnicity
People would be:
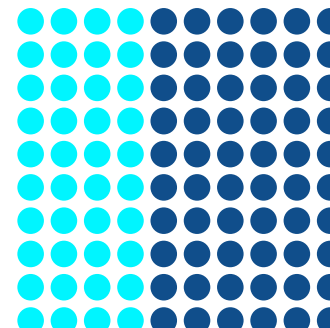
Other Ethnic Group
White British

Sessions
People would attend:
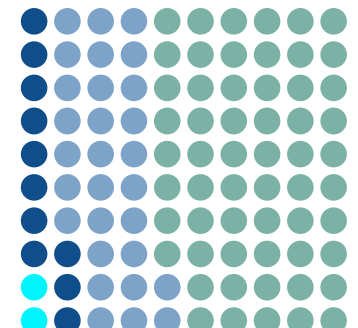
1 to 5
6 to 10
11 to 20
21 or more

CORC
Child Outcomes
Research Consortium

About the Plots

Information about how to read and interpret the plots.

What are Margins of Error?

Why show margins of error? When we make statistical comparisons, we have to take the uncertainty in the data into account. This can be caused by small sample sizes or very varied data. The margin of error gives a range of numbers which we are reasonably certain contains the true average. If the interval is narrow, we are quite certain what the true average is. If it is wide, we are not.

How can margins of error be used to evaluate change scores? When the margin of error of an average change score doesn't cross 0, it suggests that there may be a difference between two scores. When it crosses 0, there is no evidence to suggest that the two scores are different.

How can margins of error be used to compare averages?

As a rough rule, if the two margins of error overlap by more than half their average length, there is no evidence to suggest that the two averages are different.
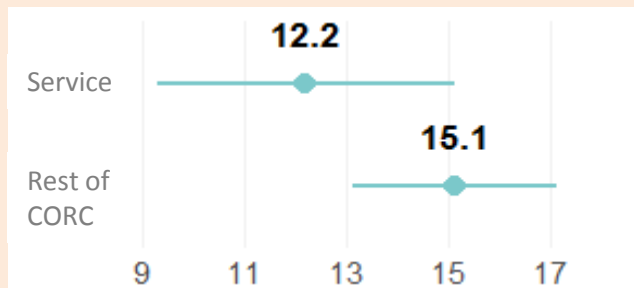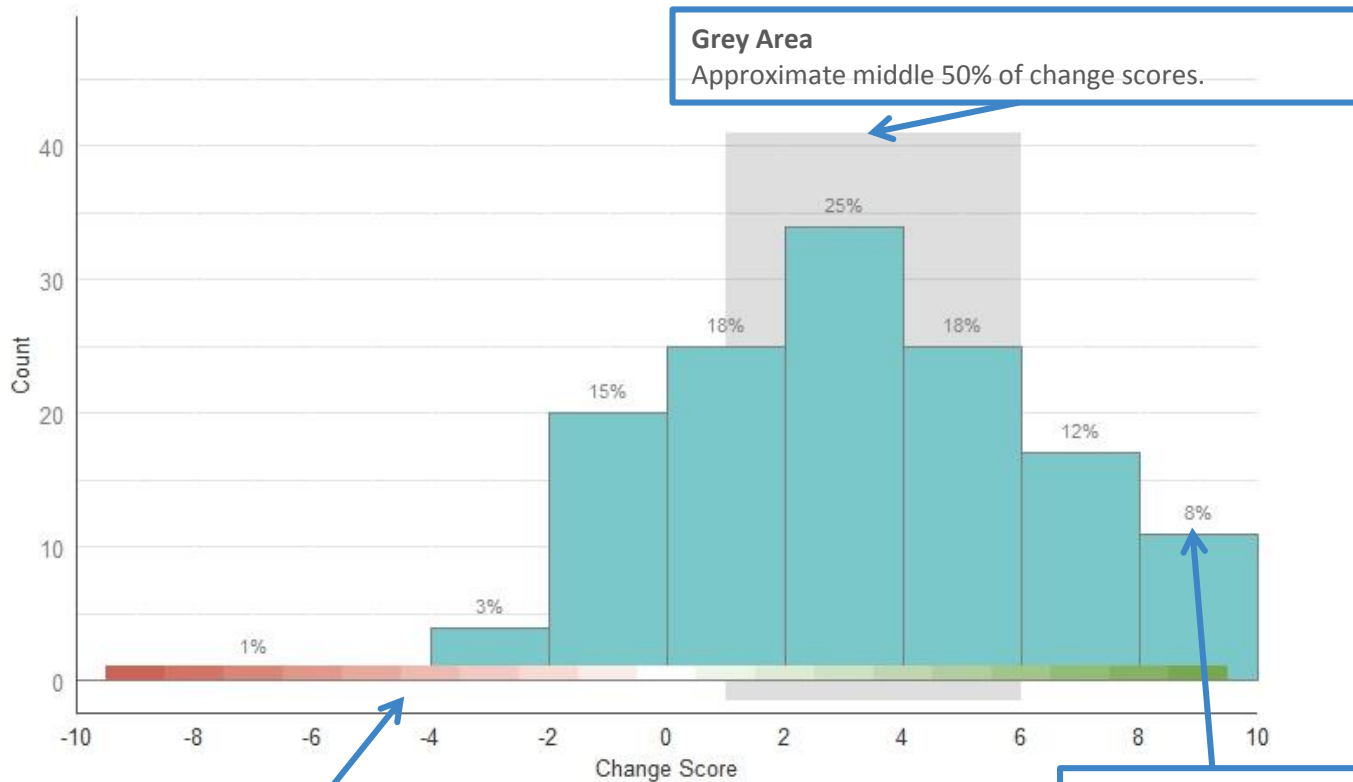
If the two margins of error overlap by less than half their average length, or not at all, we may want to investigate why there seems be a difference between the two averages.

What do the plots mean?

Child
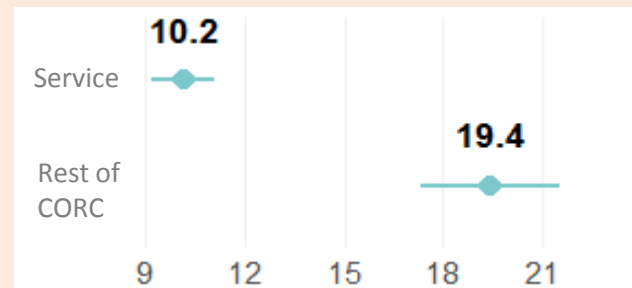
Grey Area
Approximate middle 50% of change scores.

Results included
Any paired scores. The change score is the difference between the time 1 score and the time 2 score.

What do the plots show us?
A short explanation of the graph and results of tests conducted. An indication of whether scores appear to be improving over time or not. Note that percentages in the text and graph sometimes don't fully correspond due to rounding.

Technical details are included in the appendix.

Data Completeness
The proportion of those with a time 1 score, who also have a time 2 score.

Follow up rate 17%

17% full

Progress Bar
Red if score has got **worse**,
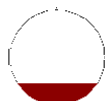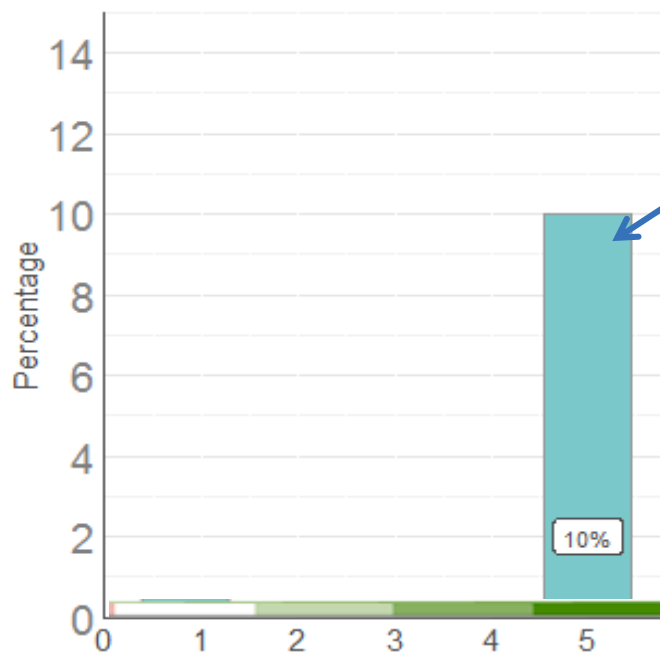White if score has **not changed**,
Green if score has **improved**.

Labels
Give the percentage of individuals with a score between the two values in the horizontal axis.

How are 'change' scores calculated?

Change= SDQ Total Difficulties (Time 2) − SDQ Total Difficulties (Time 1)

Change = 15 − 20

Change= -5

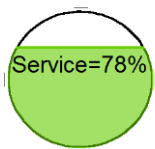10% of children's scores went down by 5 points between time 1 and time 2.
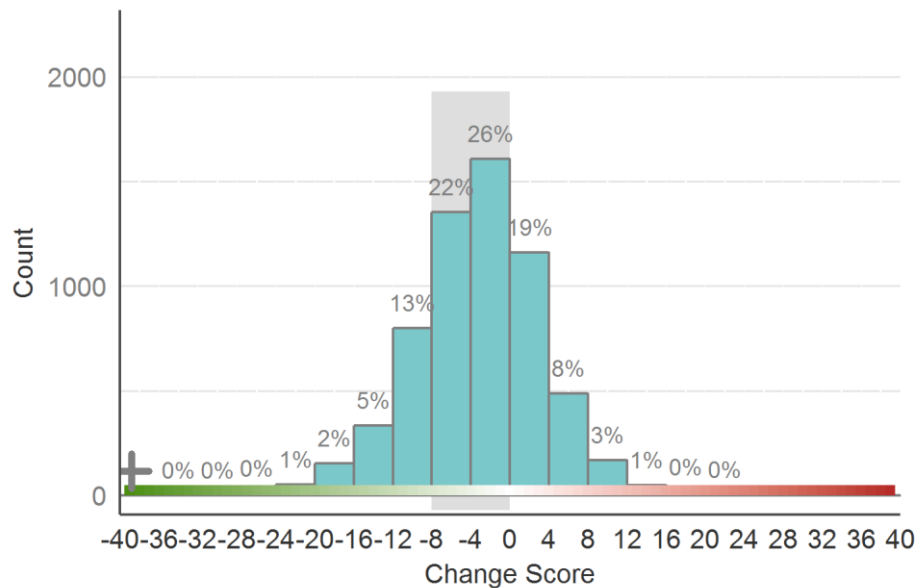
We plot these change scores to get an impression of how much all individuals have changed between time 1 and time 2.

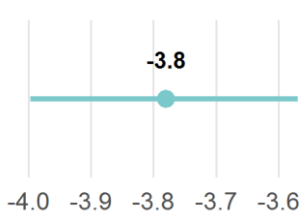# Are children, young people and their families improving?

Child SDQ Total Difficulties
Parent SDQ Total Difficulties
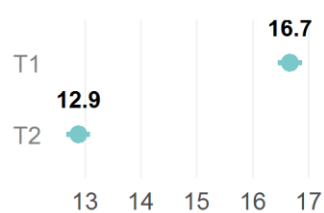Child SDQ Impact
Parent SDQ Impact

How have Child SDQ Total Difficulties scores changed between T1 and T2?

Results included

Any paired Child SDQ Total Difficulties scores, n = 6189

How representative is this sample?

78% (6189/7947) of those with a time 1 score, had a corresponding time 2 score at your service.

What do the plots show?

On average, scores changed by -4 points, with a tight margin of error.

Looking at the spread of change scores, 70% (4309/6189) of scores improved, 6% (371/6189) stayed the same, and 24% (1509/6189) deteriorated. The middle 50% of children and young people with paired scores changed by between -8 and 0 points.

The difference between time 1 and time 2 scores has a Cohen's d effect size of -0.61, in the direction of progress.

Conclusion

The majority (70%) of children and young people in your service improve on the Child SDQ Total Difficulties scale and, on average, scores improve by around 4 points. However, for a minority of children scores stay the same (6%) or get worse (24%).

A high follow-up rate on this measure means we can be certain scores can be generalised to all children and young people who complete this measure at a first time point in your service.

CORC
Child Outcomes
Research Consortium

Have Child SDQ Total Difficulties Scores Reliably Changed?
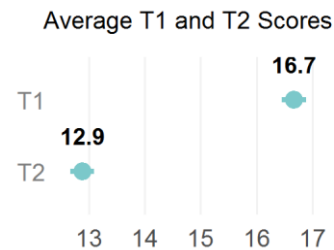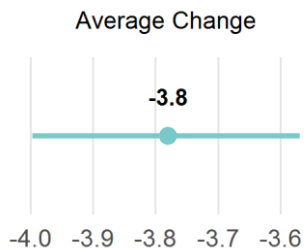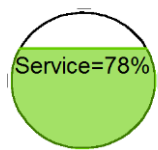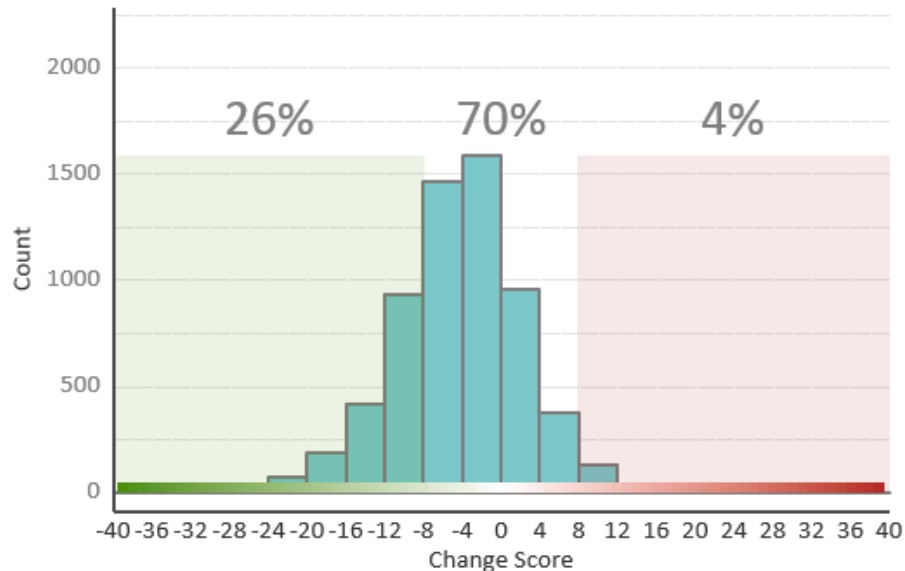
Results included

Any paired score for Child SDQ Total Difficulties (n = 6189), categorised by whether they fall above or below the threshold for reliable change. Change scores higher/lower than +/- 7.81 were classified as reliable change.

How representative is this sample?

78% (6189/7947) of those with a time 1 score, had a corresponding time 2 score at your service.

What do the plots show?

26% of children and young people showed reliable improvement between time 1 and time 2. 70% of children did not show reliable change, and a small proportion (4%) had scores that reliably deteriorated.

Conclusion

26% of children reliably improved between time 1 and time 2 on the Child SDQ Total Difficulties scale. The majority of children did not show reliable change, meaning we cannot be certain that any observed change in scores was not due to measurement error for these cases. However, this does not mean that any observed change was not meaningful to an individual child or young person or their family.

Note: The standard deviation used to calculate reliable change in this report is based on the full CYP IAPT sample up to Quarter 2, 2015.

How have Parent SDQ Total Difficulties scores changed between T1 and T2?
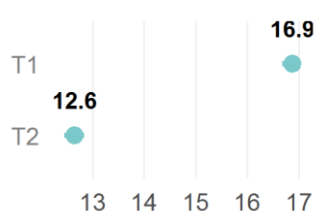
Results included

Any paired Parent SDQ Total Difficulties scores, n = 11280

How representative is this sample?

64% (11280/17659) of those with a time 1 score, had a corresponding time 2 score at your service.
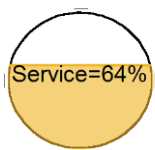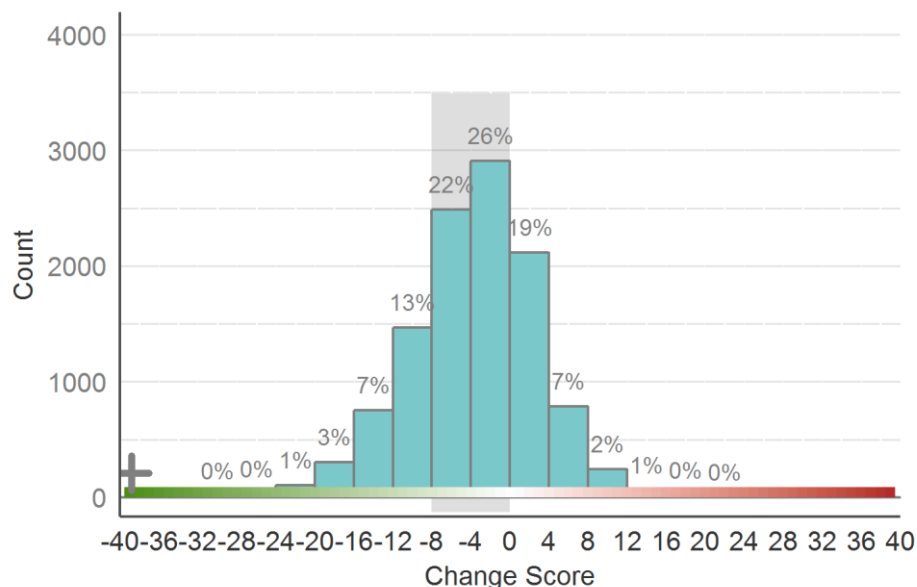
What do the plots show?

