

# Place2Be CORC report

Produced by CORC, May 2020

Reporting period: 01 May 2008 to 30 November 2019



# Contents

<b>Key Findings</b> .....	<b>5</b>
Summary of the main report	
<b>Implications</b> .....	<b>10</b>
Discussion about what the report may mean for your service	
<b>Your Report</b> .....	<b>12</b>
Information about the service and data completeness	
<b>Are children, young people and their families improving?</b> .....	<b>25</b>
<b>Are children, young people and their families improving compared to rest of CORC?</b> .....	<b>37</b>
<b>Appendices</b> .....	<b>50</b>
A: Technical Glossary	
B: Reference Tables	
C: Demographics	
D: Reliable Change for Individual SDQ Measures	
E: “Recovery” and Reliable Change	
F: Outcome Measures	
G: Our Data Slide	

# 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](http://www.corc.uk.net)

The Kantor Centre of Excellence, 4-8 Rodney Street, London, N1 9JH  
020 7443 2225

[CORC@annafreud.org](mailto:CORC@annafreud.org)

# Key Contact

The key contacts within your service are:

Nikhil Naag

Sarah Golden

# Key Findings

Are children, young people and their families improving?

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

# Key Findings

## Comparison data:

The Place2Be sample was compared to the Rest of CORC sample (124,370 cases), obtained from a majority of statutory services and some non-statutory services. Also included are members working primarily in Targeted and Specialist CAMHS (Tier 2&3) and community CAMHS. The RoC sample was filtered to only include cases where children were aged between 6 and 14 years.

## 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 69% (17362/24997). In the rest of CORC sample, this figure was 26% (32903/124370).

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 45% (164/362) for the YPCORE to 77% (7773/10049) for the Child SDQ Total Difficulties. In the rest of CORC sample, comparative follow up rates were 14% (5025/35401) for the Child SDQ Total Difficulties. The Rest of CORC sample did not contain any YPCORE data.

*\* for measures with more than 10 paired cases and a follow up rate of at least 25%*

# Key Findings

## Change in raw scores:

For this analysis, we considered that a child or young person 'improved' or 'deteriorated' on a measure if their scores changed by at least one point between time 1 and time 2.

Looking at the change of child-reported SDQ Total Difficulties individual scores, 70% (5476/7773) of cases improved (compared to 61% in the RoC sample), 6% (461/7773) stayed the same, and 24% (1836/7773) deteriorated. As for parent-reported SDQ Total Difficulties individual scores, 72% (10743/15012) of cases improved (compared to 61% in the RoC sample), 7% (979/15012) stayed the same, and 22% (3290/15012) deteriorated.

Looking at the change of individual scores on child-reported SDQ Impact, 55% (3980/7242) of cases improved (compared to 56% in the RoC sample), 31% (2254/7242) stayed the same, and 14% (1008/7242) deteriorated. For parent-reported scores, 63% (8595/13648) of cases improved (compared to 64% in the RoC sample), 25% (3446/13648) stayed the same, and 12% (1607/13648) deteriorated.

Looking at the change of individual scores from the YPCORE, 68% (111/164) of scores improved, 6% (10/164) stayed the same, and 26% (43/164) deteriorated. The Rest of CORC sample did not contain YPCORE data.

# 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 rest of CORC sample moved from above cut-off at time 1 to below cut-off at time 2.

## **Reliable Change on individual measures:**

On request by Place2Be, we calculated reliable change for individual SDQ 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.

With this approach to analysing the data, the majority of children and young people in your service did not show reliable improvement on SDQ scales. Between 20% and 23% of children reliably improved, while a small number (up to 3%) falling in the range of reliable deterioration. The proportion of children who showed reliable improvement at Place2Be was higher compared to the rest of CORC sample for the SDQ Total Difficulties scales (Child-rated: 22% at Place2Be vs. 15% for Rest of CORC; Parent-rated: 23% vs. 16%). The proportion of children showing reliable improvement on the Child-rated SDQ Impact scale was 20% for both samples, but the proportion of children showing reliable improvement on the Parent-rated SDQ Impact scale was lower for Place2Be (22%) compared to the rest of CORC (27%).



# Key Findings

## **'Recovery' and Reliable Change in relation to multiple scales:**

This approach combines information from multiple scales (e.g. SDQ sub-scales, YP-CORE) to provide an overall measure of recovery or reliable change for each case (Wolpert et al. 2016).

5005 of the 24997 cases (children and young people's periods of contact) seen by your service met criteria for analysis of child-reported reliable change and 'recovery' (including 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).

Taking into account uncertainty in the data due to random variation, we estimate that between 43% and 46% reliably improved, between 45% and 47% '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 a national study from Children and Young People's Mental Health services (Wolpert et al., 2016; see Appendix D).

11112 of the 24997 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 43% 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 from the Wolpert et al. (2016) study (see Appendix D).

Care should be taken when comparing results between Place2Be and the Wolpert et al. (2016) study. Firstly, Wolpert et al. (2016) describe routinely collected outcome data in child mental health as flawed, uncertain, proximate and sparse (FUPS), and provide recommendations for considering and discussing results generated from FUPS data. Secondly, differences are likely to exist between the Place2Be and Wolpert et al. (2016) samples with regard to the demographics of the children and young people included, the range and severity of mental health difficulties experienced and the types and duration of interventions accessed. Finally, a wider range of outcome measures were used to calculate reliable change in the Wolpert et al. (2016) study compared to Place2Be. Overall these differences across samples should be kept in mind when comparing and interpreting findings.

# Implications

What are the main implications?

# Implications

## Use of measures

Your service is collecting information from different perspectives, i.e. children and parents and carers, and we would encourage you to continue to do so.

## Data collection

The paired outcome rate in your service is higher than in the Rest of CORC sample and we would encourage you to maintain this rate.

## Outcomes

Improvement in scores was generally higher than the rest of CORC services. It may be interesting to explore any patterns/trends for particular sub-groups, for example by age. You may also wish to consider other factors that could be influencing outcomes, such as complexity of needs and length and appropriateness of interventions.

# 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) (<https://digital.nhs.uk/home>).

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

# Data Completion: Outcome Measures

What is the sample size?

Who is in your sample?

Paired outcome rates

Follow up rates

# What is the sample size?

Date from: 01  
May 2008

Date to: 30  
November 2019

24,997 cases - service

---

124,370 cases – from the rest of CORC (RoC) sample,  
made up of 62 services

'Rest of CORC (RoC)' is made up of a majority of statutory services and some non-statutory services. Also included are members working primarily in Targeted and Specialist CAMHS (Tier 2&3) and community CAMHS. The RoC sample was filtered to only include cases where children were aged between 6 and 14 years.

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.

What is the sample size?

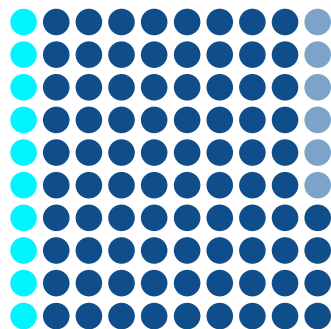
# Who is in your sample?

## 24997 people were seen by the service

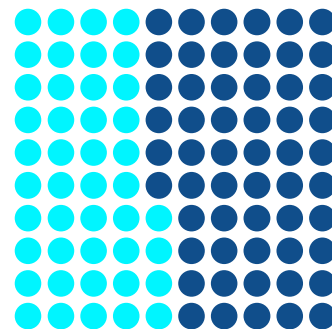
This slide displays a summary of age, gender, ethnicity and sessions people would attend for cases shown in further detail in the report. This sample may not be representative of every child seen by your service as data may not have been submitted for all children, and completeness may vary across variables.