On average, scores changed by -4 points, with a tight margin of error.

Looking at the spread of change scores, 71% (8057/11280) of scores improved, 6% (724/11280) stayed the same, and 22% (2499/11280) deteriorated. The middle 50% of children and with paired scores changed by between -8 and 0 points.

The difference between time 1 and time 2 scores has a Cohen's d effect size of -0.61, in the direction of progress.

Conclusion

The majority (71%) of children and young people in your service improve on the Parent SDQ Total Difficulties scale and, on average, scores improve by around 4 points. However, for a minority of children scores stay the same (6%) or get worse (22%).

A good follow-up rate on this measure means we can be fairly certain scores can be generalised to all children and young people whose parents complete this measure at a first time point in your service.

Have Parent SDQ Total Difficulties Scores Reliably Changed?
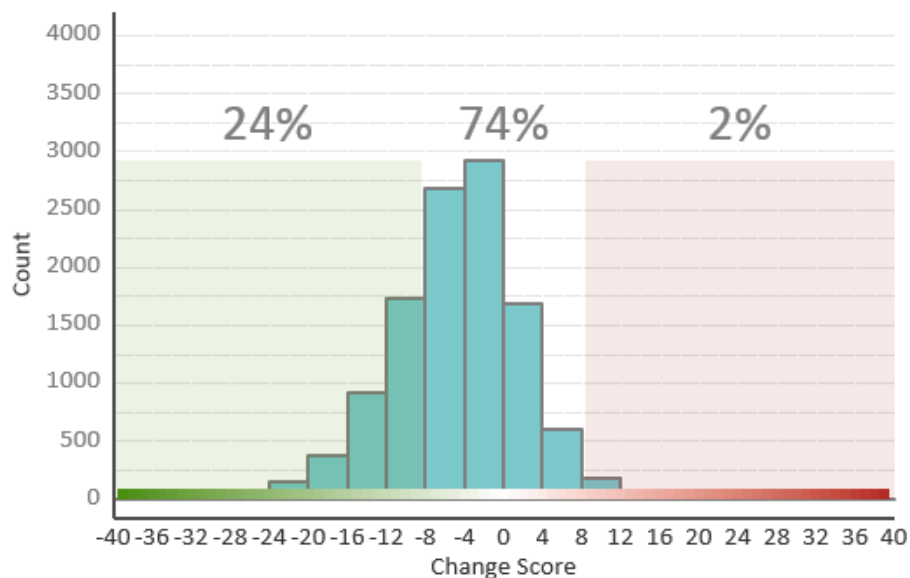
Results included

Any paired score for Parent SDQ Total Difficulties (n = 11280), categorised by whether they fall above or below the threshold for reliable change. Change scores higher/lower than +/- 8.37 were classified as reliable change.

How representative is this sample?

64% (11280/17659) of those with a time 1 score, had a corresponding time 2 score at your service.

What do the plots show?

24% of children and young people showed reliable improvement between time 1 and time 2. 74% of children did not show reliable change, and a small proportion (2%) had scores that reliably deteriorated.

Conclusion

23% of children reliably improved between time 1 and time 2 on the Parent SDQ Total Difficulties scale. The majority of children did not show reliable change, meaning we cannot be certain that any observed change in scores was not due to measurement error for these cases. However, this does not mean that any observed change was not meaningful to an individual child or young person or their family.

Note: The standard deviation used to calculate reliable change in this report is based on the full CYP IAPT sample up to Quarter 2, 2015.

How have Child SDQ Impact scores changed between T1 and T2?
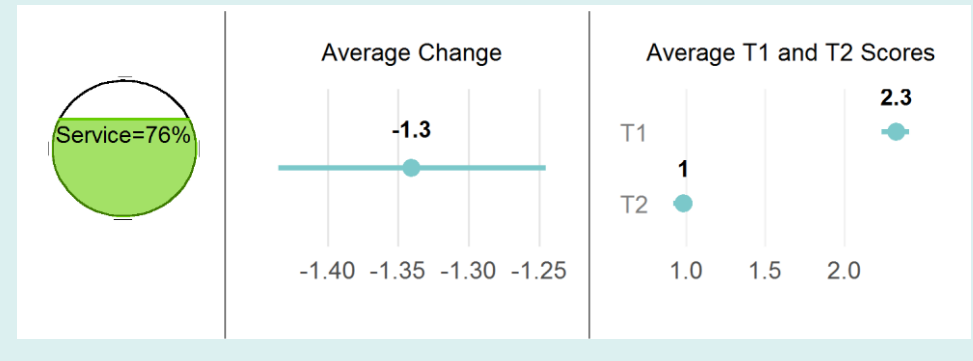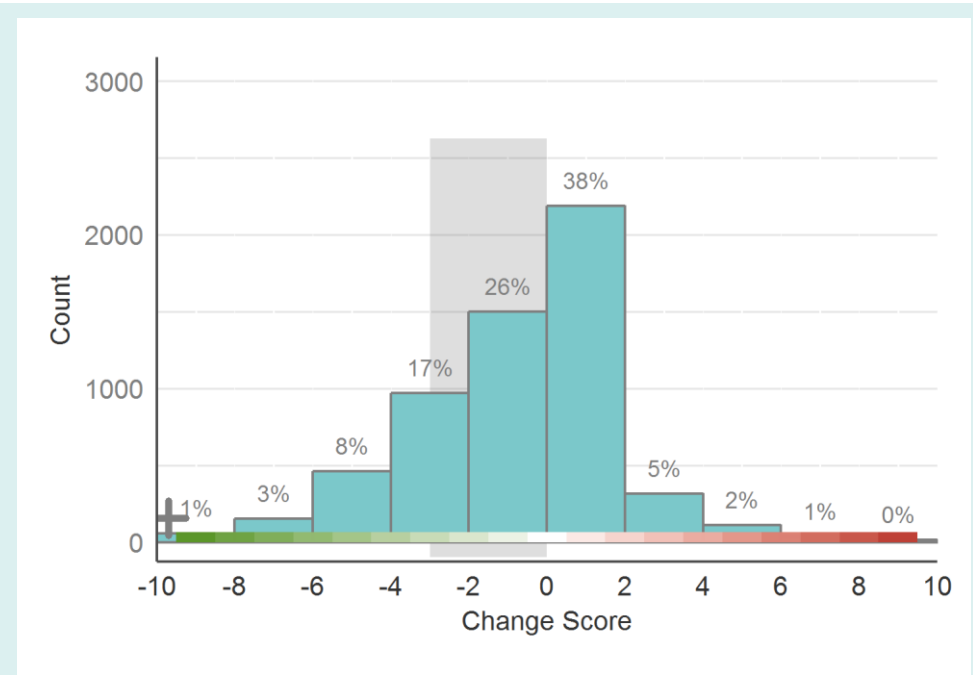
Results included

Any paired Child SDQ Impact scores, n = 5807

How representative is this sample?

76% (5807/7608) of those with a time 1 score, had a corresponding time 2 score at your service.

What do the plots show?

On average, scores changed by -1 point, with a tight margin of error.

Looking at the spread of change scores, 54% (3139/5807) of scores improved, 32% (1860/5807) stayed the same, and 14% (808/5807) deteriorated. The middle 50% of children and young people with paired scores changed by between -3 and 0 points.

The difference between time 1 and time 2 scores has a Cohen's d effect size of -0.55, in the direction of progress.

Conclusion

Around half (54%) of children and young people in your service improve on the Child SDQ Total Impact scale and, on average, scores improve by around 1 point. Scores remained stable for 32% of children, and got worse for 14%. Low time 1 scores may contribute to a higher proportion of children with change scores in the range of 0 to 2.

A high follow-up rate on this measure means we can be certain scores can be generalised to all children and young people who complete this measure at a first time point in your service.

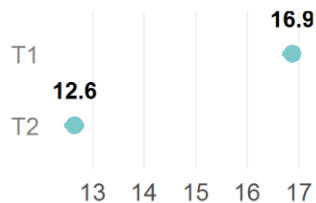# Have Child SDQ Total Impact Scores Reliably Changed?
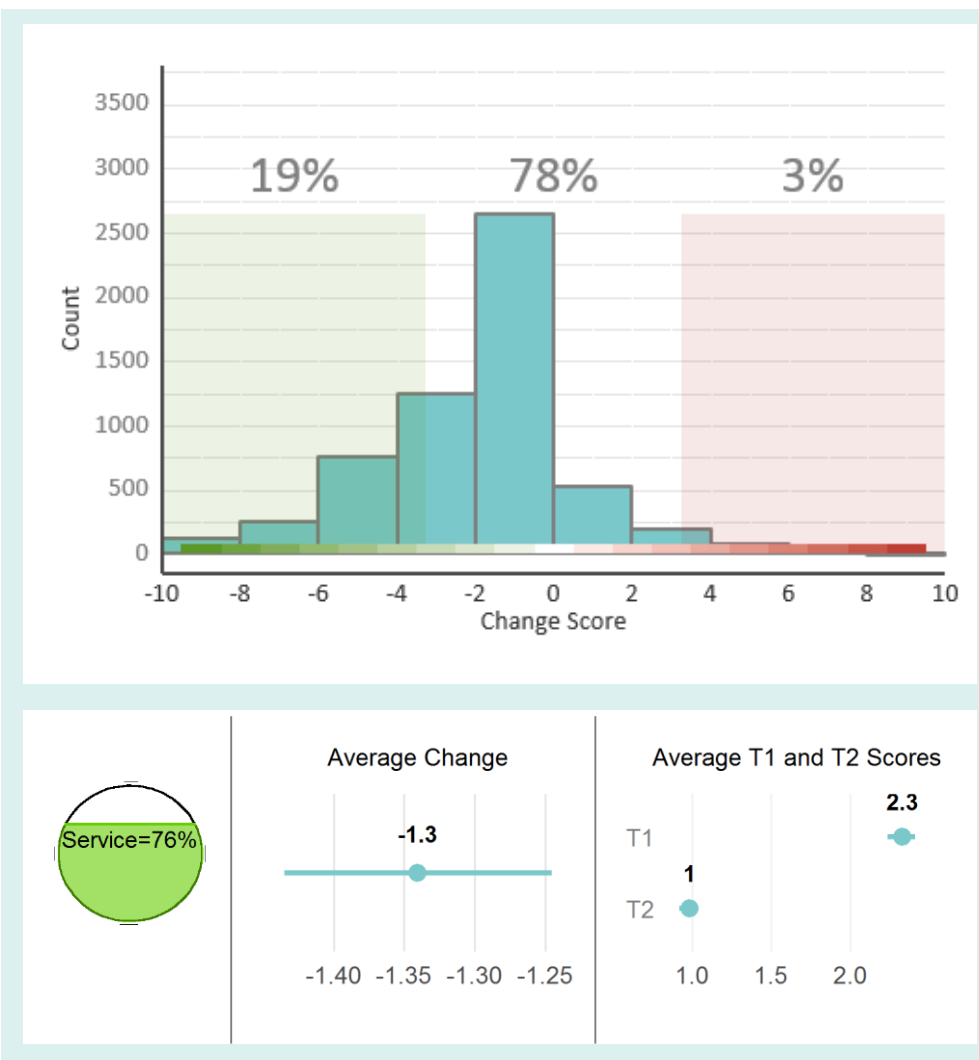
Results included

Any paired score for Child SDQ Total Impact (n = 5807) categorised by whether they fall above or below the threshold for reliable change. Change scores higher/lower than +/- 3.25 were classified as reliable change.

How representative is this sample?

76% (5807/7608) of those with a time 1 score, had a corresponding time 2 score at your service.

What do the plots show?

19% of children and young people showed reliable improvement between time 1 and time 2. 78% of children did not show reliable change, and a small proportion (3%) had scores that reliably deteriorated.

Low time 1 scores may contribute to a higher proportion of children not showing reliable change on this measure, since the threshold for reliable change is higher than the average time 1 score.

Conclusion

19% of children reliably improved between time 1 and time 2 on the Child SDQ Total Impact scale. The majority of children did not show reliable change, meaning we cannot be certain that any observed change in scores was not due to measurement error for these cases. However, this does not mean that any observed change was not meaningful to an individual child or young person or their family.

Note: The standard deviation used to calculate reliable change in this report is based on the full CYP IAPT sample up to Quarter 2, 2015.

How have Parent SDQ Impact scores changed between T1 and T2?
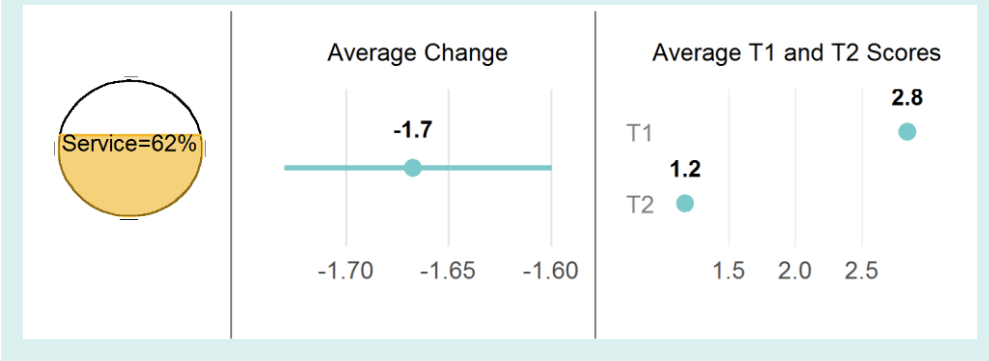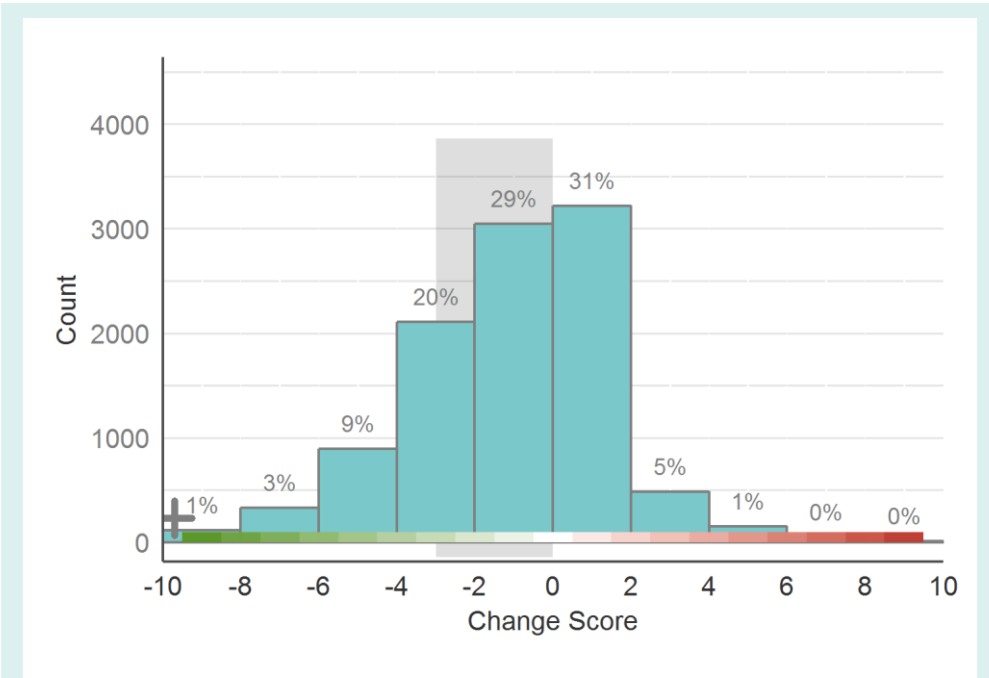
Results included

Any paired Parent SDQ Impact scores, n = 10409

How representative is this sample?

62% (10409/16816) of those with a time 1 score, had a corresponding time 2 score at your service.

What do the plots show?

On average, scores changed by -2 points, with a tight margin of error.

Looking at the spread of change scores, 62% (6495/10409) of scores improved, 26% (2681/10409) stayed the same, and 12% (1233/10409) deteriorated. The middle 50% of children and young people with paired scores changed by between -3 and 0 points.

The difference between time 1 and time 2 scores has a Cohen's d effect size of -0.68, in the direction of progress.

Conclusion

More than half (62%) of children and young people in your service improve on the Parent SDQ Total Impact scale and, on average, scores improve by around 2 points. Scores remained stable for 26% of children, and got worse for 12%. Low time 1 scores may contribute to a higher proportion of children with change scores in the range of 0 to 2.

A good follow-up rate on this measure means we can be fairly certain scores can be generalised to all children and young people whose parents complete this measure at a first time point in your service.

CORC
Child Outcomes
Research Consortium

31

Have Parent SDQ Total Impact Scores Reliably Changed?
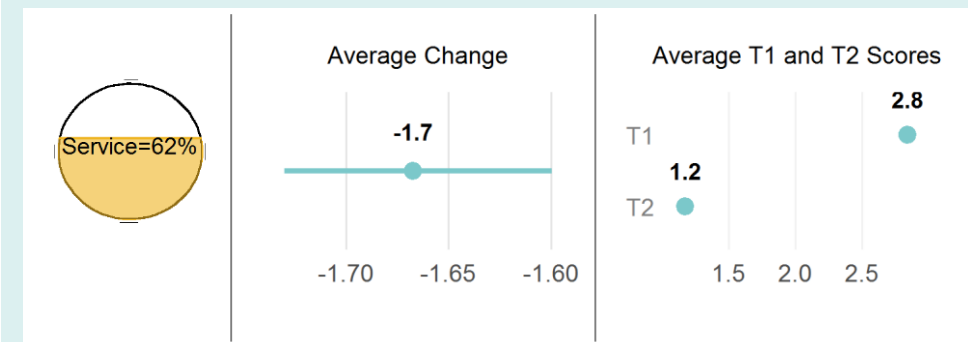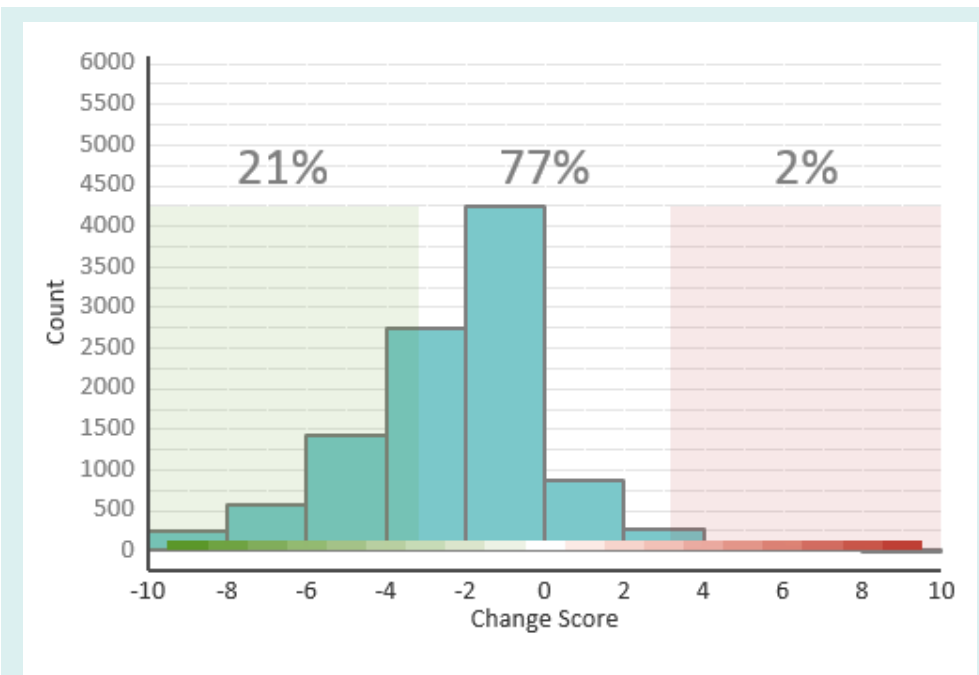
Results included

Any paired score for Parent SDQ Total Impact (n = 10409), categorised by whether they fall above or below the threshold for reliable change. Change scores higher/lower than +/- 3.19 were classified as reliable change.

How representative is this sample?

62% (10409/16816) of those with a time 1 score, had a corresponding time 2 score at your service.

What do the plots show?

21% of children and young people showed reliable improvement between time 1 and time 2. 77% of children did not show reliable change, and a small proportion (2%) had scores that reliably deteriorated.

Low time 1 scores may contribute to a higher proportion of children not showing reliable change on this measure, since the threshold for reliable change is higher than the average time 1 score.

Conclusion

21% of children reliably improved between time 1 and time 2 on the Parent SDQ Total Impact scale. The majority of children did not show reliable change, meaning we cannot be certain that any observed change in scores was not due to measurement error for these cases. However, this does not mean that any observed change was not meaningful to an individual child or young person or their family.

Note: The standard deviation used to calculate reliable change in this report is based on the full CYP IAPT sample up to Quarter 2, 2015.

"Recovery" and Reliable Change

Child-reported measures: Note that figures represent analyses conducted on a filtered 'paired clinical sample' to assess "recovery" and reliable change as defined below, following Wolpert et al. (2016). Figures may therefore differ from reliable change results reported elsewhere in the report and should be compared with caution, taking into account differences in samples and analytical approach.

Indicator	Definition	N	% of paired clinical sample [95% Margins of Error]
'Recovery'	Moved from above a clinical threshold on at least one paired measure at a first time point, to below on all completed measures at a last time point	1695	46% [45%-48%]
Reliable Improvement	Change from a first to a last time point was more than what would be expected due to measurement error, in a positive direction, on at least one measure, and no measure reliably deteriorated	1631	45% [43%-46%]
No reliable change	Change from a first to a last time point was less than what would be expected due to measurement error	1729	47% [46%-49%]
Reliable deterioration	Change from a first to a last time point was more than what would be expected due to measurement error, in a negative direction, on at least one measure	295	8% [6%-10%]
Reliable 'Recovery'	Moved from above a clinical threshold on at least one paired measure at a first time point, to below on all completed measures at a last time point, and the change was reliable in a positive direction, with no measures reliably deteriorating	1000	27% [26%-29%]

The following measures are included: Child SDQ Emotional, Child SDQ Conduct, Child SDQ Hyperactivity, Child SDQ Impact

Sample included

Any closed cases with three or more recorded events, and above a clinical threshold on at least one paired child-reported measure at a first time point (the 'paired clinical sample'), n= 3655

How representative is this sample?

19% (3655/19536) of all submitted cases fell in the 'paired clinical sample'.

What does the table show?

Scores for 46% (margin of error between 45% and 48%) of children and young people showed 'recovery', 45% (margin of error between 43% and 46%) showed reliable improvement, and 27% (margin of error between 26% and 29%) showed 'reliable recovery'.

Conclusion

Taking into account uncertainty in the data, figures for 'recovery' were higher, and figures for reliable improvement were slightly lower in your service than those reported on child-reported measures in Wolpert et al. (2016). Scores for reliable 'recovery' were similar in the two samples. It is worth noting that analysis included only Child SDQ subscales for your service, whereas analyses in Wolpert et al. (2016) included a suite of child-reported routine outcome measures.

Wolpert, M., Jacob, J., Napoleone, E., Whale, A., Calderon, A., & Edbrooke-Childs, J. (2016). *Child- and Parent-reported Outcomes and Experience from Child and Young People's Mental Health Services 2011-2015.* London: CAMHS Press

CORC
Child Outcomes
Research Consortium

"Recovery" and Reliable Change

Parent-reported measures: Note that figures represent analyses conducted on a filtered 'paired clinical sample' to assess "recovery" and reliable change as defined below, following Wolpert et al. (2016). Figures may therefore differ from reliable change results reported elsewhere in the report and should be compared with caution, taking into account differences in samples and analytical approach.

Indicator	Definition	N	% of paired clinical sample [95% Margins of Error]
'Recovery'	Moved from above a clinical threshold on at least one paired measure at a first time point, to below on all completed measures at a last time point	3351	42% [41%-43%]
Reliable Improvement	Change from a first to a last time point was more than what would be expected due to measurement error, in a positive direction, on at least one measure, and no measure reliably deteriorated	3490	43% [42%-45%]
No reliable change	Change from a first to a last time point was less than what would be expected due to measurement error	4079	51% [50%-52%]
Reliable deterioration	Change from a first to a last time point was more than what would be expected due to measurement error, in a negative direction, on at least one measure	460	6% [5%-7%]
Reliable 'Recovery'	Moved from above a clinical threshold on at least one paired measure at a first time point, to below on all completed measures at a last time point, and the change was reliable in a positive direction, with no measures reliably deteriorating	1913	24% [23%-25%]

The following measures are included: Parent SDQ Emotional, Parent SDQ Conduct, Parent SDQ Hyperactivity, Parent SDQ Impact

Sample included
Any closed cases with three or more recorded events, and above a clinical threshold on at least one paired parent-reported measure at a first time point (the 'paired clinical sample'), n= 8029

How representative is this sample?
41% (8029/19536) of all submitted cases fell in the 'paired clinical sample'.

What does the table show?
Scores for 42% (margin of error between 41% and 43%) of children and young people showed 'recovery', 43% (margin of error between 42% and 45%) showed reliable improvement, and 24% (margin of error between 23% and 25%) showed 'reliable recovery'.

Conclusion
Taking into account uncertainty in the data, figures for 'recovery' and reliable 'recovery' were higher in your service than those reported on parent-reported measures in Wolpert et al. (2016). Scores for reliable improvement were similar in the two samples. It is worth noting that analysis included only Parent SDQ subscales for your service, whereas analyses in Wolpert et al. (2016) included a suite of parent-reported routine outcome measures.

Wolpert, M., Jacob, J., Napoleone, E., Whale, A., Calderon, A., & Edbrooke-Childs, J. (2016). *Child- and Parent-reported Outcomes and Experience from Child and Young People's Mental Health Services 2011-2015.* London: CAMHS Press

Are children, young people and their families improving compared to the Rest of CORC sample?

Child SDQ Total Difficulties
Parent SDQ Total Difficulties
Child SDQ Impact
Parent SDQ Impact
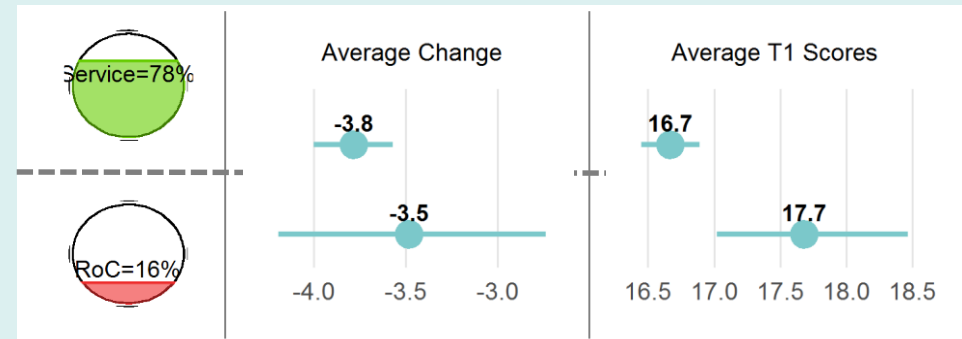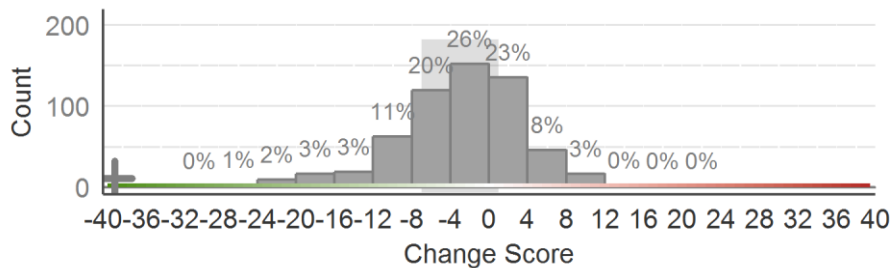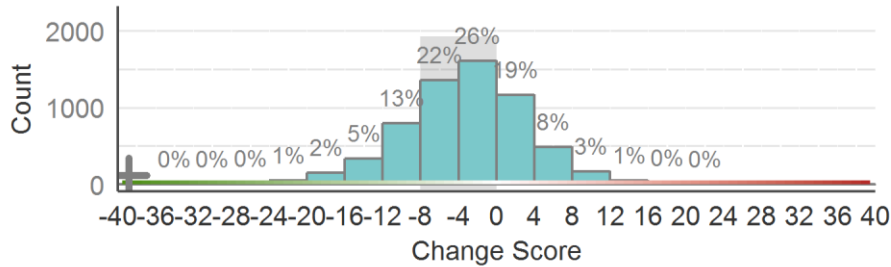
Results included

Any paired Child SDQ Total Difficulties scores: Service n = 6189, Rest of CORC n = 582

How representative is this sample?

78% (6189/7947) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 16% (582/3688).

What do the plots show?

On average, scores changed by -4 points, with a tight margin of error. In RoC services, the average change score was -3, between -4 and -3.

Looking at the spread of change scores, 70% (4309/6189) of scores improved in your service, and 65% (381/582) in the RoC sample. The middle 50% of children and young people with paired scores changed by between -8 and 0 points for your service, and between -7 and 1 for the RoC sample.

The difference between time 1 and time 2 scores has a Cohen's d effect size of -0.61 in your service and -0.60 in the RoC sample, both in the direction of progress.

Conclusion

A majority of children improve on the Child SDQ in both your service and the rest of CORC (70% and 65%, respectively) and children progress similarly, on average, in the two samples. Average time 1 scores indicate children in your service had a slightly lower level of starting difficulties than children in the rest of CORC.

Due to a low follow-up rate in the rest of CORC sample, we can be more sure for your service than for the rest of CORC that the findings can be generalised to all children and young people who complete the SDQ at time 1.

Have Child SDQ Total Difficulties Scores Reliably Changed compared to the Rest of CORC Sample?
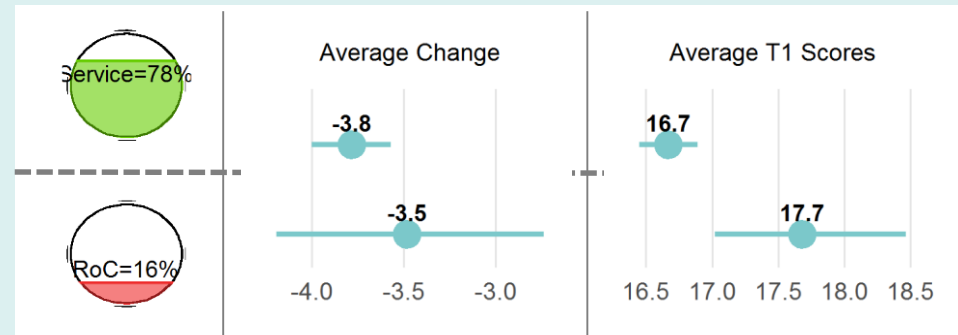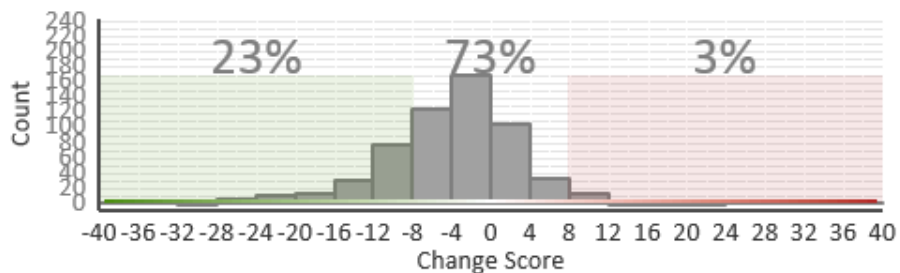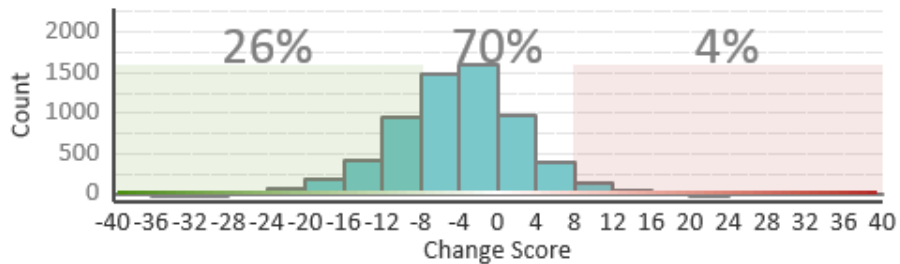
Results included

Any paired Child SDQ Total Difficulties score, with scores categorised by whether they fall above or below the threshold for reliable change. Change scores higher/lower than +/- 7.81 were classified as reliable change. Service n = 6189, Rest of CORC n = 582.

How representative is this sample?

78% (6189/7947) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 16% (582/3688).

What do the plots show?

Place2Be and the rest of CORC had similar proportions of children showing reliable improvement (26%, compared with 23%), no change (70%, compared with 73%) and reliable deterioration (4%, compared with 3%).

Conclusion

While 23-26% of children in your service and in the rest of CORC reliably improved, the majority of children with time 1 Child SDQ data did not show reliable change. This means that we cannot be certain that any observed change in scores was not due to measurement error for these cases. However, this does not mean that any observed change was not meaningful to the individual child or young person, or their family.

Note: The standard deviation used to calculate reliable change in this report is based on the full CYP IAPT sample up to Quarter 2, 2015.

How have Child SDQ Total Difficulties scores changed between T1 and T2 compared to the Rest of CORC?