### If the service saw 100 people:

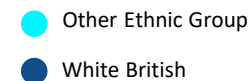
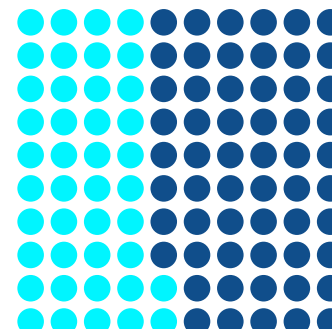
Age  
People would be:



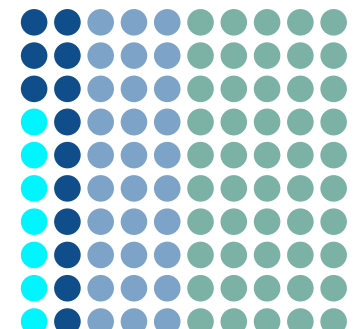
Gender  
People would be:



Ethnicity  
People would be:



Sessions  
People would attend:





# 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.

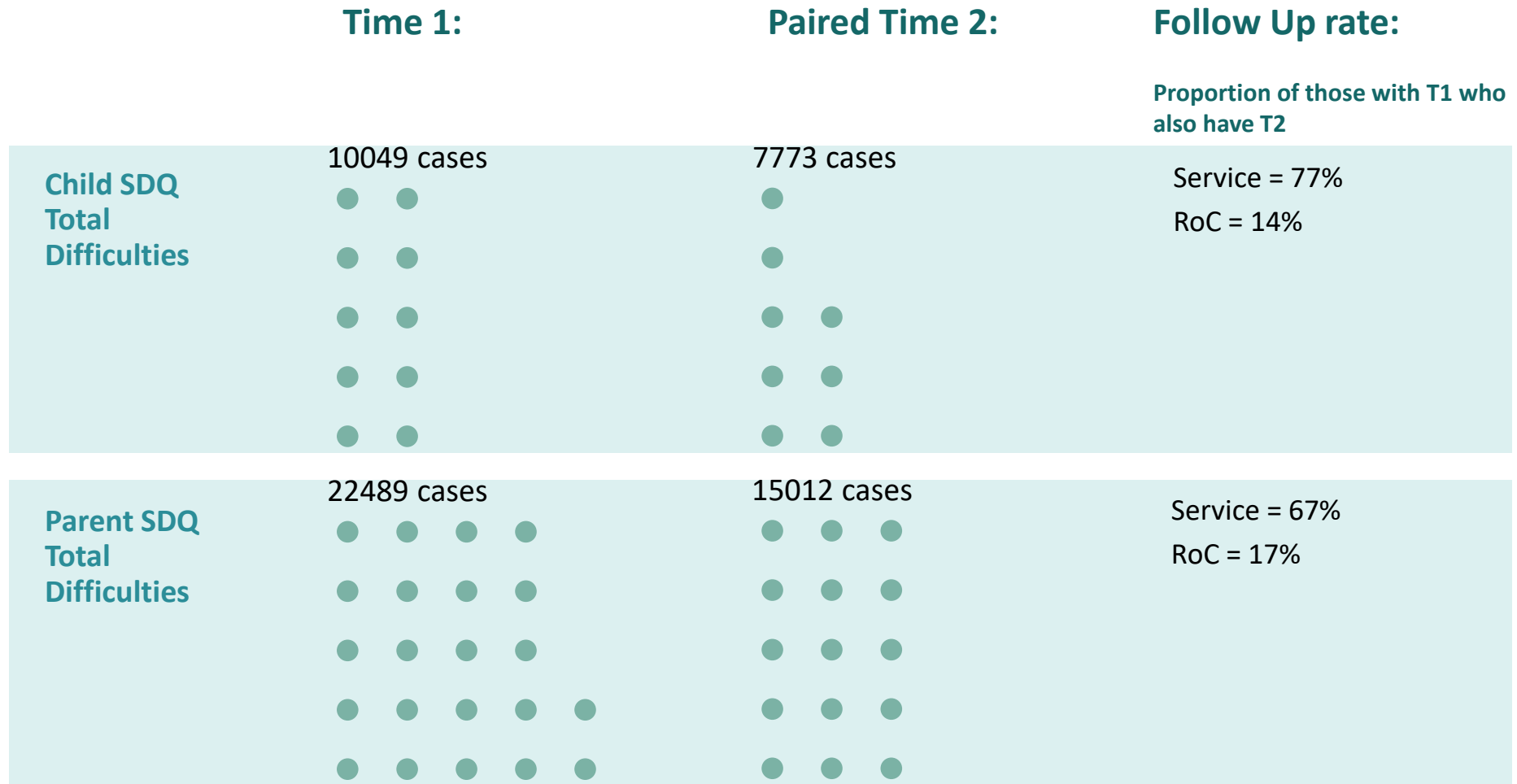
Service = 69%, 17362 cases

Rest of CORC = 26%, 32903 cases



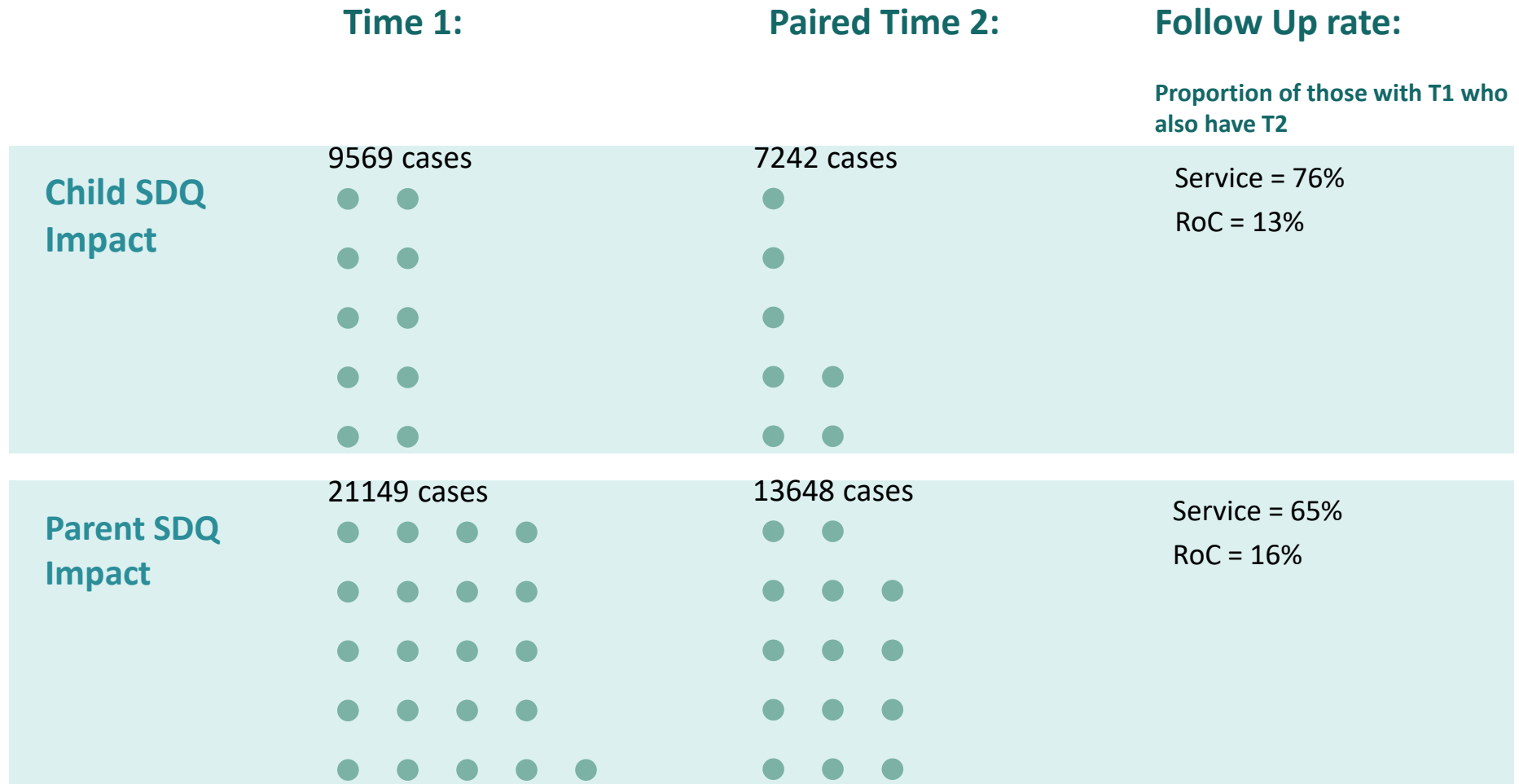
How many children and young people have a paired outcome measure?

# What are the Follow Up rates?



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

# What are the Follow Up rates?



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

# What are the Follow Up rates?

	Time 1:	Paired Time 2:	Follow Up rate:
YPCORE	362 cases	164 cases	Proportion of those with T1 who also have T2 Service = 45% RoC = No data

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

# About the Plots

Information about how to read and interpret the plots.