	Place2Be	Rest of CORC
Stayed Above	16.8% (1039)	28.2% (164)
Improved (above to below)	28.2% (1747)	23.0% (134)
Deteriorated (below to above)	6.5% (402)	6.4% (37)
Stayed Below	48.5% (3001)	42.4% (247)
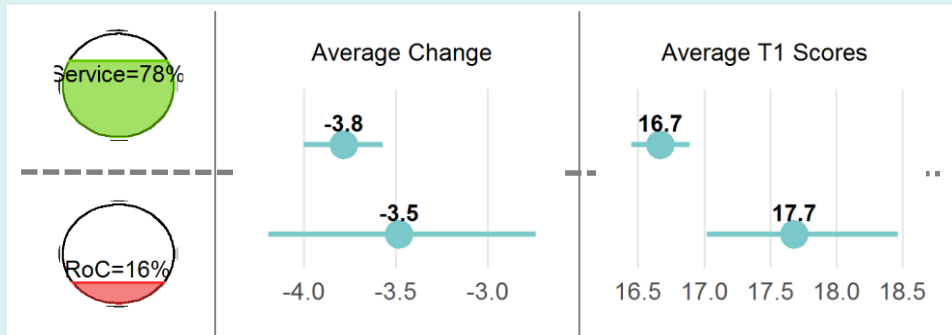
Results included

Any paired Child SDQ Total Difficulties score: Service n = 6189, Rest of CORC n = 582. Scores were classified as above cut-off if they were in the "high"/"very high" range (top 10% of scores for a community sample*), and below if they were in the "close to average"/"slightly raised" range.

How representative is this sample?

78% (6189/7947) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 16% (582/3688).

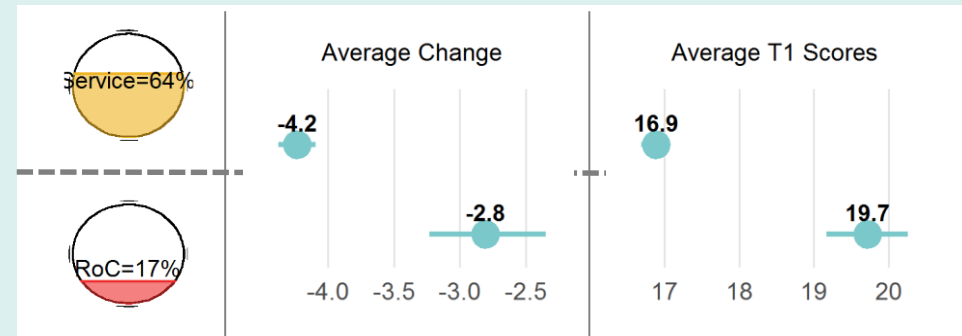**What does the table show?**

49% children and young people seen by your service had scores that were below cut-off at both time points (compared to 42% in the rest of CORC). 28% of children started above cut-off at time 1 and moved to below cut-off at time 2 (compared to 23% in the rest of CORC), while 17% remained above cut-off at both time points (compared to 28% in the rest of CORC). A small proportion (6-7%) of children in both groups declined from time 1 to time 2.

Conclusion

Compared to children in the rest of CORC, a higher proportion of children in your service either improved from time 1 to time 2 or stayed below cut-off at both time points on the Child SDQ Total Difficulties scale. A smaller proportion of children in your service stayed above cut-off at both time points compared to children in the rest of CORC.

Service=78%
RoC=16%

Average Change: -3.8, -3.5
Average T1 Scores: 16.7, 17.7

* Cut-offs are based on the 4-band categorisation of scores in the SDQ Scoring Guidelines ("Close to average", "Slightly raised", "High" and "Very high"). In the original 3-band categorisation of SDQ scores ("normal", "borderline" and "abnormal"), "abnormal" corresponds to the top 10% of a community sample. For more information, see *Instructions in English for scoring by hand SDQs for 4-17 year olds,* available at this link: http://www.sdqinfo.com/py/sdqinfo/c0.py.

CORC
Child Outcomes
Research Consortium

How have Parent SDQ Total Difficulties scores changed between T1 and T2 compared to the rest of CORC (RoC) sample?

Results included

Any paired Parent SDQ Total Difficulties scores: Service n = 11280, Rest of CORC n = 1246

How representative is this sample?

64% (11280/17659) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 17% (1246/7137).
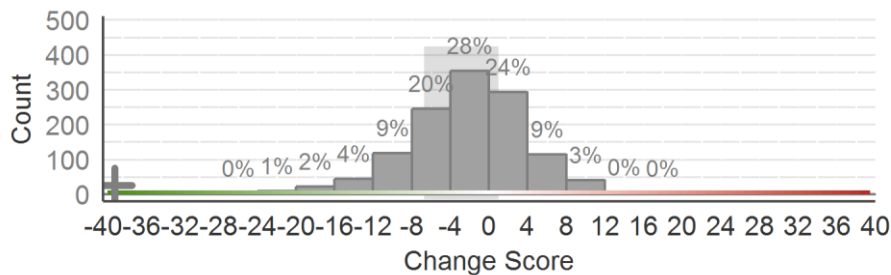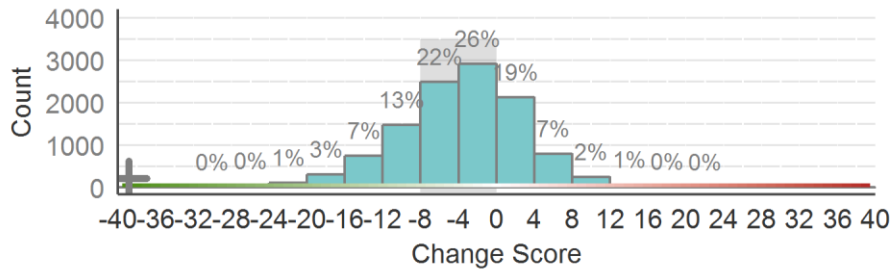
What do the plots show?

On average, scores changed by -4 points, with a tight margin of error in your service. In RoC services, the average change score was -3, between -3 and -2.

Looking at the spread of change scores, 71% (8057/11280) of scores improved in your service, and 64% (793/1246) in the RoC sample. The middle 50% of children with paired scores changed by between -8 and 0 points for your service, and between -7 and 1 for the RoC sample.

The difference between time 1 and time 2 scores has a Cohen's d effect size of -0.61 in your service and -0.38 in the RoC sample, both in the direction of progress.

Conclusion

Most children improve on the Parent SDQ in both your service and the rest of CORC (71% and 64%, respectively), however, children in your service progress more, on average, than children in the rest of CORC. Average time 1 scores indicate children in your service had a slightly lower level of starting difficulties compared to children in the rest of CORC.

Due to a low follow-up rate in the rest of CORC, we can be more sure for your service than for the rest of CORC that the findings can be generalised to all children and young people with a Parent SDQ at time 1.

Have Parent SDQ Total Difficulties Scores Reliably Changed compared to the Rest of CORC Sample?
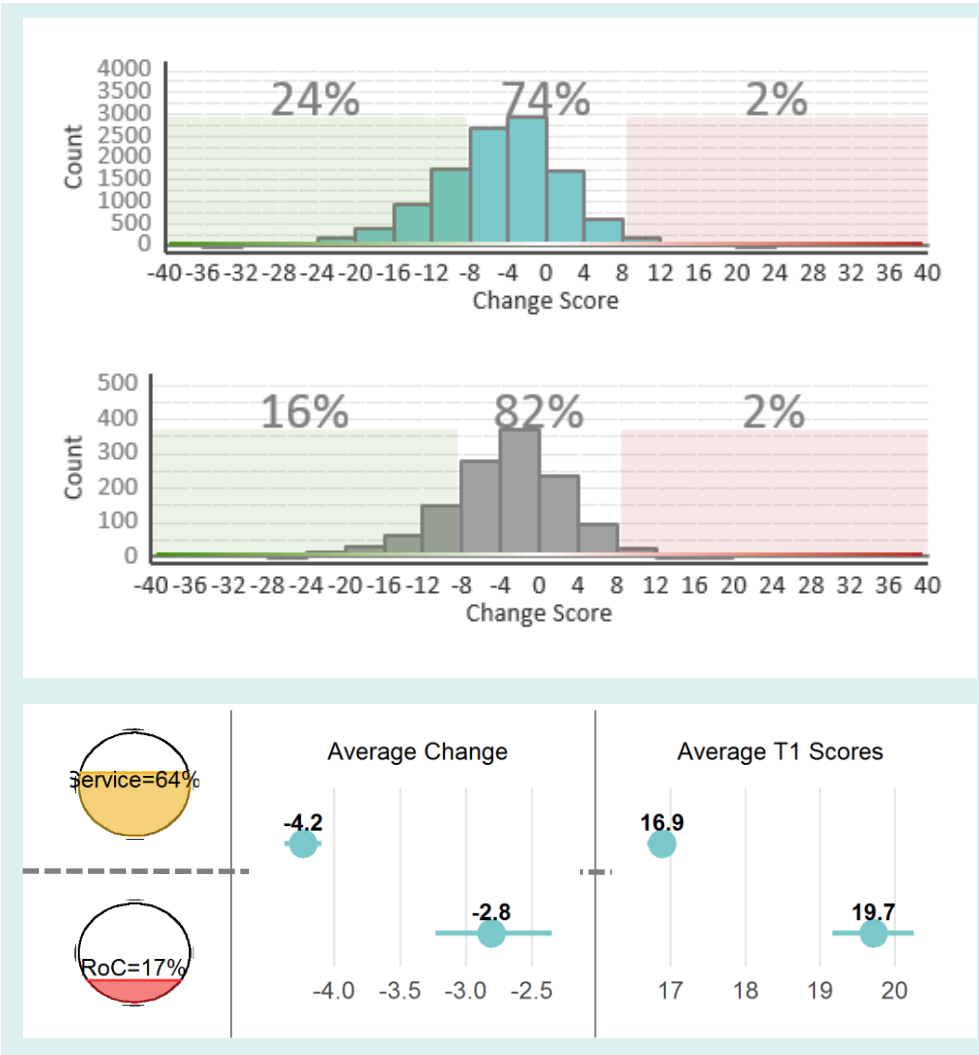
Results included

Any paired Parent SDQ Total Difficulties score, with scores categorised by whether they fall above or below the threshold for reliable change. Change scores higher/lower than +/- 8.37 were classified as reliable change. Service n = 11280, Rest of CORC n = 1246

How representative is this sample?

64% (11280/17659) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 17% (1246/7137).

What do the plots show?

Compared to the rest of CORC, more children seen by Place2Be reliably improved (24%, compared with 16%). The same percentage of children reliably deteriorated (2%) in both services. The majority of children in both services did not see any reliable change on the Parent SDQ.

Conclusion

While 24% of children in your service and 16% of children in the rest of CORC reliably improved, the majority of children with time 1 Parent SDQ data did not show reliable change. This means we cannot be certain that any observed change in scores was not due to measurement error for these cases. However, this does not mean that any observed change was not meaningful to the individual child or young person, or their family.

Note: The standard deviation used to calculate reliable change in this report is based on the full CYP IAPT sample up to Quarter 2, 2015.

How have Parent SDQ Total Difficulties scores changed between T1 and T2 compared to the Rest of CORC?

	Place2Be	Rest of CORC
Stayed Above	22.7% (2566)	46.5% (580)
Improved (above to below)	28.0% (3154)	22.1% (275)
Deteriorated (below to above)	4.6% (523)	6.0% (75)
Stayed Below	44.7% (5037)	25.4% (316)

Results included

Any paired Parent SDQ Total Difficulties score: Service n = 11280, Rest of CORC n = 1246. Scores were classified as above cut-off if they were in the "high"/"very high" range (top 10% of scores for a community sample*), and below if they were in the "close to average"/"slightly raised" range.

How representative is this sample?

64% (11280/17659) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 17% (1246/7137).
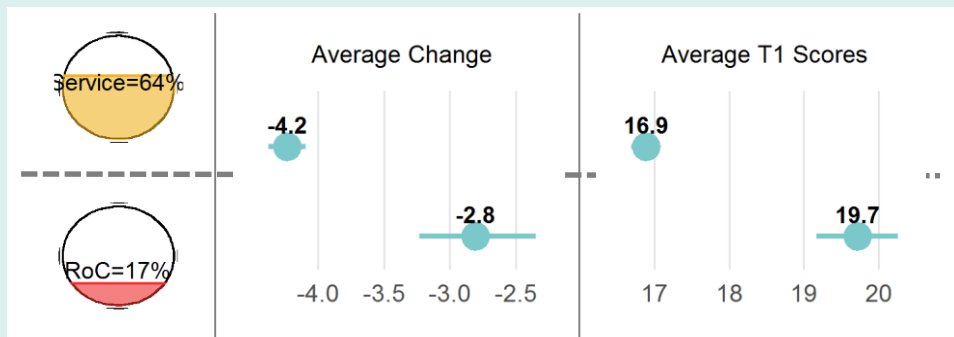
What does the table show?

45% children and young people seen by your service had scores that were below cut-off at both time points (compared to 25% in the rest of CORC). 28% of children started above cut-off at time 1 and moved to below cut-off at time 2 (compared to 22% in the rest of CORC), while 23% remained above cut-off at both time points (compared to 47% in the rest of CORC). A small proportion (5-6%) of children in both groups declined from time 1 to time 2.

Conclusion

Compared to children in the rest of CORC, a higher proportion of children in your service either improved from time 1 to time 2 or stayed below cut-off at both time points on the Parent SDQ Total Difficulties scale. A smaller proportion of children in your service stayed above cut-off at both time points compared to children in the rest of CORC.

These results may reflect the lower average starting level of difficulties parents in your service report, compared to parents in the rest of CORC.

Service=64%

RoC=17%

Average Change

-4.2

-2.8

-4.0 -3.5 -3.0 -2.5

Average T1 Scores

16.9

19.7

17   18   19   20

* Cut-offs are based on the 4-band categorisation of scores in the SDQ Scoring Guidelines ("Close to average", "Slightly raised", "High" and "Very high". In the original 3-band categorisation of SDQ scores ("normal", "borderline" and "abnormal"), "abnormal" corresponds to the top 10% of a community sample. For more information, see *Instructions in English for scoring by hand SDQs for 4-17 year olds,* available at this link: http://www.sdqinfo.com/py/sdqinfo/c0.py.

CORC
Child Outcomes
Research Consortium

Added Value Score for Parent SDQ Total Difficulties compared to the Rest of CORC

Mean Added Value Score (Effect size)

Service — 0.5

Rest of CORC — 0.1

0.0   0.2   0.4

Service = 13%   ROC = 9%
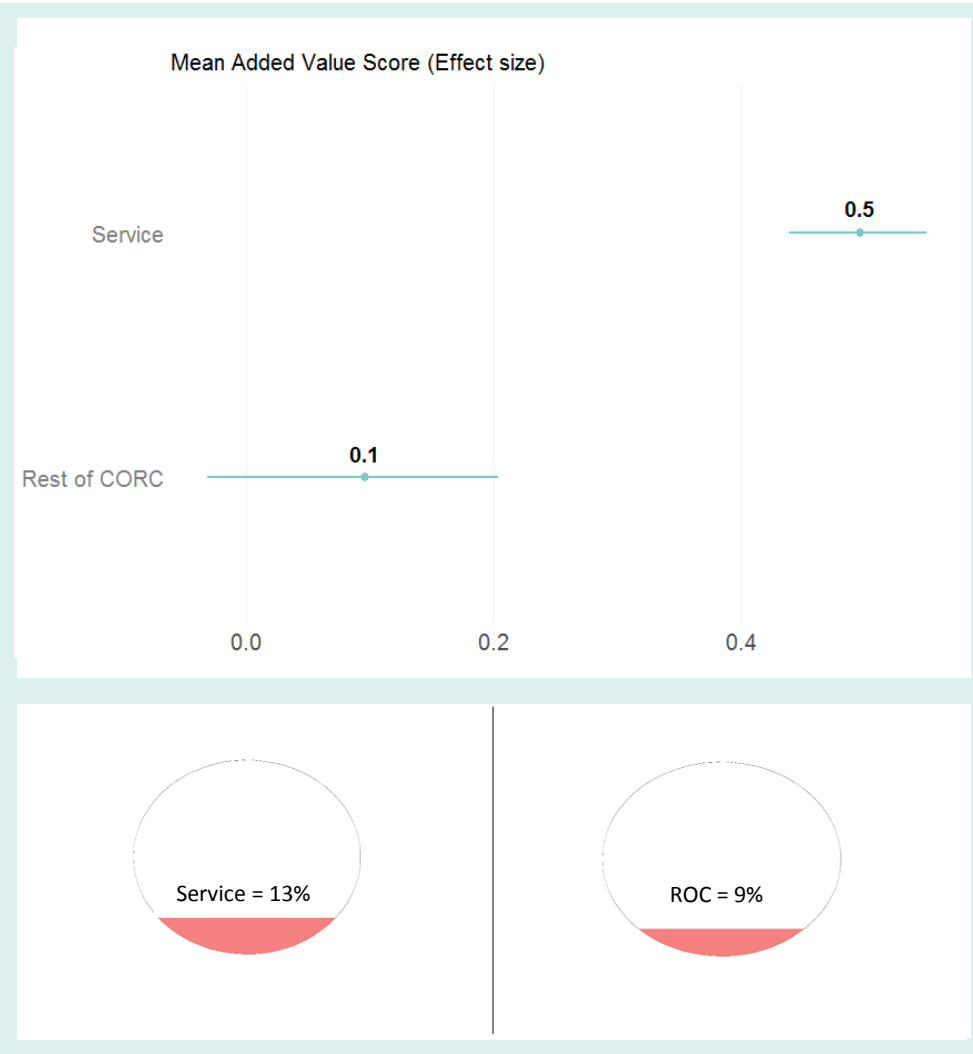
Results included

Cases meeting criteria for analysis of Added Value Scores*: all paired Parent SDQ Total Difficulties results where the time 2 SDQ was completed within 4-8 months of the time 1 SDQ. Service n = 2354, Rest of CORC n = 635.

How representative is this sample?