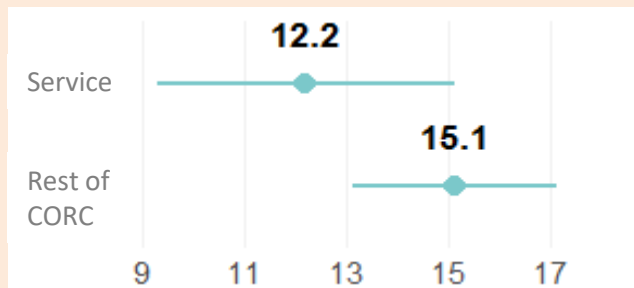
# What are Margins of Error?

**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.

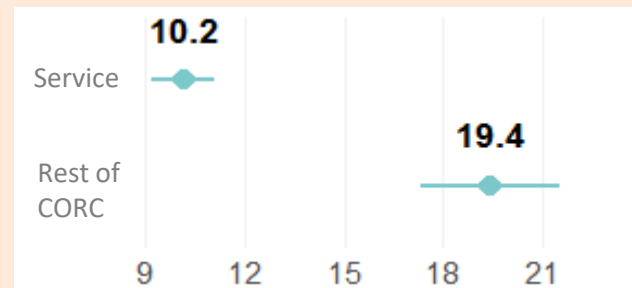
**How can margins of error be used to evaluate change scores?** When the margin of error of an average change score doesn't cross 0, it suggests that there may be a difference between two scores. When it crosses 0, there is no evidence to suggest that the two scores are different.

**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 to be a difference between the two averages.

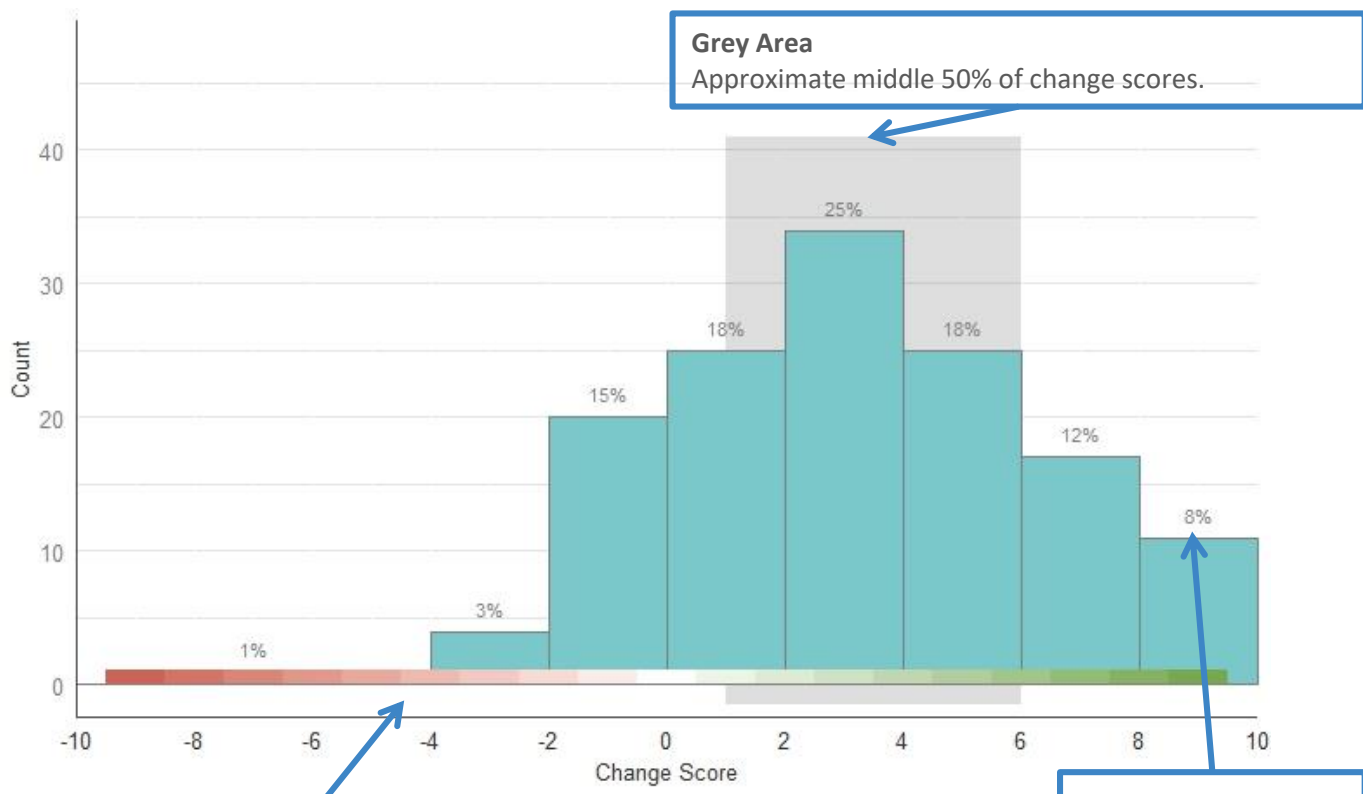


# What do the plots mean?

**Icon**  
Indicates who completed the outcome measure.



Child



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

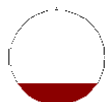
**Labels**  
Give the percentage of individuals with a score between the two values in the horizontal axis.

**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

What do the plots mean?

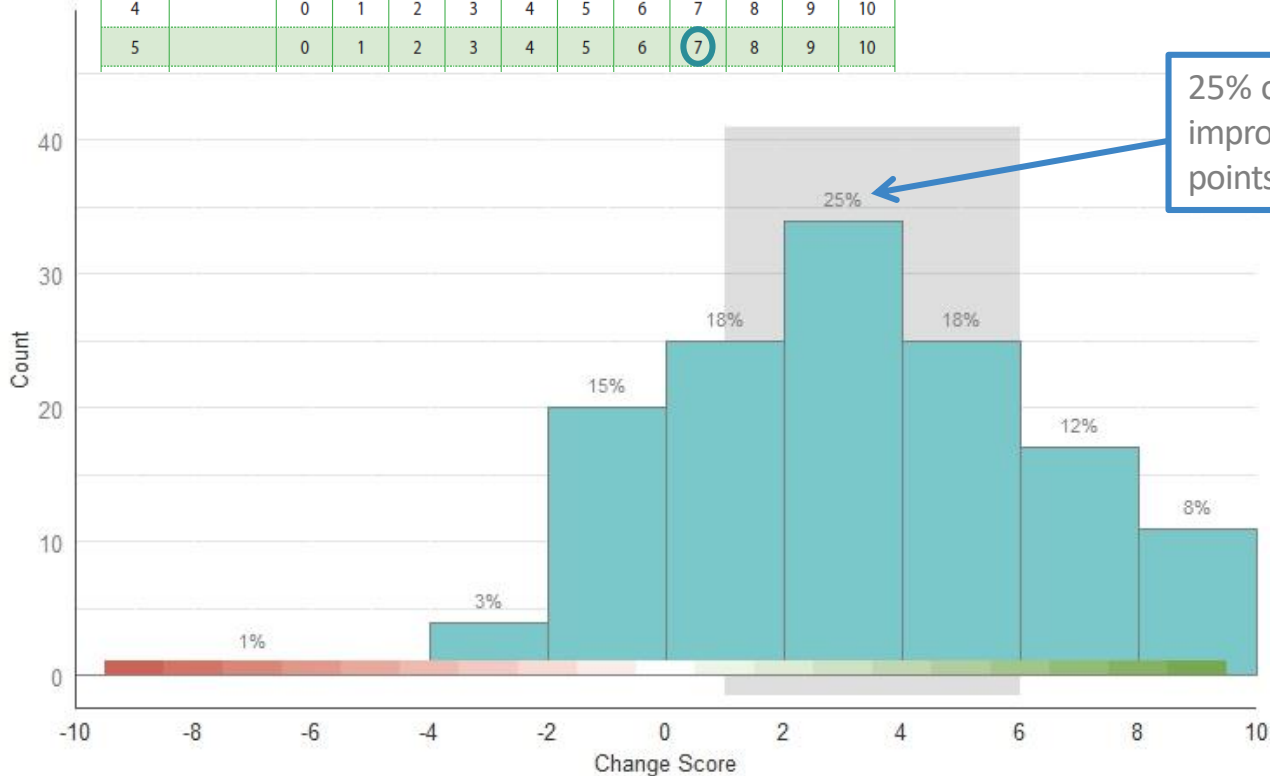
# How are 'change' scores calculated?

Session	Date	Today I would rate my progress to this goal? (please circle the appropriate number below)										
Remember a score of zero means no progress has been made towards a goal, a score of ten means a goal has been reached fully, and a score of five is exactly half way between the two												
1		0	1	2	3	4	5	6	7	8	9	10
2		0	1	2	3	4	5	6	7	8	9	10
3		0	1	2	3	4	5	6	7	8	9	10
4		0	1	2	3	4	5	6	7	8	9	10
5		0	1	2	3	4	5	6	7	8	9	10

Change = Goal progress (last session) – Goal progress (first session)

Change = 7 – 2

Change = 5



25% of children's scores improved by between 2 and 4 points from time 1 to time 2.

We plot these change scores to get an impression of how much all individuals have changed between time 1 and time 2.

How are 'change' scores calculated?

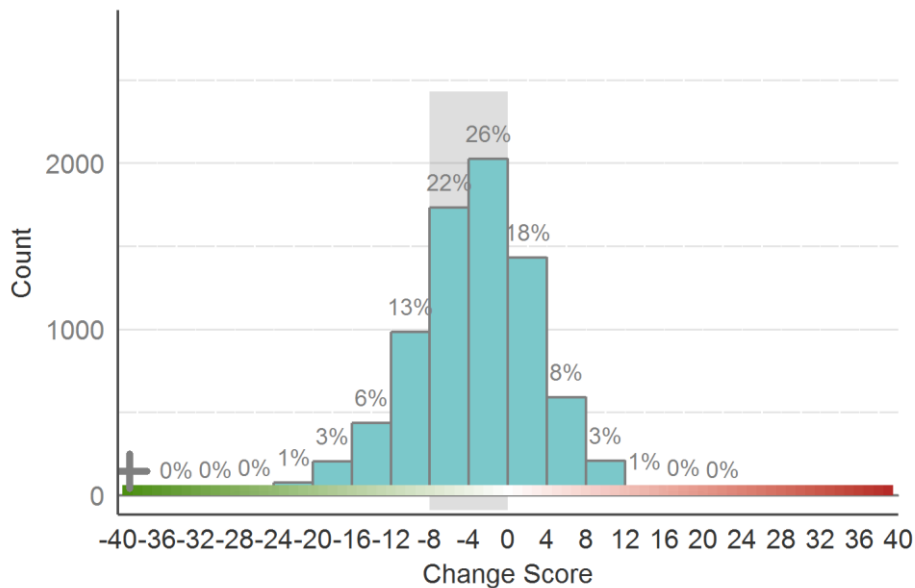


# Are children, young people and their families improving?

SDQ  
YP CORE

Are children, young people and their families improving?

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



## Results included

Any paired Child SDQ Total Difficulties scores, n = 7773

## How representative is this sample?

77% (7773/10049) of those with a time 1 score, had a corresponding time 2 score at your service.

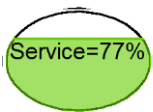
## What do the plots show?

On average, scores changed by -3.9 points (paired sample t-test,  $t(7772) = -53$ ,  $p < .001$ ; moderate effect size, Cohen's  $d = .60$ ), with a margin of error between -4.1 and -3.8.

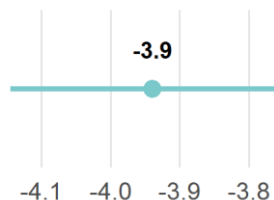
Looking at the change of individual scores, 70% (5476/7773) of scores improved, 6% (461/7773) stayed the same, and 24% (1836/7773) deteriorated.

The middle 50% of children and young people with paired scores changed by between -8 and 0 points.

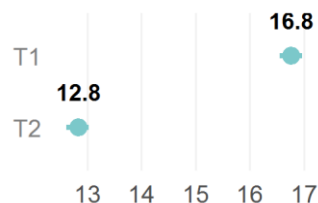
### Follow Up Rate



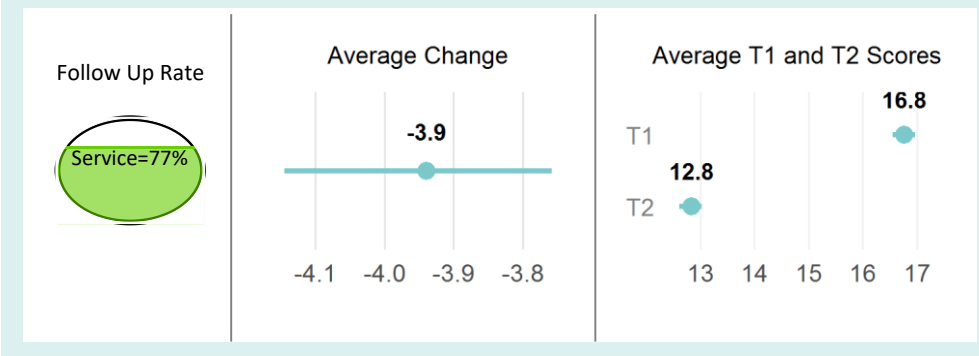
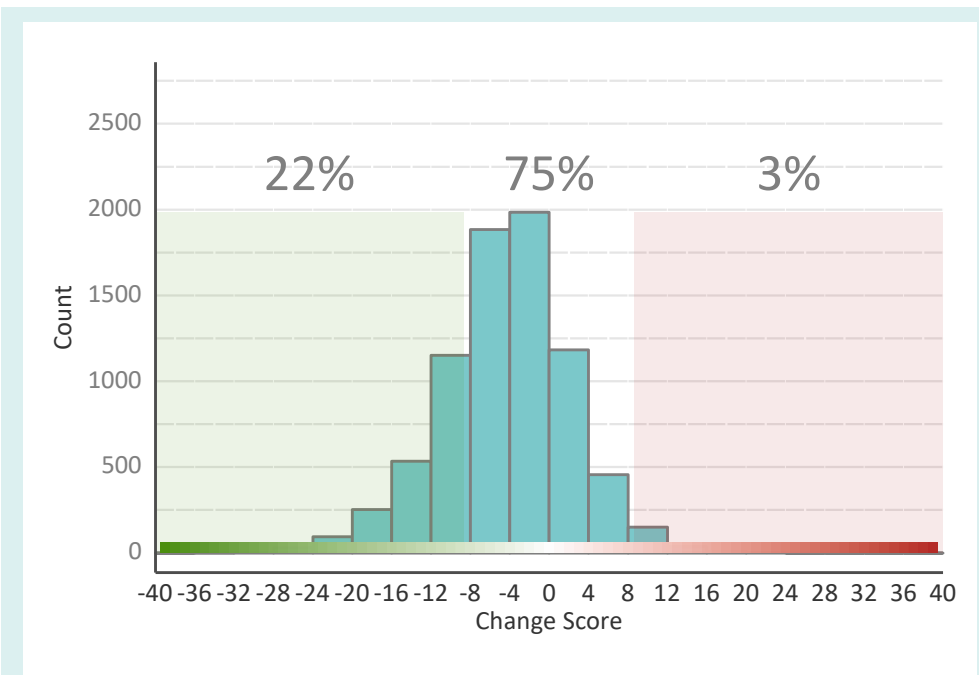
### Average Change



### Average T1 and T2 Scores



# Have Child SDQ Total Difficulties Scores Reliably Changed?



## Results included

Any paired score for Child SDQ Total Difficulties (n = 7773), categorised by whether they fall above or below the threshold for reliable change. Change scores higher/lower than +/- 8.70 were classified as reliable change.

## How representative is this sample?

77% (7773/10049) of those with a time 1 score, had a corresponding time 2 score at your service.