13% of those with a time 1 Parent SDQ score, had a corresponding time 2 score that meets the above criteria for inclusion at your service. For the RoC sample this figure was 9%. Due to the low follow-up rates in both groups for this analysis, we cannot be sure that the samples are representative of all children and young people whose parents completed an SDQ at time 1 in your service and in the rest of CORC.

What do the plots show?

Both Place2Be and the rest of CORC have a positive added value score. The score for your service (0.50, margin of error between 0.44 and 0.55) is higher than that for rest of CORC (0.10, margin of error between -0.03 and 0.20). The margin of error for the RoC sample, which crosses zero, indicates uncertainty around the score.

Conclusion

From this data, Place2Be appears to have a positive effect on the children and young people for whom parent-reported SDQ data has been collected within a 4-8 month time frame. The score for rest of CORC suggests uncertainty around the direction of impact in that sample.

We would need a higher follow-up rate within the time frame for calculating Added Value Scores to be able to generalise these results.

*The added value score is a method of comparing change in the symptoms of a child being seen by a service (as measured by the Parent SDQ Total Difficulties score) to expected change for children with similar problems generally not seen by services. This is calculated using an algorithm developed by Ford et al. (2009), and is expressed as an effect size, with effect sizes greater than 0 suggesting the service has had a **positive** effect.

CORC
Child Outcomes
Research Consortium

How have Child SDQ Impact scores changed between T1 and T2 compared to the rest of CORC (RoC) sample?
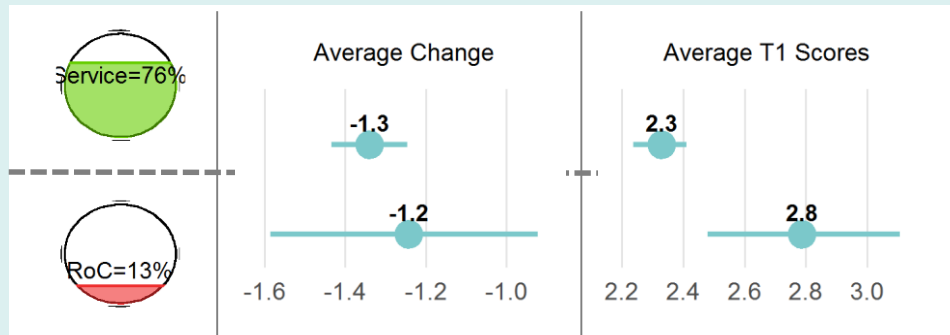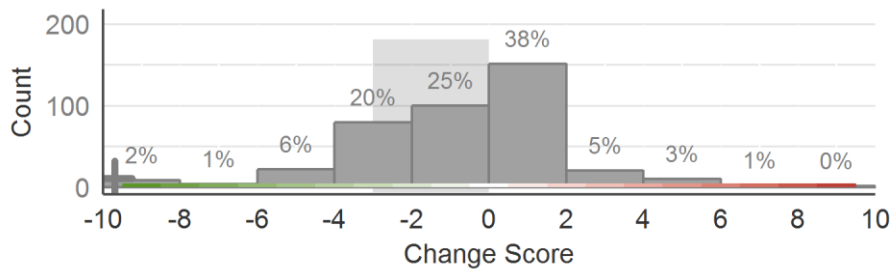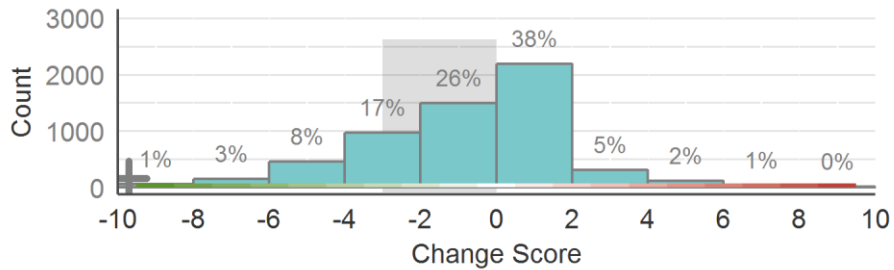
Results included

Any paired Child SDQ Impact scores: Service n = 5807, Rest of CORC n = 397

How representative is this sample?

76% (5807/7608) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 13% (397/3099).

What do the plots show?

On average, scores changed by -1 points, with a tight margin of error in your service. In RoC services, the average change score was -1, between -2 and -1.

Looking at the spread of change scores, 54% (3139/5807) of scores improved in your service, and 54% (213/397) in the RoC sample. The middle 50% of children and young people with paired scores changed by between -3 and 0 points for both samples.

The difference between time 1 and time 2 scores has a Cohen's d effect size of -0.55 in your service and -0.49 in the RoC sample, both in the direction of progress.

Conclusion

Around half of children (54%) improve on the Child SDQ Impact in both your service and the rest of CORC and children progress similarly, on average, in the two samples. Average time 1 scores indicate children in your service had a slightly lower level of starting impact compared to children in the rest of CORC.

Due to a low follow-up rate in the rest of CORC sample, we can be more sure for your service than for the rest of CORC that the findings can be generalised to all children and young people who complete the SDQ Impact at time 1.

Have Child SDQ Impact Scores Reliably Changed compared to the Rest of CORC Sample?
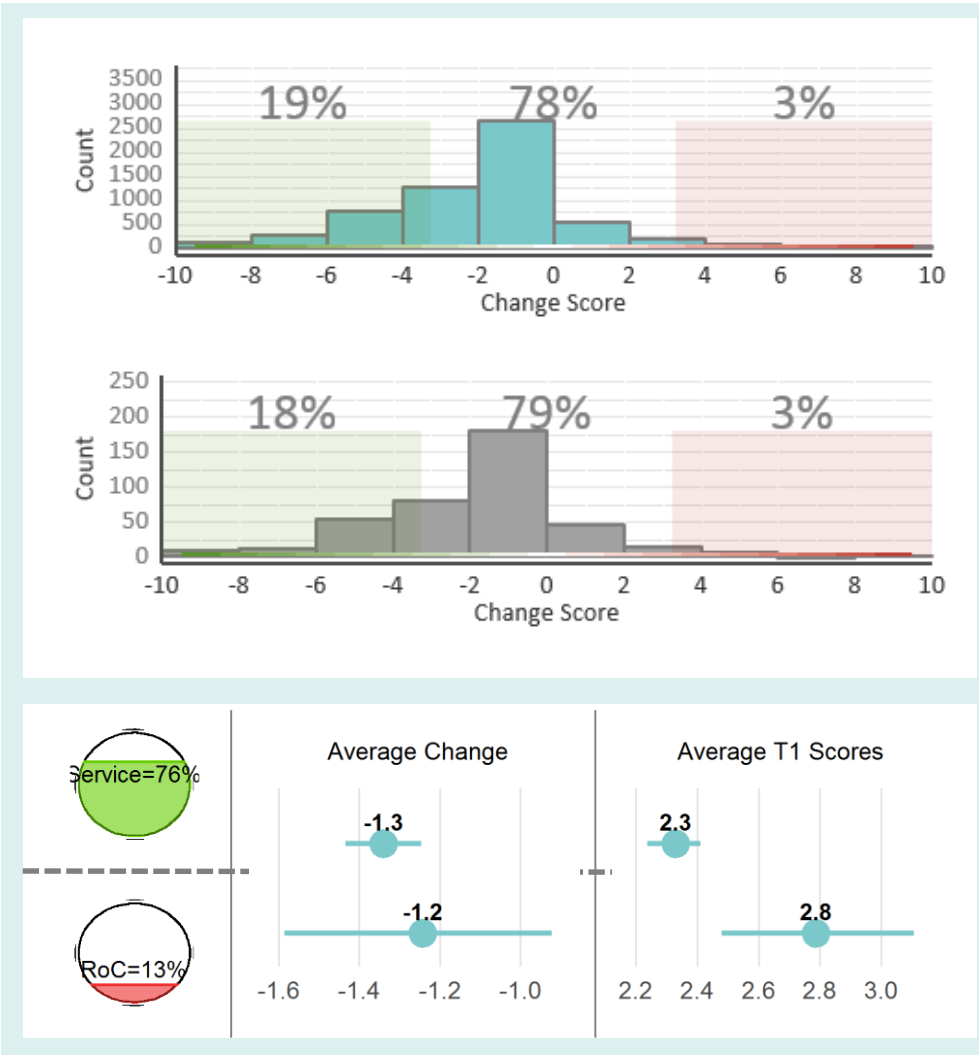
Results included

Any paired Child SDQ Total Impact score, with scores categorised by whether they fall above or below the threshold for reliable change. Change scores higher/lower than +/- 3.25 were classified as reliable change. Service n = 5807, Rest of CORC n = 397

How representative is this sample?

76% (5807/7608) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 13% (397/3099).

What do the plots show?

Place2Be and the rest of CORC had similar proportions of children showing reliable improvement (19%, compared with 18%), no change (78%, compared with 79%) and reliable deterioration (3% for both samples).

Conclusion

While 18-19% of children in both samples reliably improved, the majority of children with time 1 Child SDQ Impact data did not show reliable change. This means we cannot be certain that any observed change in scores was not due to measurement error for these cases. However, this does not mean that any observed change was not meaningful to the individual child or young person, or their family.

Note: The standard deviation used to calculate reliable change in this report is based on the full CYP IAPT sample up to Quarter 2, 2015.

How have Child SDQ Impact scores changed between T1 and T2 compared to the Rest of CORC?

	Place2Be	Rest of CORC
Stayed Above	17.5% (1016)	32.7% (130)
Improved (above to below)	36.1% (2098)	30.7% (122)
Deteriorated (below to above)	6.6% (382)	5.3% (21)
Stayed Below	39.8% (2311)	31.2% (124)
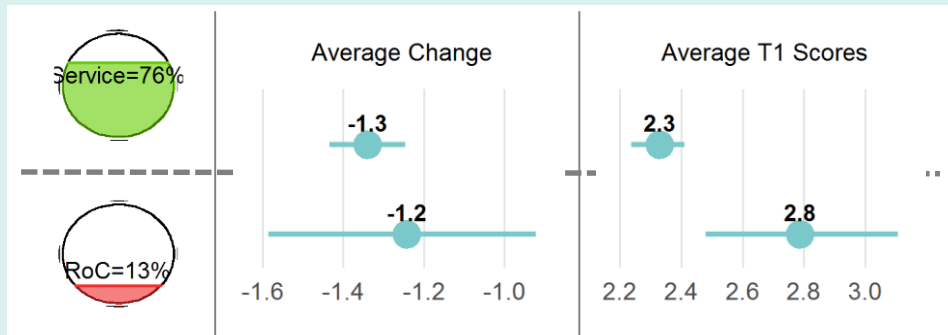
Results included

Any paired Child SDQ Total Impact score: Service n = 5807, Rest of CORC n = 397. Scores were classified as above cut-off if they were in the "high"/"very high" range (top 10% of scores for a community sample), and below if they were in the "close to average"/"slightly raised" range.

How representative is this sample?

76% (5807/7608) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 13% (397/3099).
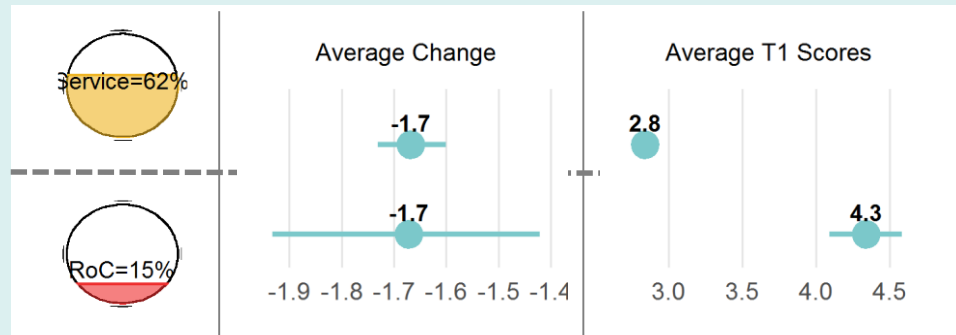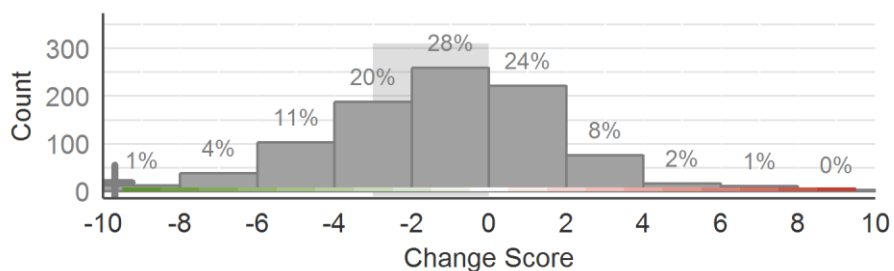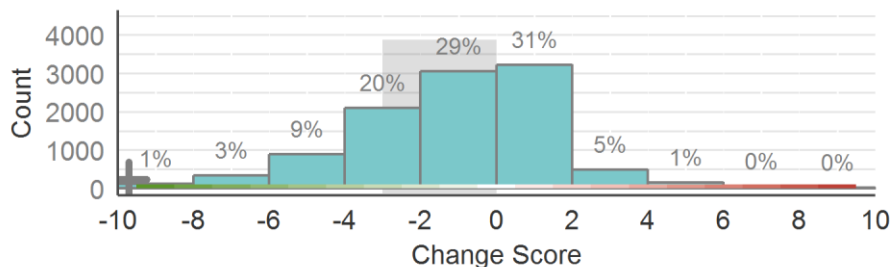
What does the table show?

40% children and young people seen by your service had scores that were below cut-off at both time points (compared to 31% in the rest of CORC). 36% of children started above cut-off at time 1 and moved to below cut-off at time 2 (compared to 31% in the rest of CORC), while 18% remained above cut-off at both time points (compared to 33% in the rest of CORC). A small proportion (5-7%) of children in both groups declined from time 1 to time 2.

Conclusion

Compared to children in the rest of CORC, a higher proportion of children in your service either improved from time 1 to time 2 or stayed below cut-off at both time points on the Child SDQ Total Impact scale. A smaller proportion of children in your service stayed above cut-off at both time points compared to children in the rest of CORC.

* Cut-offs are based on the 4-band categorisation of scores in the SDQ Scoring Guidelines ("Close to average", "Slightly raised", "High" and "Very high". In the original 3-band categorisation of SDQ scores ("normal", "borderline" and "abnormal"), "abnormal" corresponds to the top 10% of a community sample. For more information, see *Instructions in English for scoring by hand SDQs for 4-17 year olds,* available at this link: http://www.sdqinfo.com/py/sdqinfo/c0.py.

How have Parent SDQ Impact scores changed between T1 and T2 compared to the rest of CORC (RoC) sample?

Results included

Any paired Parent SDQ Impact scores: Service n = 10409, Rest of CORC n = 921

How representative is this sample?

62% (10409/16816) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 15% (921/6080).

What do the plots show?

On average, scores changed by -2 points, with a tight margin of error in your service. In RoC services, the average change score was -2, between -2 and -1.

Looking at the spread of change scores, 62% (6495/10409) of scores improved in your service, and 65% (597/921) in the RoC sample. The middle 50% of children and young people with paired scores changed by between -3 and 0 points in both samples.

The difference between time 1 and time 2 scores has a Cohen's d effect size of -0.68 in your service and -0.61 in the rest of CORC sample, both in the direction of progress.

Conclusion

More than half of children in your service and in the rest of CORC improve on the Parent SDQ Impact (62% and 65%, respectively), and children progress similarly, on average, in the two samples. Average time 1 scores indicate children in your service had a lower level of starting impact compared to children in the rest of CORC.

Due to a low follow-up rate in the rest of CORC sample, we can be more sure for your service than for the rest of CORC that the findings can be generalised to all children and young people whose parents complete the SDQ Impact at time 1.