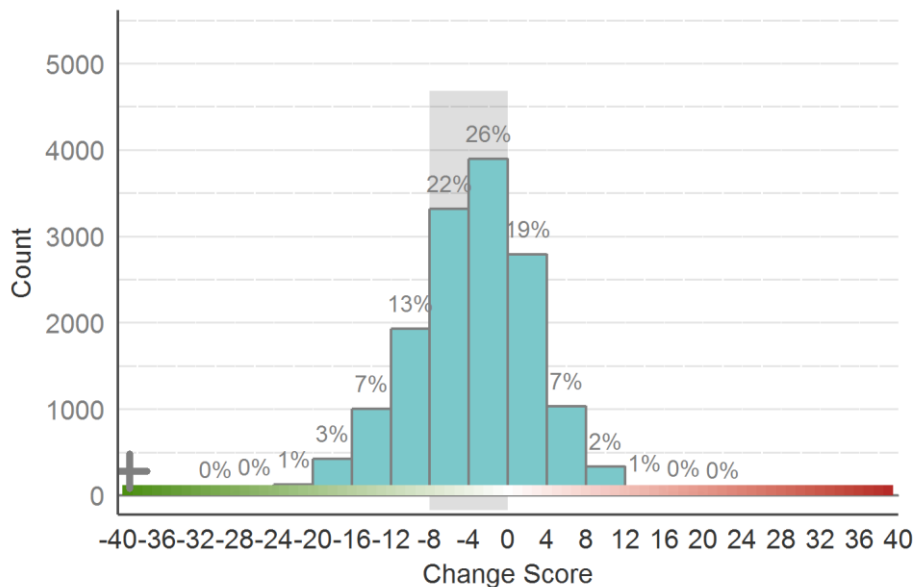
## What do the plots show?

22% of children and young people showed reliable improvement between time 1 and time 2. 75% of children did not show reliable change, and a small proportion (3%) had scores that reliably deteriorated.

## Conclusion

22% 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.

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



## Results included

Any paired Parent SDQ Total Difficulties scores, n = 15012

## How representative is this sample?

67% (15012/22489) of those with a time 1 score, had a corresponding time 2 score at your service.

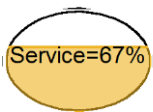
## What do the plots show?

On average, scores changed by -4.2 points (paired sample t-test,  $t(15011) = -79, p < .001$ ; moderate effect size, Cohen's  $d = .64$ ), with a margin of error between -4.4 and -4.1.

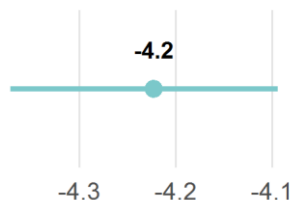
Looking at the change of individual scores, 72% (10743/15012) of scores improved, 7% (979/15012) stayed the same, and 22% (3290/15012) deteriorated.

The middle 50% of children and young people with paired scores changed by between -8 and 0 points.

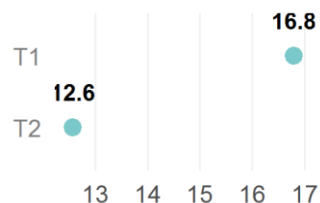
### Follow Up Rate



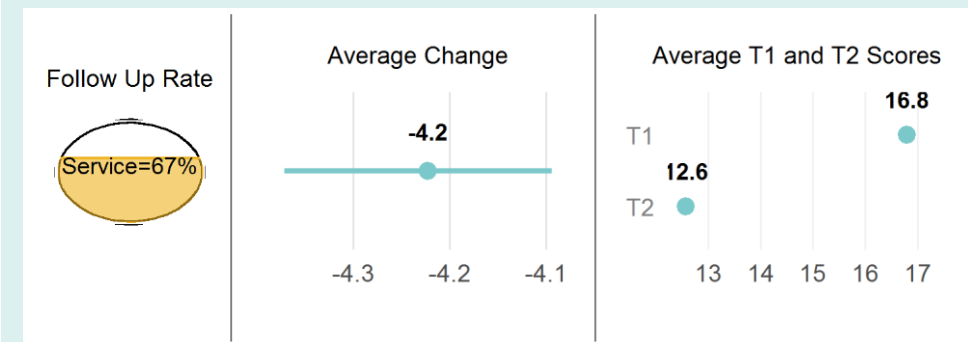
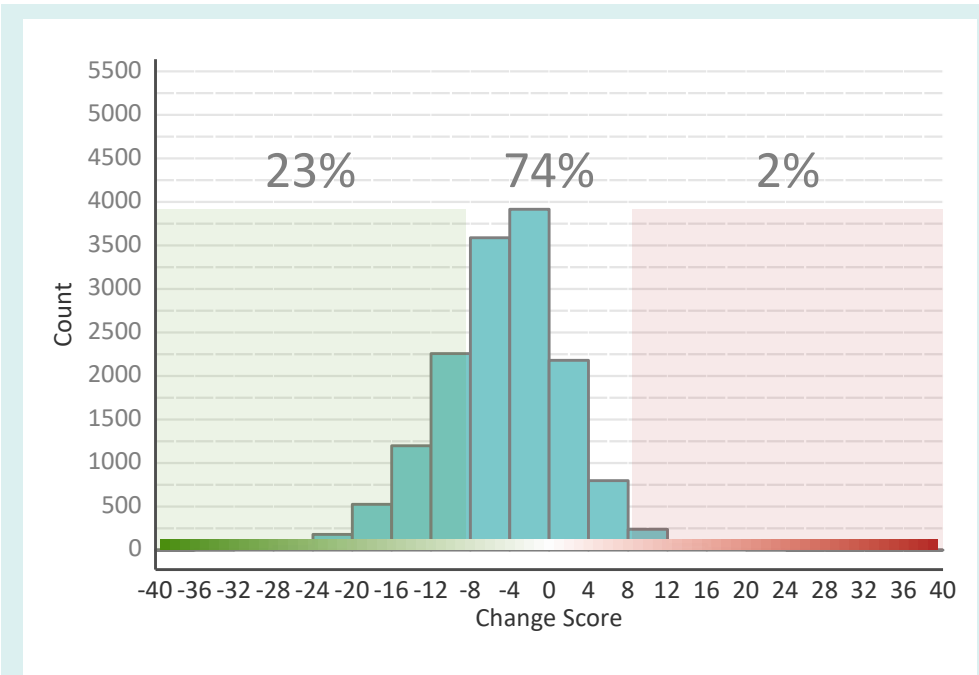
### Average Change



### Average T1 and T2 Scores



# Have Parent SDQ Total Difficulties Reliably Changed?



## Results included

Any paired Parent SDQ Total Difficulties scores (n = 15012), categorised by whether they fall above or below the threshold for reliable change. Change scores higher/lower than +/- 8.47 were classified as reliable change.

## How representative is this sample?

67% (15012/22489) of those with a time 1 score, had a corresponding time 2 score at your service.

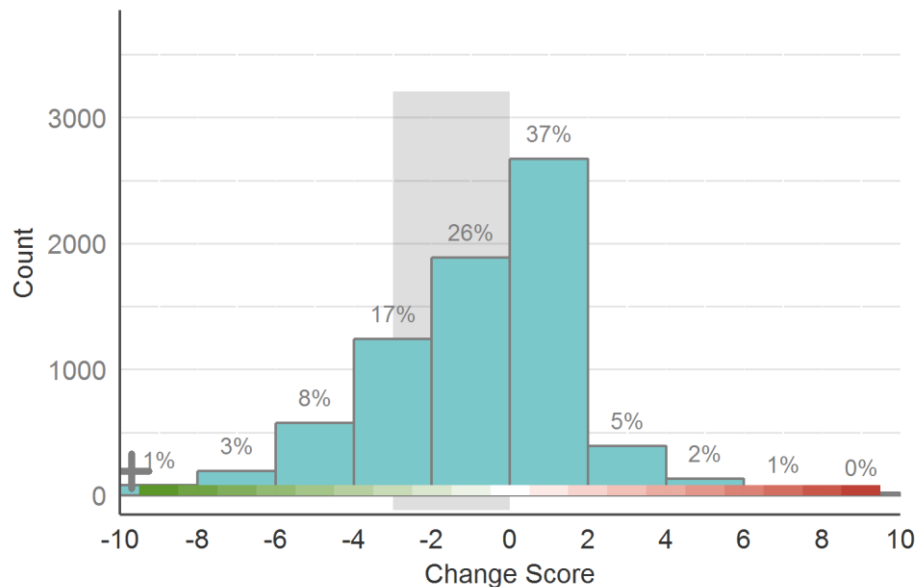
## What do the plots show?

23% of children and young people in your sample reliably improved. 74% of children and young people did not show reliable change, and a small proportion (2%) had scores that reliably deteriorated.

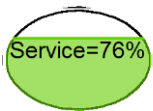
## Conclusion

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.

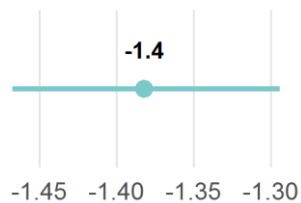
# How have Child SDQ Impact scores changed between T1 and T2?



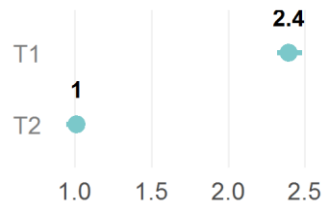
Follow Up Rate



Average Change



Average T1 and T2 Scores



## Results included

Any paired Child SDQ Impact scores, n = 7242

## How representative is this sample?

76% (7242/9569) of those with a time 1 score, had a corresponding time 2 score at your service.

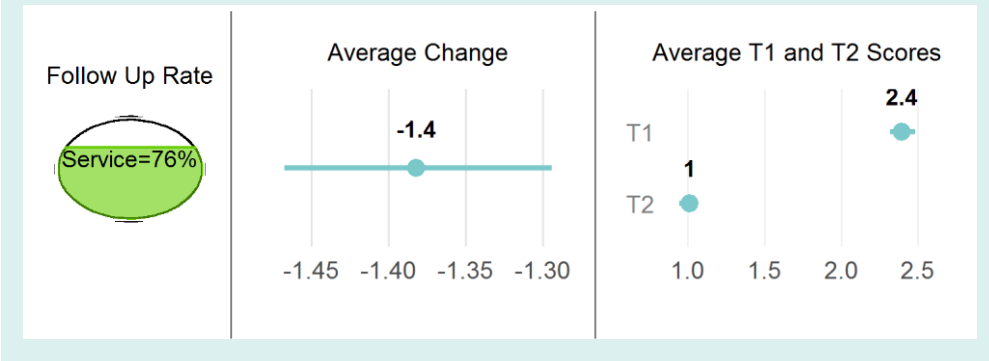
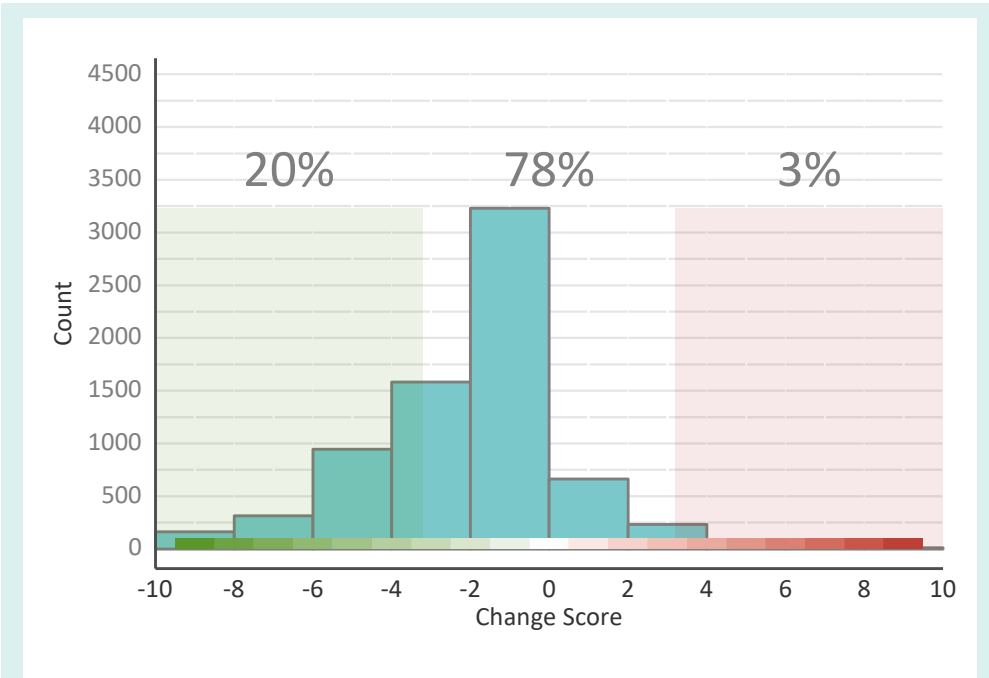
## What do the plots show?

On average, scores changed by -1.4 points (paired sample t-test,  $t(7241) = -45, p < .001$ ; moderate effect size, Cohen's  $d = .53$ ), with a margin of error between -1.5 and -1.3.

Looking at the change of individual scores, 55% (3980/7242) of scores improved, 31% (2254/7242) stayed the same, and 14% (1008/7242) deteriorated.

The middle 50% of children and young people with paired scores changed by between -3 and 0 points.

# Have Child SDQ Impact scores Reliably Changed?



## Results included

Any paired Child SDQ Impact scores (n = 7242), categorised by whether they fall above or below the threshold for reliable change. Change scores higher/lower than +/- 3.21 were classified as reliable change.

## How representative is this sample?

76% (7242/9569) of those with a time 1 score, had a corresponding time 2 score at your service.

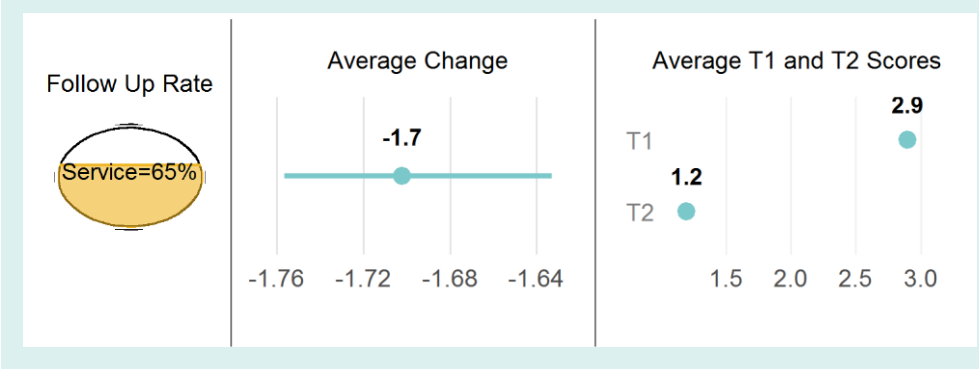
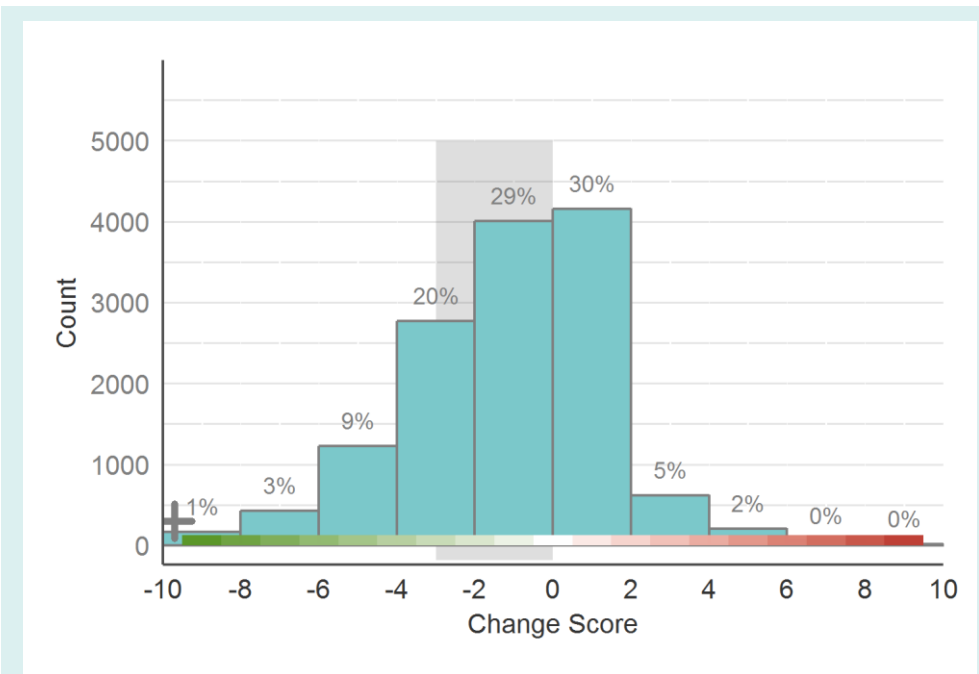
## What do the plots show?

20% of children and young people reliably improved. 78% of children and young people did not show reliable change, and 3% of children and young people reliably deteriorated.

## Conclusion

20% of children reliably improved between time 1 and time 2 on the Child SDQ 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.