Have Parent SDQ Impact Scores Reliably Changed compared to the Rest of CORC Sample?
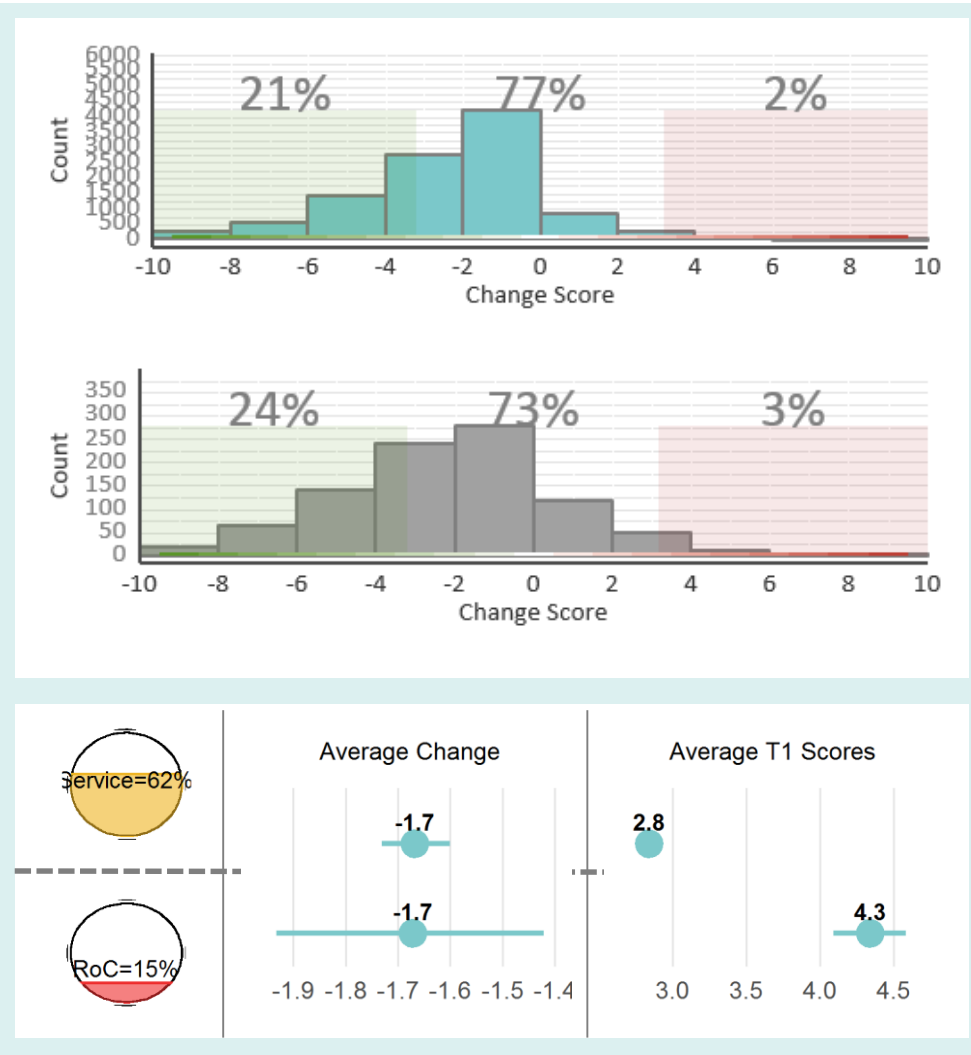
Results included

Any paired Parent SDQ Total Impact score, with scores categorised by whether they fall above or below the threshold for reliable change. Change scores higher/lower than +/- 3.19 were classified as reliable change. Service n = 10409, Rest of CORC n = 921

How representative is this sample?

62% (10409/16816) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 15% (921/6080).

What do the plots show?

Place2Be and the rest of CORC had similar proportions of children showing reliable improvement (21%, compared with 24%), no change (77%, compared with 73%) and reliable deterioration (2%, compared with 3%).

Conclusion

While 21-24% of children in both samples reliably improved, the majority of children with time 1 Parent SDQ Impact data did not show reliable change. This means we cannot be certain that any observed change in scores was not due to measurement error for these cases. However, this does not mean that any observed change was not meaningful to the individual child or young person, or their family.

Note: The standard deviation used to calculate reliable change in this report is based on the full CYP IAPT sample up to Quarter 2, 2015.

How have Parent SDQ Impact scores changed between T1 and T2 compared to the Rest of CORC?

	Place2Be	Rest of CORC
Stayed Above	23.5% (2444)	49.7% (458)
Improved (above to below)	40.8% (4243)	33.1% (305)
Deteriorated (below to above)	4.2% (439)	4.9% (45)
Stayed Below	31.5% (3283)	12.3% (113)
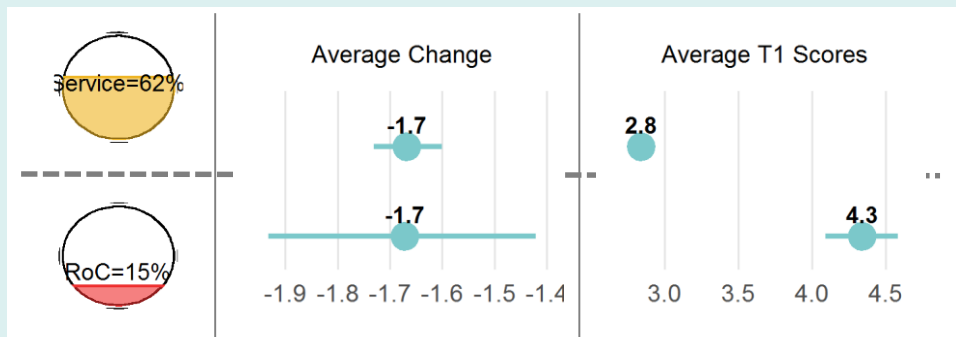
Results included

Any paired Parent SDQ Total Impact score: Service n = 10409, Rest of CORC n = 921. Scores were classified as above cut-off if they were in the "high"/"very high" range (top 10% of scores for a community sample), and below if they were in the "close to average"/"slightly raised" range.

How representative is this sample?

62% (10409/16816) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 15% (921/6080).

What does the table show?

32% children and young people seen by your service had scores that were below cut-off at both time points (compared to 12% in the rest of CORC). 41% of children started above cut-off at time 1 and moved to below cut-off at time 2 (compared to 33% in the rest of CORC), while 24% remained above cut-off at both time points (compared to 50% in the rest of CORC). A small proportion (4-5%) of children in both groups declined from time 1 to time 2.

Conclusion

Compared to children in the rest of CORC, a higher proportion of children in your service either improved from time 1 to time 2 or stayed below cut-off at both time points on the Parent SDQ Total Impact scale. A smaller proportion of children in your service stayed above cut-off at both time points compared to children in the rest of CORC.

These results may reflect the lower average starting level of impact parents in your service report, compared to parents in the rest of CORC.

* Cut-offs are based on the 4-band categorisation of scores in the SDQ Scoring Guidelines ("Close to average", "Slightly raised", "High" and "Very high"). In the original 3-band categorisation of SDQ scores ("normal", "borderline" and "abnormal"), "abnormal" corresponds to the top 10% of a community sample. For more information, see *Instructions in English for scoring by hand SDQs for 4-17 year olds,* available at this link: http://www.sdqinfo.com/py/sdqinfo/c0.py.

Appendix A

Technical Glossary

CORC
Child Outcomes
Research Consortium

Technical Glossary

Term	Definition	Methodology
Paired Outcome Rate	The percentage of children and young people who have **any** paired outcome measure.	The number of individuals with at least one paired outcome measure (completed by child, parent or clinician) is divided by the total number of individuals in the data submitted.
Follow up rate	**For each** outcome measure, out of those with a time 1 measure, the percentage of those with a time 2.	The number of individuals with both a time 1 and time 2 measure is divided by the number of those with a time 1 measure, for each outcome measure separately.
Margin of Error	A confidence interval: a numeric interval around an estimated number (for example, the mean), which contains the mean with a certain level of confidence.	Confidence intervals are calculated at the 99% level, using the 'basic' non-parametric bootstrap method.

CORC
Child Outcomes
Research Consortium

Appendix B

Reference Tables

Completion rates and counts

Data Counts

Sample	Time	Child SDQ Total Difficulties	Parent SDQ Total Difficulties	Child SDQ Impact	Parent SDQ Impact
Service	1	7947	17659	7608	16816
Service	2	6189	11280	5807	10409
Rest of CORC	1	3688	7137	3099	6080
Rest of CORC	2	582	1246	397	921

Corresponds to slide: What are the Follow Up Rates?

Follow Up Rates

Sample	Child SDQ Total Difficulties	Parent SDQ Total Difficulties	Child SDQ Impact	Parent SDQ Impact
Service	78%	64%	76%	62%
Rest of CORC	16%	17%	13%	15%

Corresponds to slide: What are the Follow Up Rates?

Cases with any paired outcome measure

Sample	Percentage	Count
Service	66.54%	13000
RoC	30.69%	2430

Corresponds to slide: How many paired outcomes were present?

Time 1 and Time 2 Means and Confidence Intervals

Means and confidence intervals: Service

Measure	Time	Mean	Lower limit	Upper limit
Child SDQ Total Difficulties	1	16.7	16.4	16.9
Child SDQ Total Difficulties	2	12.9	12.7	13.1
Parent SDQ Total Difficulties	1	16.9	16.7	17.0
Parent SDQ Total Difficulties	2	12.6	12.5	12.8
Child SDQ Impact	1	2.3	2.2	2.4
Child SDQ Impact	2	1.0	0.9	1.0
Parent SDQ Impact	1	2.8	2.8	2.9
Parent SDQ Impact	2	1.2	1.1	1.2

Corresponds to slides on individual outcome measures

Means and confidence intervals: Rest of CORC

Measure	Time	Mean	Lower limit	Upper limit
Child SDQ Total Difficulties	1	17.7	17.0	18.5
Child SDQ Total Difficulties	2	14.2	13.3	15.0
Parent SDQ Total Difficulties	1	19.7	19.2	20.3
Parent SDQ Total Difficulties	2	16.9	16.2	17.6
Child SDQ Impact	1	2.8	2.5	3.1
Child SDQ Impact	2	1.5	1.3	1.8
Parent SDQ Impact	1	4.3	4.1	4.6
Parent SDQ Impact	2	2.7	2.4	2.9

Corresponds to slides on individual outcome measures

Change Score Means and Confidence Intervals

Means and confidence intervals: Service

Measure	Mean	Lower limit	Upper limit
Child SDQ Total Difficulties	-3.8	-4.0	-3.6
Parent SDQ Total Difficulties	-4.2	-4.4	-4.1
Child SDQ Impact	-1.3	-1.4	-1.2
Parent SDQ Impact	-1.7	-1.7	-1.6

Corresponds to slides on individual outcome measures

**Means and confidence intervals: Rest of CORC**

Measure	Mean	Lower limit	Upper limit
Child SDQ Total Difficulties	-3.5	-4.2	-2.7
Parent SDQ Total Difficulties	-2.8	-3.2	-2.3
Child SDQ Impact	-1.2	-1.6	-0.9
Parent SDQ Impact	-1.7	-1.9	-1.4

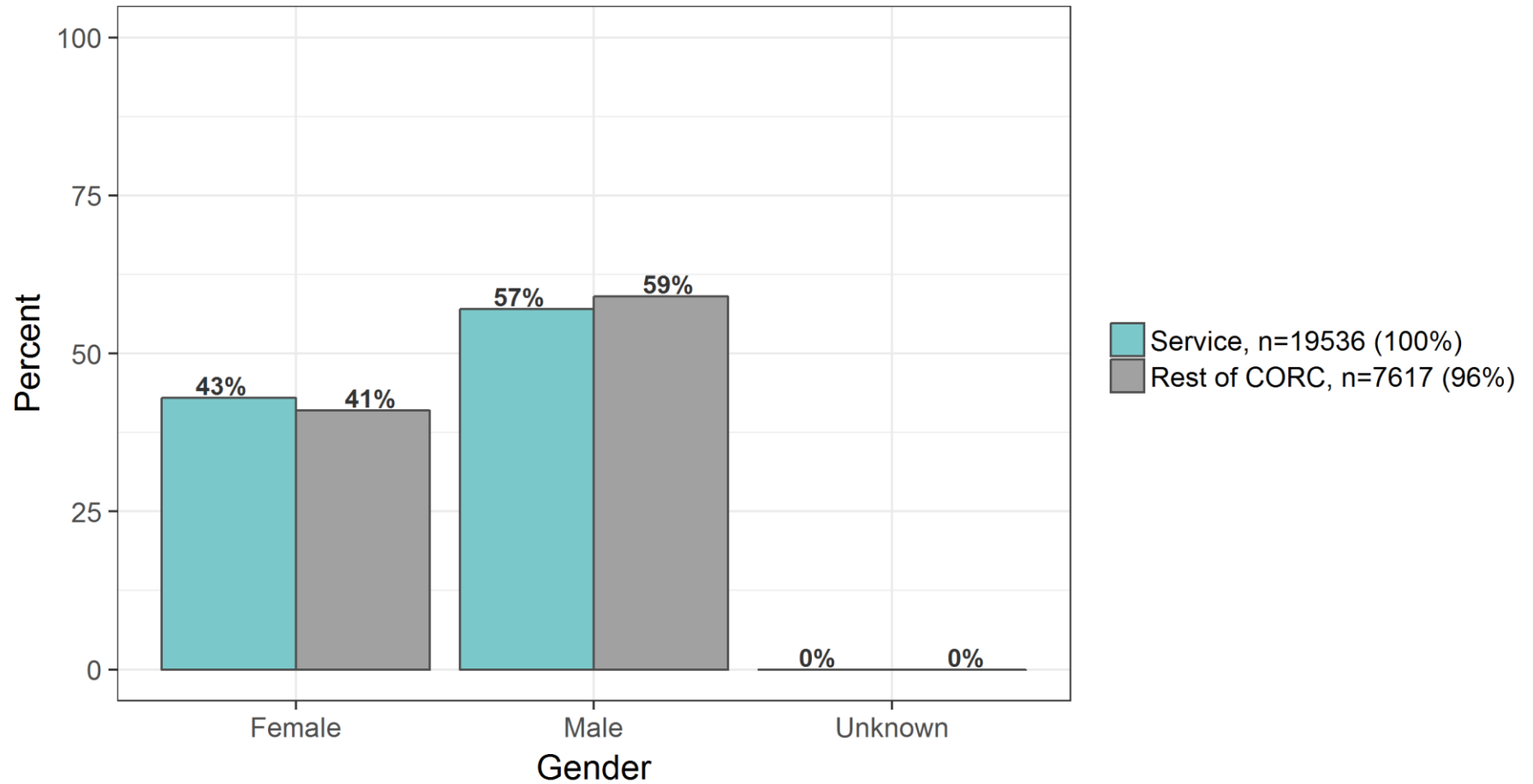*Corresponds to slides on individual outcome measures*