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



## Results included

Any paired Parent SDQ Impact scores, n = 13648

## How representative is this sample?

65% (13648/21149) of those with a time 1 score, had a corresponding time 2 score at your service.

## What do the plots show?

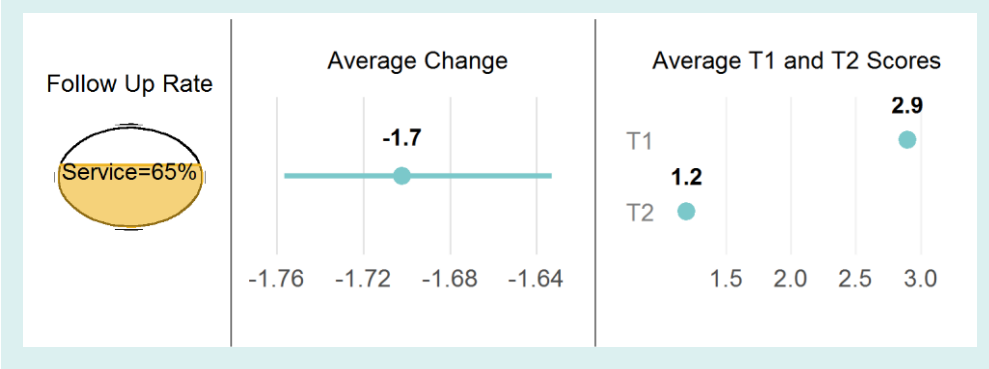
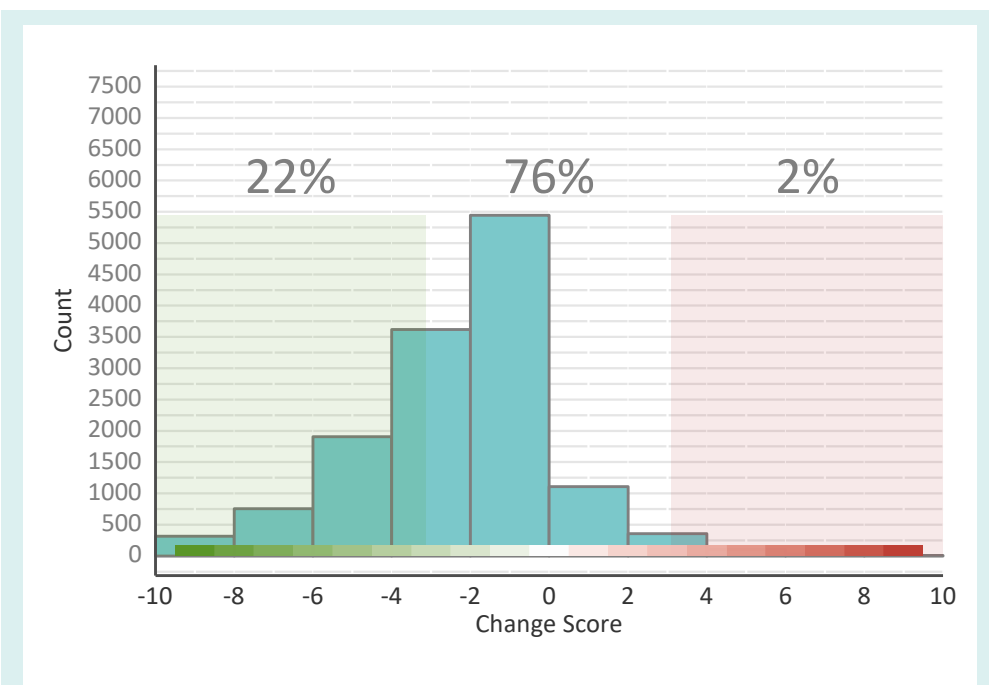
On average, scores changed by -1.7 points (paired sample t-test,  $t(13647) = -78, p < .001$ ; moderate effect size, Cohen's  $d = .67$ ), with a margin of error between -1.8 and -1.6.

Looking at the change of individual scores, 63% (8595/13648) of scores improved, 25% (3446/13648) stayed the same, and 12% (1607/13648) deteriorated.

The middle 50% of children and young people with paired scores changed by between -3 and 0 points.



# Have Parent SDQ Impact scores Reliably Changed?



## Results included

Any paired Parent SDQ Impact scores (n = 13648), categorised by whether they fall above or below the threshold for reliable change. Change scores higher/lower than +/- 3.12 were classified as reliable change.

## How representative is this sample?

65% (13648/21149) of those with a time 1 score, had a corresponding time 2 score at your service.

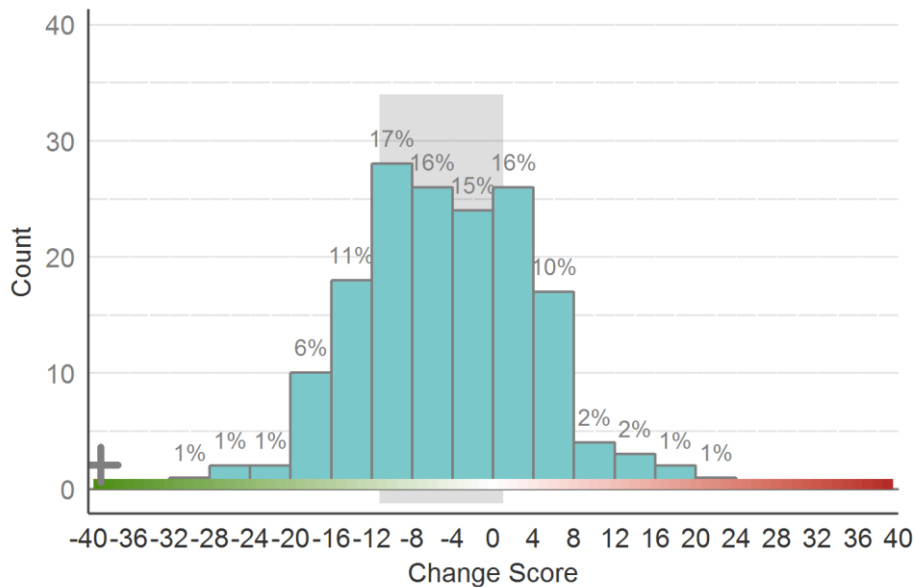
## What do the plots show?

22% of children and young people reliably improved. 76% of children and young people did not show reliable change, and 2% of children and young people reliably deteriorated.

## Conclusion

22% of children reliably improved between time 1 and time 2 on the Parent SDQ 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.

# How have YPCORE scores changed between T1 and T2?



## Results included

Any paired YPCORE scores, n = 164

## How representative is this sample?

45% (164/362) 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 -5.1 points (paired sample t-test,  $t(163) = -7.43$ ,  $p < .001$ ; moderate effect size, Cohen's  $d = .58$ ), with a margin of error between -7.1 and -3.4.

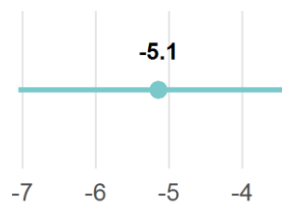
Looking at the change of individual scores, 68% (111/164) of scores improved, 6% (10/164) stayed the same, and 26% (43/164) deteriorated.

The middle 50% of children and young people with paired scores changed by between -11.25 and 1 points.

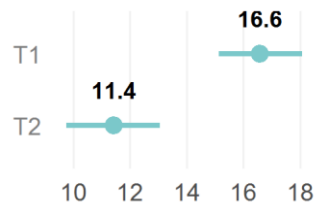
### Follow Up Rate



### Average Change



### Average T1 and T2 Scores



## 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	2304	46% [45%-47%]
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	2221	44% [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	2438	49% [47%-50%]
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	346	7% [5%-8%]
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	1382	28% [26%-29%]

*The following measures are included: Child SDQ Emotional, Child SDQ Conduct, Child SDQ Hyperactivity, Child SDQ Impact, YP-CORE*

### 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= 5005

### How representative is this sample?

78% (5005/6455) of those in the clinical sample fell in the 'paired clinical sample'.

### What does the table show?

Scores for 46% (margin of error between 45% and 47%) of children and young people showed 'recovery', 44% (margin of error between 43% and 46%) showed reliable improvement, and 28% (margin of error between 26% and 29%) showed 'reliable recovery'.