Appendix C

Demographics & Case Characteristics

CORC
Child Outcomes
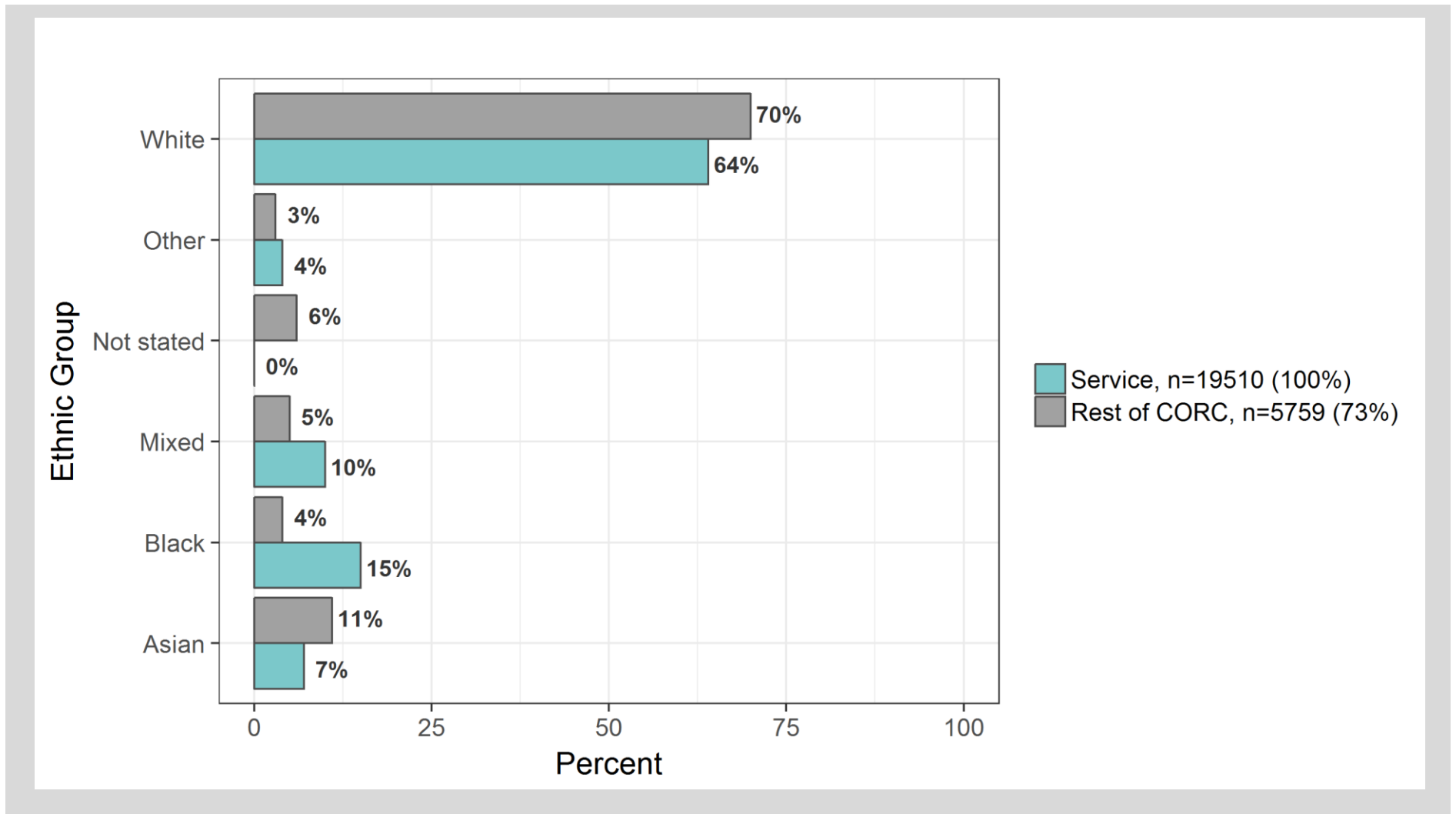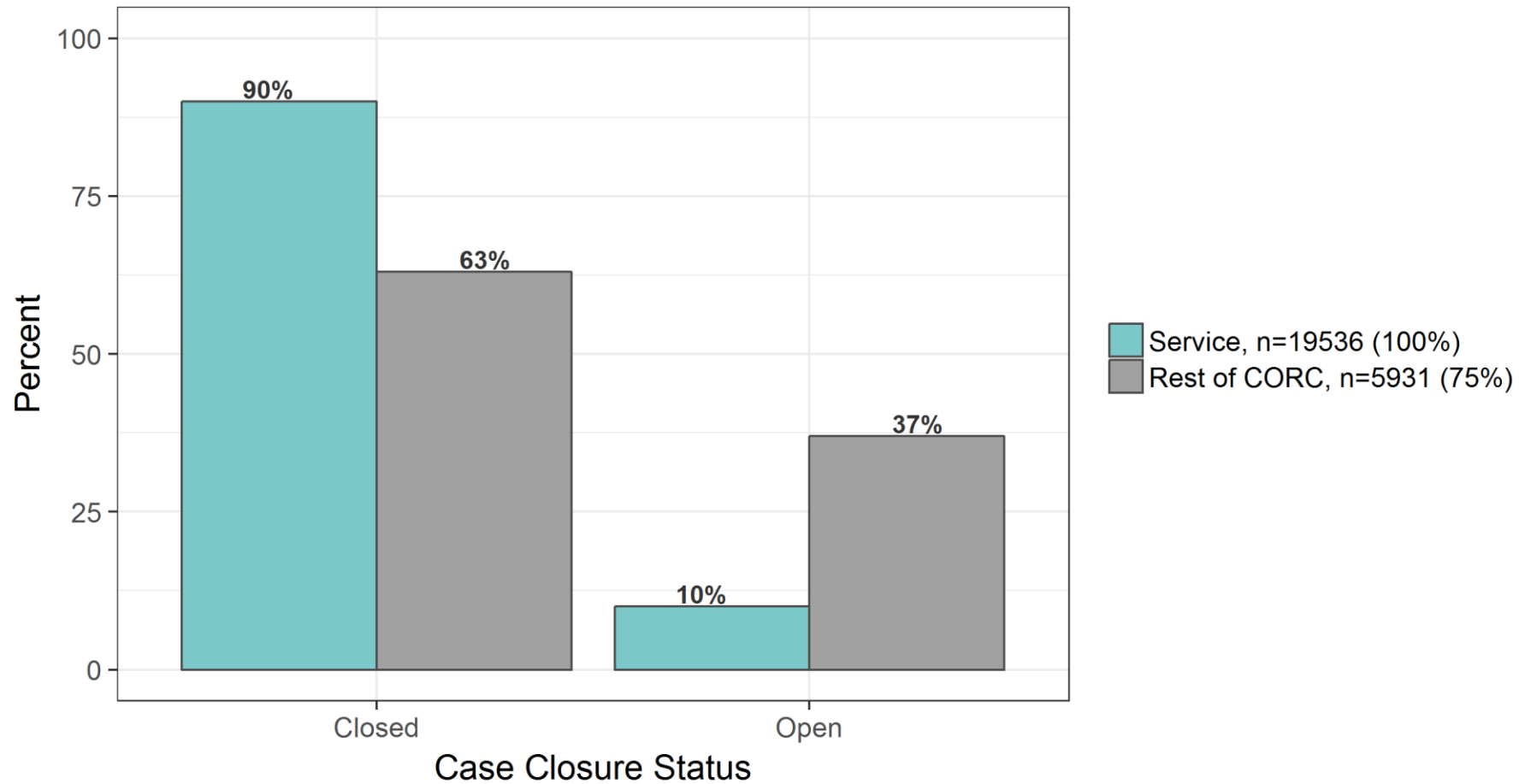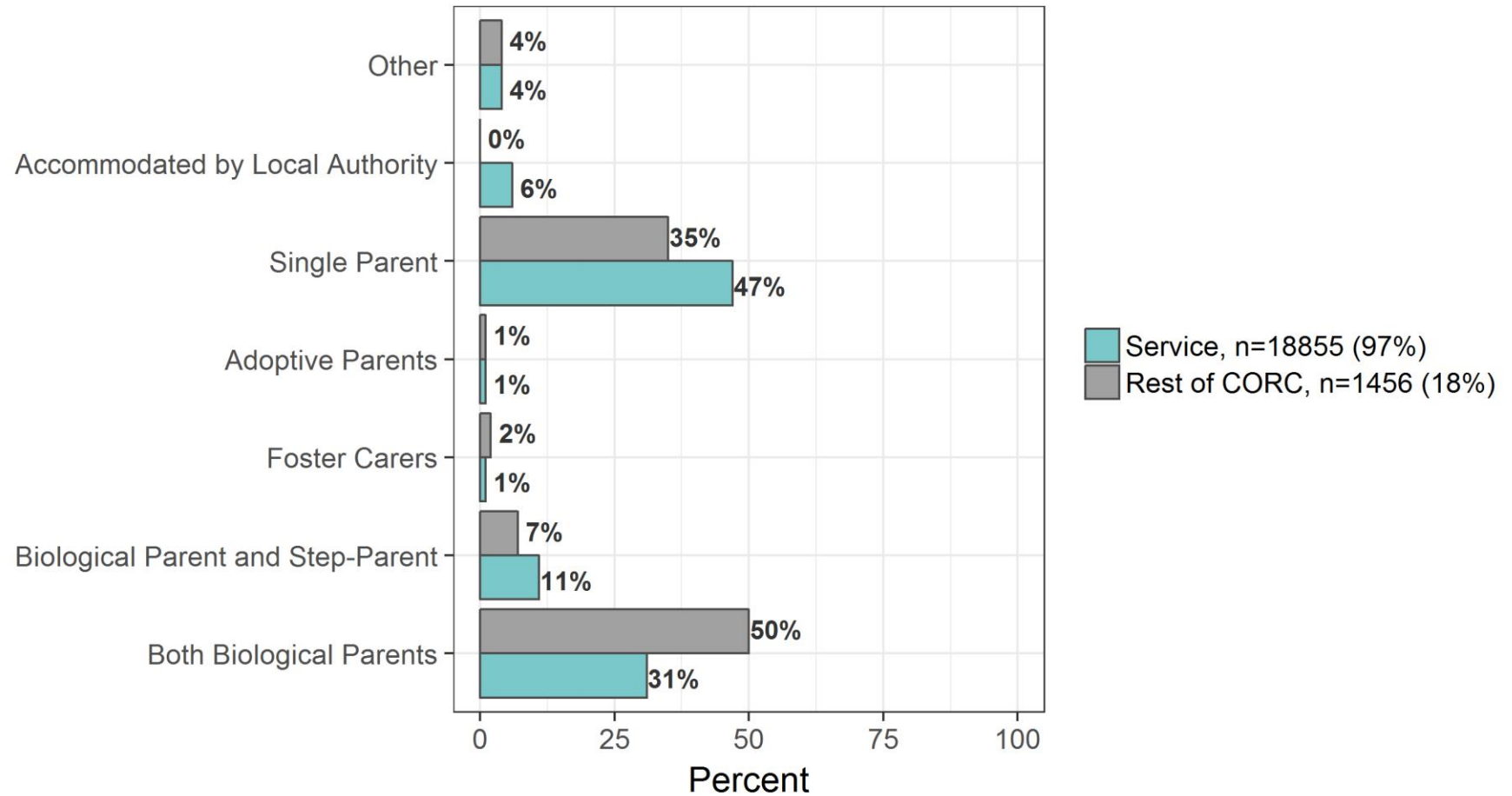Research Consortium

Age Bands
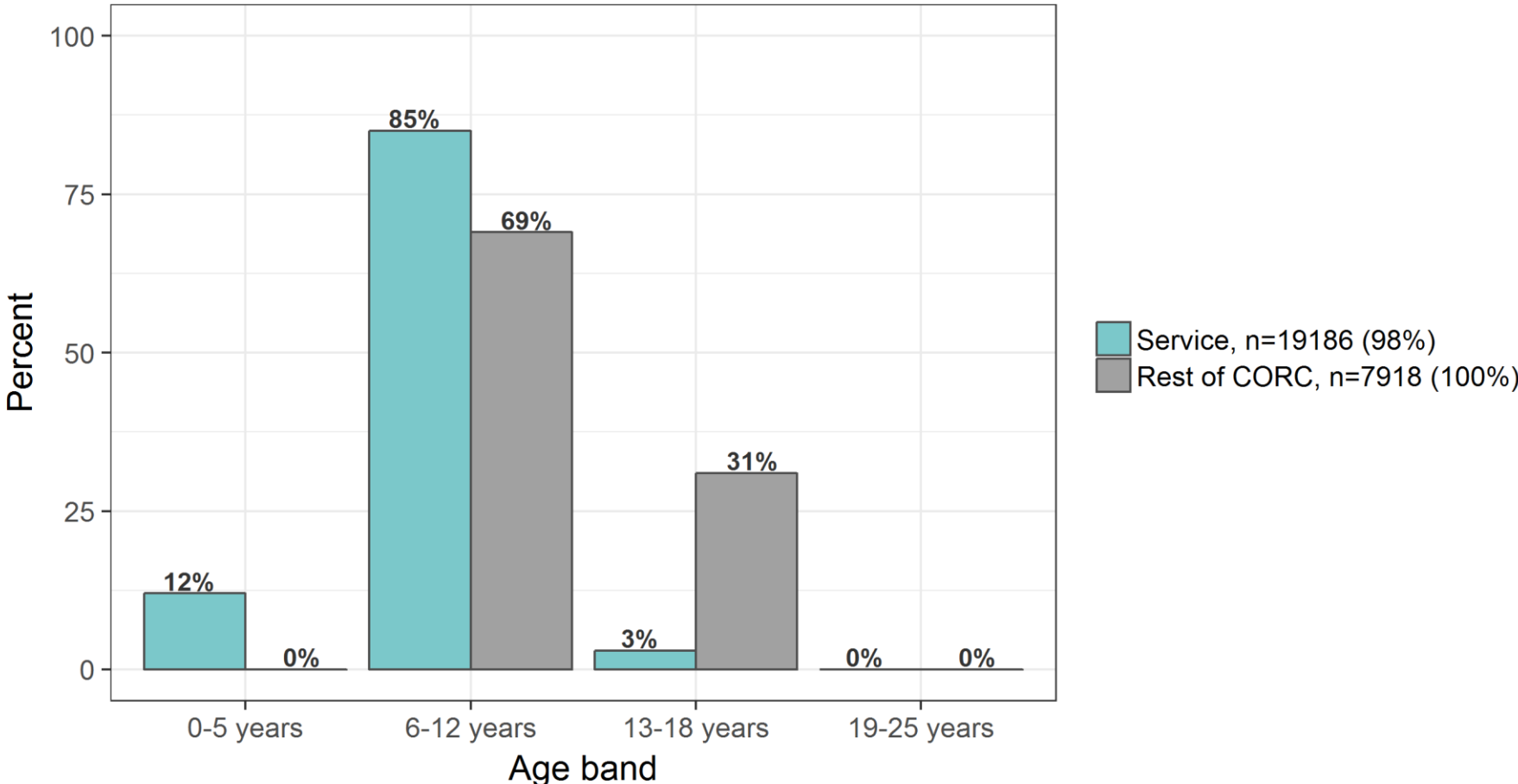
Gender

Ethnicity

Significance testing on this variable requested by service. Using the Chi-Square test, there is a statistically significant difference in ethnicity (White vs. all other) between the two groups (p<0.001). The absolute difference in proportions is 6%.

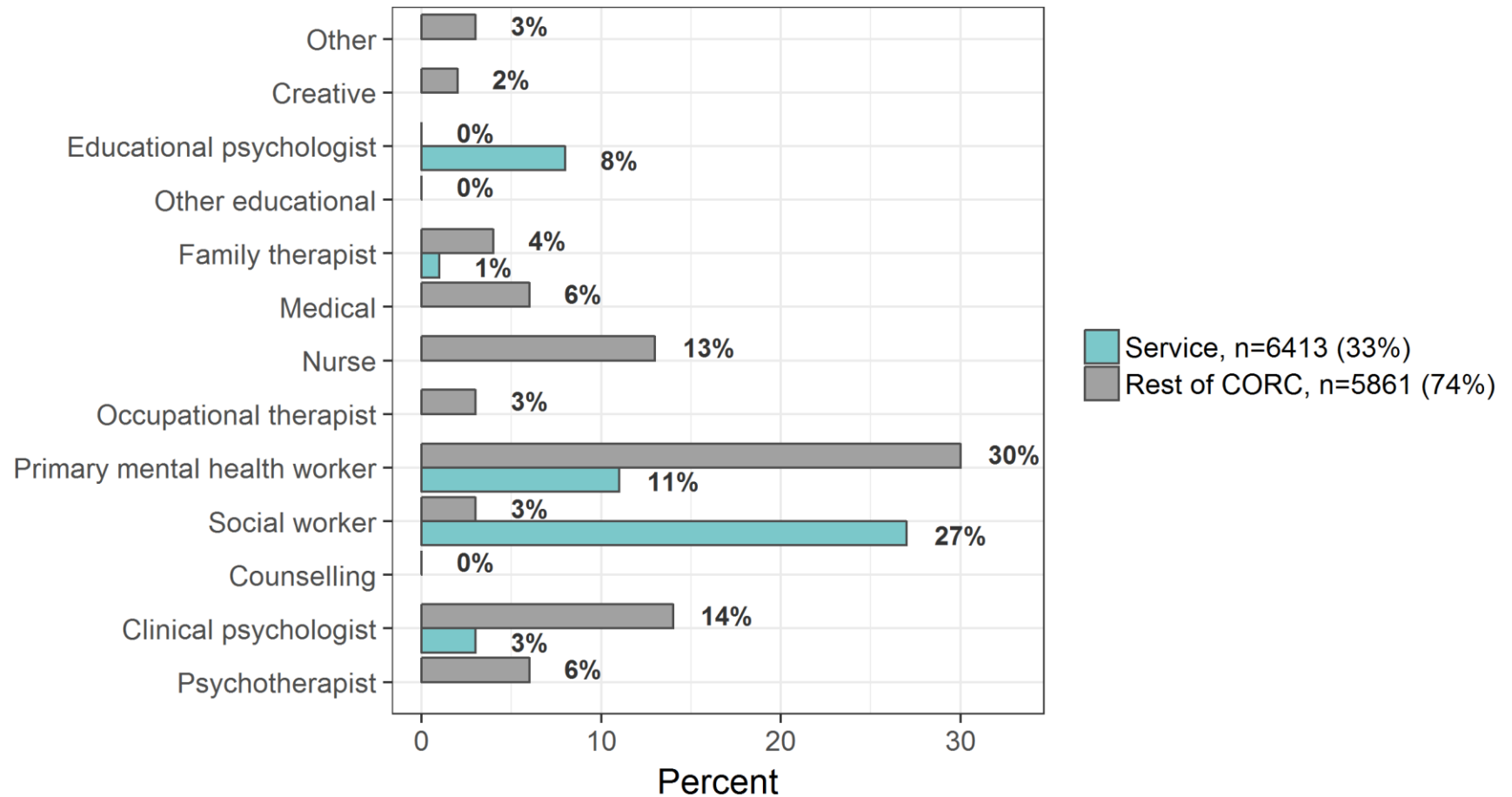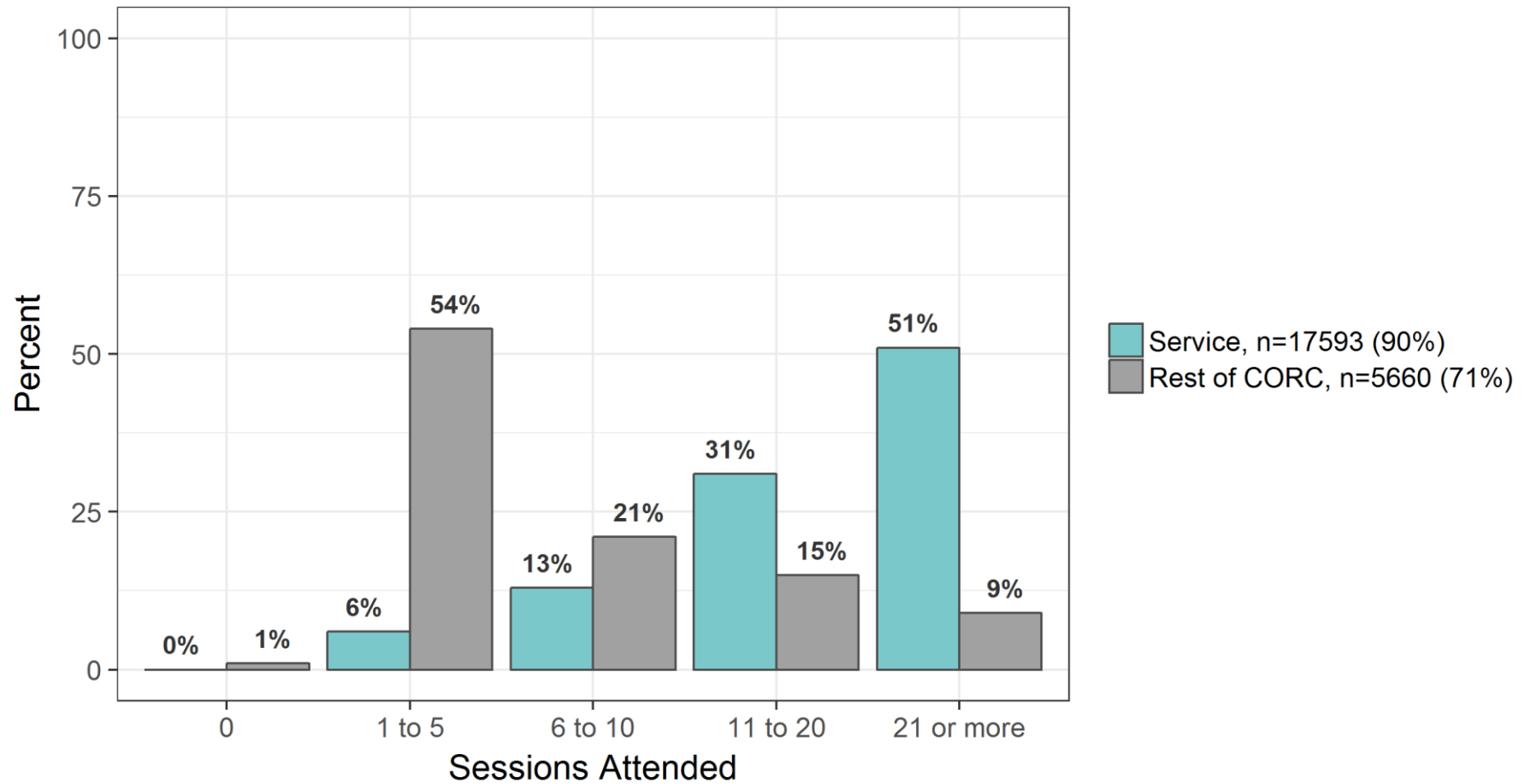# Case Closure Status

Family Composition

Type of Professionals

Sessions Attended

Child Protection Registration
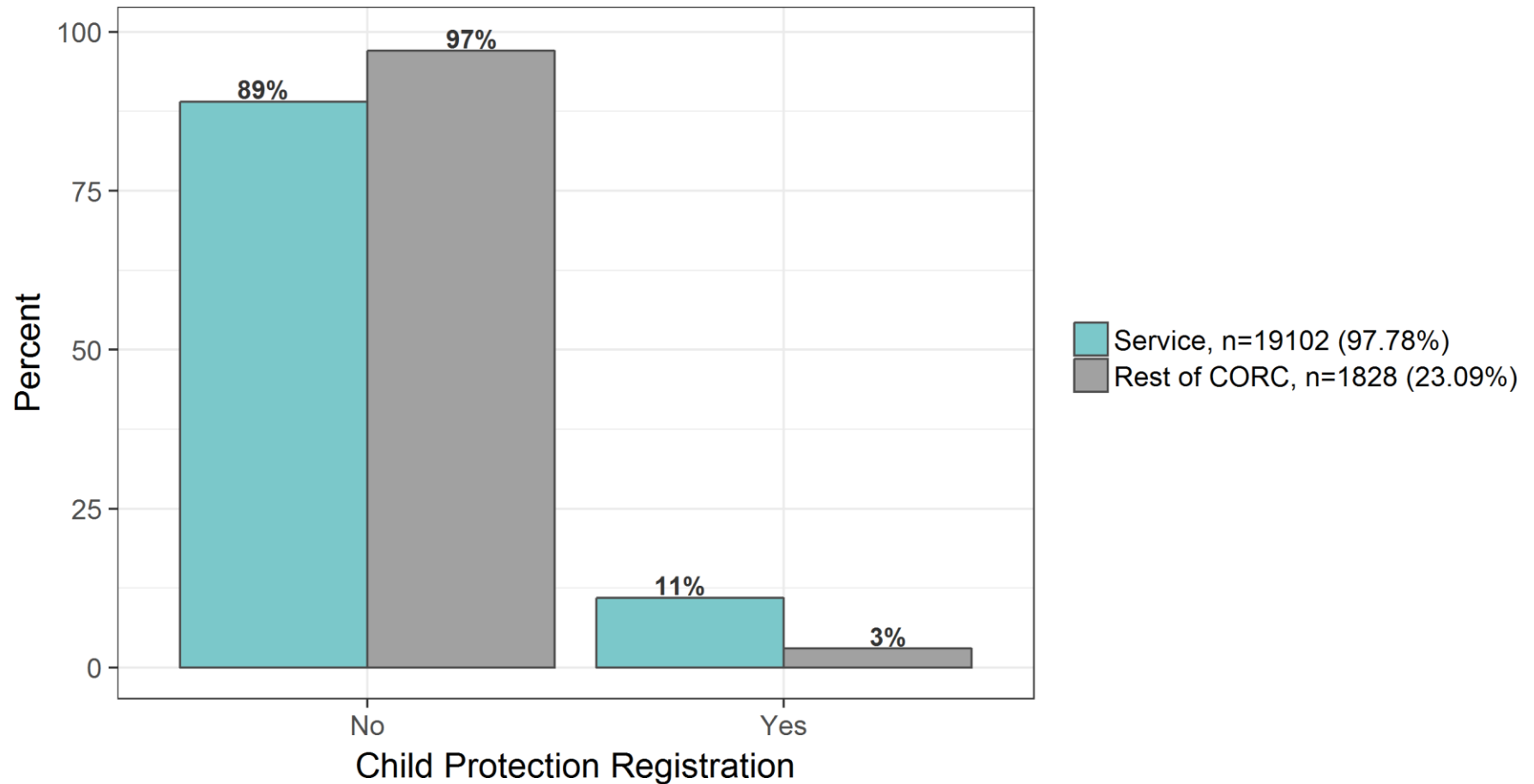
Significance testing on this variable requested by service. Using the Chi-Square test, there is a statistically significant difference in the number of children registered with Child Protection between the two groups (p<0.001). The absolute difference in proportions is 8%.

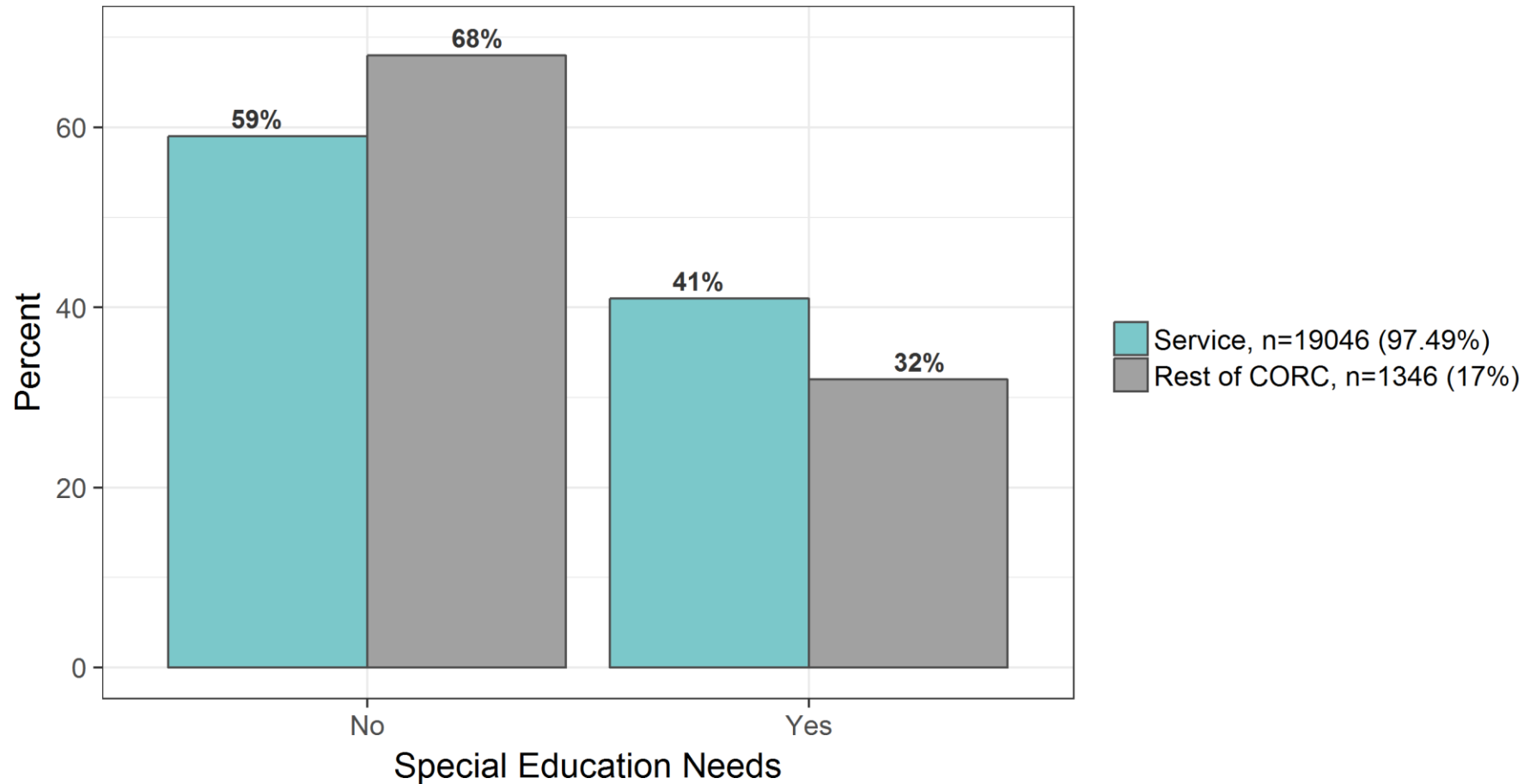# Special Education Needs

Significance testing on this variable requested by service. Using the Chi-Square test, there is a statistically significant difference in the number of children with Special Education Needs between the two groups (p<0.001). The absolute difference in proportions is 9%.

Appendix D

"Recovery" and Reliable Change (Wolpert
et al. 2016)

CORC
Child Outcomes
Research Consortium

"Recovery" and Reliable Change (Wolpert et al. 2016)

Child Reported Measures

Indicator	Definition	N	% of paired clinical sample [95% Margins of Error]
"Recovery"	Moved from above a clinical threshold on at least one paired measure at a first time point, to below on all completed measures at a last time point	2117	36% [35% - 37%]
Reliable Improvement	Change from a first to a last time point was more than what would be expected due to measurement error, in a positive direction, on at least one measure, and no measure reliably deteriorated	3056	52% [51% - 53%]
No Reliable Change	Change from a first to a last time point was less than what would be expected due to measurement error	2223	38% [36% - 39%]
Reliable Deterioration	Change from a first to a last time point was more than what would be expected due to measurement error, in a negative direction, on at least one measure	617	11% [9% - 12%]
Reliable "Recovery"	Moved from above a clinical threshold on at least one paired measure at a first time point, to below on all completed measures at a last time point, and the change was reliable in a positive direction, with no measures reliably deteriorating	1569	27% [25% - 28%]

Sample included

The results are those reported in Wolpert et al., (2016). Any closed cases with three or more recorded events, and above a clinical threshold on at least one paired child-reported measure at a first time point (the 'paired clinical sample'), n = 5,896.

How representative is this sample?

Of all closed treatment cases (n= 23,373), 25% fell in the 'paired clinical sample'. This means we cannot be confident this sample represents all children and young people who were seen for a course of treatment (of at least three events, excluding assessment only).

What does the table show?

Scores for 36% (margin of error between 35% and 37%) of children and young people showed "recovery", 52% (margin of error between 51% and 53%) showed reliable improvement, and 27% (margin of error between 25% and 28%) showed reliable "recovery".

Wolpert, M., Jacob, J., Napoleone, E., Whale, A., Calderon, A., & Edbrooke-Childs, J. (2016). *Child- and Parent-reported Outcomes and Experience from Child and Young People's Mental Health Services 2011-2015*. London: CAMHS Press

CORC
Child Outcomes Research Consortium

"Recovery" and Reliable Change (Wolpert et al. 2016)

Parent Reported Measures

Indicator	Definition	N	% of paired clinical sample [95% Margins of Error]
"Recovery"	Moved from above a clinical threshold on at least one paired measure at a first time point, to below on all completed measures at a last time point	1038	28% [26% - 29%]
Reliable Improvement	Change from a first to a last time point was more than what would be expected due to measurement error, in a positive direction, on at least one measure, and no measure reliably deteriorated	1503	41% [39% - 42%]
No Reliable Change	Change from a first to a last time point was less than what would be expected due to measurement error	1878	51% [49% - 52%]
Reliable Deterioration	Change from a first to a last time point was more than what would be expected due to measurement error, in a negative direction, on at least one measure	326	9% [7% - 11%]
Reliable "Recovery"	Moved from above a clinical threshold on at least one paired measure at a first time point, to below on all completed measures at a last time point, and the change was reliable in a positive direction, with no measures reliably deteriorating	609	16% [15% - 18%]

Sample included

The results are those reported in Wolpert et al., (2016). Any closed cases with three or more recorded events, and above a clinical threshold on at least one paired child-reported measure at a first time point (the 'paired clinical sample'), n = 3,707.

How representative is this sample?

Of all closed treatment cases (n= 23,373), 16% fell in the 'paired clinical sample'. This means we cannot be confident this sample represents all children and young people who were seen for a course of treatment (of at least three events, excluding assessment only).

What does the table show?

Scores for 28% (margin of error between 26% and 29%) of children and young people showed "recovery", 41% (margin of error between 39% and 42%) showed reliable improvement, and 16% (margin of error between 15% and 18%) showed reliable "recovery".

Wolpert, M., Jacob, J., Napoleone, E., Whale, A., Calderon, A., & Edbrooke-Childs, J. (2016). *Child- and Parent-reported Outcomes and Experience from Child and Young People's Mental Health Services 2011-2015*. London: CAMHS Press

Appendix E

References for Outcome measures and
Analysis

CORC
Child Outcomes
Research Consortium

Web Links to Outcome Measures

Strengths and Difficulties Questionnaire (SDQ):
http://www.sdqinfo.com/py/sdqinfo/c0.py

For additional information and resources on outcome measures,
visit the CORC website:
www.corc.uk.net

References for Additional Analysis

Reliable Change

Reliable change figures were calculated using the Reliable Change Index (RCI) formula taken from:

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12-19.

Reliability values used in the RCI formula taken from:

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *J. Am. Acad. Child Adolesc. Psychiatry, 40*(11), 1337-1346.

Cut-Offs

Values taken from: SDQ Scoring Guide

Added Value Score

Calculated using the added value score formula taken from:

Ford, T., Hutchings, J., Bywater, T., Goodman, A., & Goodman, R. (2009). Strengths and Difficulties Questionnaire Added Value Scores: Evaluating effectiveness in child mental health interventions. *The British Journal of Psychiatry, 194*(6), 552-558. doi:10.1192/bjp.bp.108.052373

Appendix F

The following slide can be used to display a summary of your data returns to service users

CORC
Child Outcomes
Research Consortium

OUR DATA

We collect information with your permission, we then take off your name and pass it to CORC, a research team who look at how to improve our service, national services and inform national policy. Your data makes a big difference to us, the general population and others with mental health difficulties.

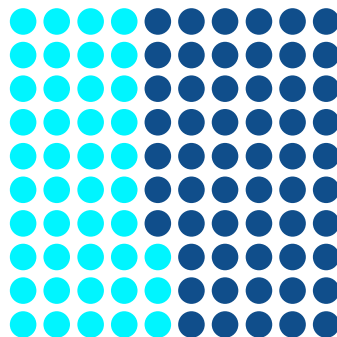**19536 people were seen by the service**

If the service saw 100 people:
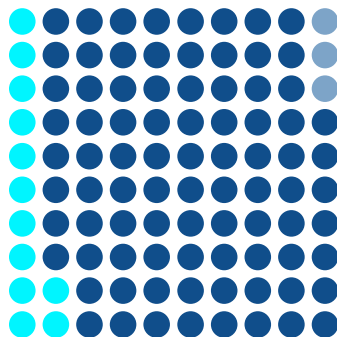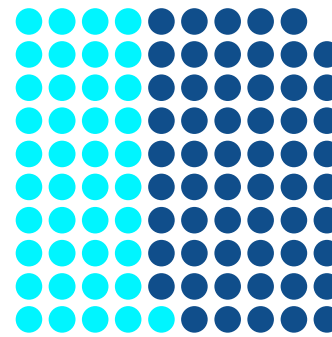
Age
People would be:

Gender
People would be:

Ethnicity
People would be:

Sessions
People would attend:

Age
- ● 0-5 years
- ● 6-12 years
- ● 13-18 years

Due to rounding, percentages may not add up to 100.

Gender
- ● Female
- ● Male

Ethnicity
- ● Other Ethnic Group
- ● White British

Sessions
- ● 1 to 5
- ● 6 to 10
- ● 11 to 20
- ● More than 21
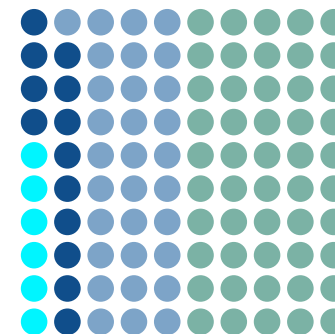
67%

Current completeness
This shows how many people completed questionnaires at their first and at a return visit.

CORC
Child Outcomes
Research Consortium

The Child Outcomes Research Consortium (CORC) is the UK's leading membership organisation that collects and uses evidence to improve children and young people's mental health and wellbeing.