## 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	4670	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	4845	44% [43%-45%]
No reliable change	Change from a first to a last time point was less than what would be expected due to measurement error	5659	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	608	5% [4%-6%]
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	2690	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= 11112

### How representative is this sample?

69% (11112/16216) of those in the clinical sample fell in the 'paired clinical sample'.

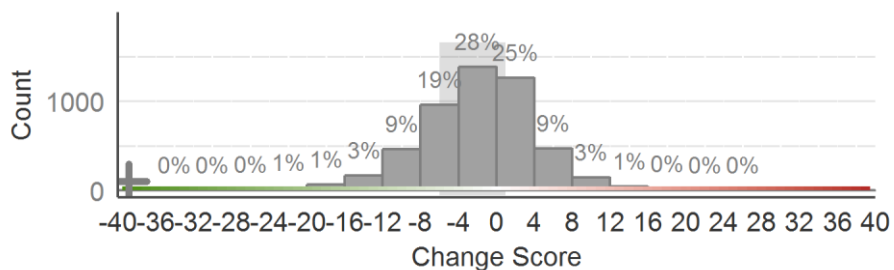
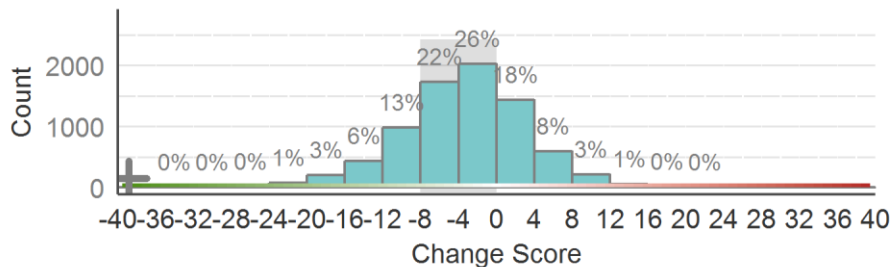
### What does the table show?

Scores for 42% (margin of error between 41% and 43%) of children and young people showed 'recovery', 44% (margin of error between 43% and 45%) showed reliable improvement, and 24% (margin of error between 23% and 25%) showed 'reliable recovery'.

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

SDQ

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



Follow Up Rate

Service=77%

Follow Up Rate

RoC=14%

Average Change

-3.9

-4.0

-3.5

-3.0

-2.5

Average T1 Scores

16.8

17.0

17.5

18.0

## Results included

Any paired Child SDQ Total Difficulties scores: Service n = 7773, Rest of CORC n = 5025

## How representative is this sample?

77% (7773/10049) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 14% (5025/35401).

## What do the plots show?

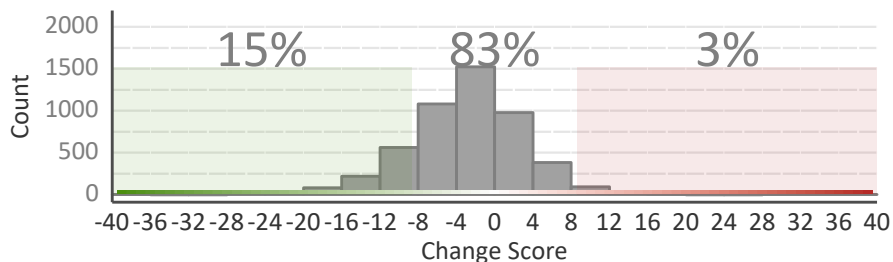
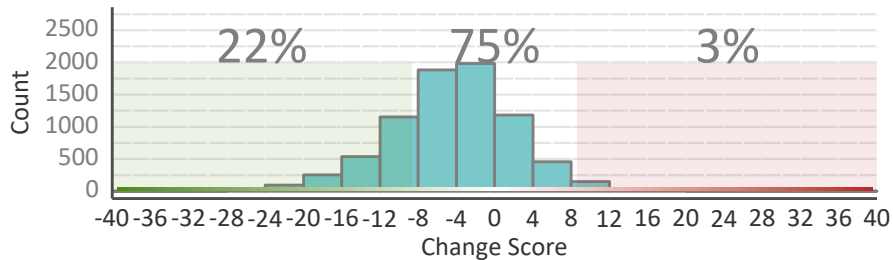
On average, scores changed by -3.9 points, with a margin of error between -4.1 and -3.8 in your service.

In RoC services, the average change score was -2.5, between -2.7 and -2.3.

Looking at the change of individual scores, 70% (5476/7773) of scores improved in your service, and 61% (3083/5025) 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 -6 and 1 for the RoC sample.

# Have Child SDQ Total Difficulties scores Reliably Changed compared to the rest of CORC (RoC) sample?



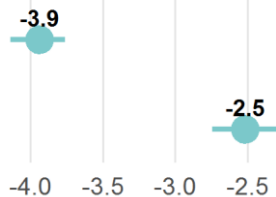
Follow Up Rate

Service=77%

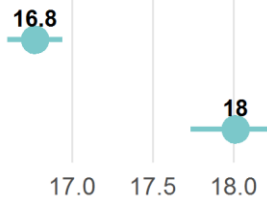
Follow Up Rate

RoC=14%

Average Change



Average T1 Scores



## Results included

Any paired Child SDQ Total Difficulties scores, with scores categorised by whether they fall above or below the threshold for reliable change. Change scores higher/lower than +/- 8.70 were classified as reliable change. Service n = 7773, Rest of CORC n = 5025

## How representative is this sample?

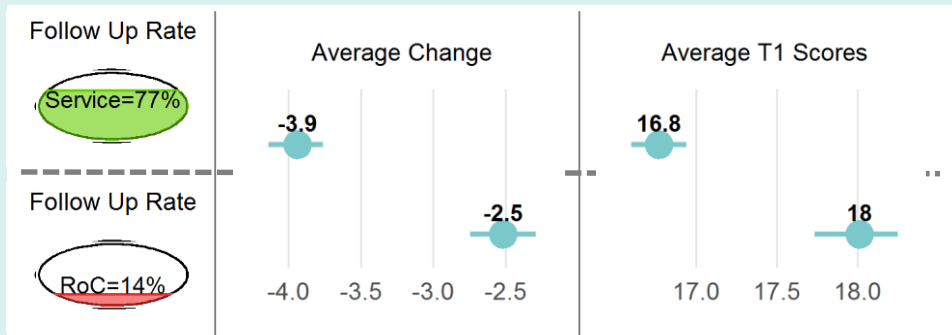
77% (7773/10049) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 14% (5025/35401).

## What do the plots show?

Place2Be had a higher proportions of children showing reliable improvement compared to the rest of CORC (22%, compared with 15%), lower proportion of children showing no change (75%, compared with 83%) and similar proportions showing reliable deterioration (3%, compared with 3%).

# How have Child SDQ Total Difficulties scores changed in relation to a cut-off, compared to the Rest of CORC?

	Place2Be	Rest of CORC
Stayed Above	17.2% (1340)	32.6% (1638)
Improved (above to below)	28.6% (2223)	20.8% (1043)
Deteriorated (below to above)	6.1% (477)	7.4% (370)
Stayed Below	48.0% (3733)	39.3% (1974)



## Results included

Any paired Child SDQ Total Difficulties score: Service n = 7773, Rest of CORC n = 5025. 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?

77% (7773/10049) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 14% (5025/35401).

## What does the table show?

48% children and young people seen by your service had scores that were below cut-off at both time points (compared to 39% in the rest of CORC). 29% of children started above cut-off at time 1 and moved to below cut-off at time 2 (compared to 21% in the rest of CORC), while 17% remained above cut-off at both time points (compared to 33% in the rest of CORC). A small proportion (6-7%) of children in both groups declined from time 1 to time 2.

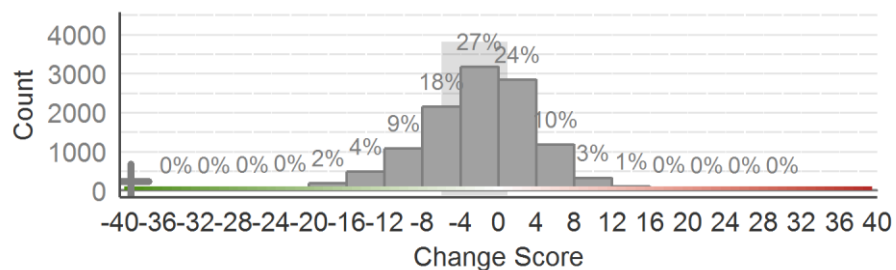
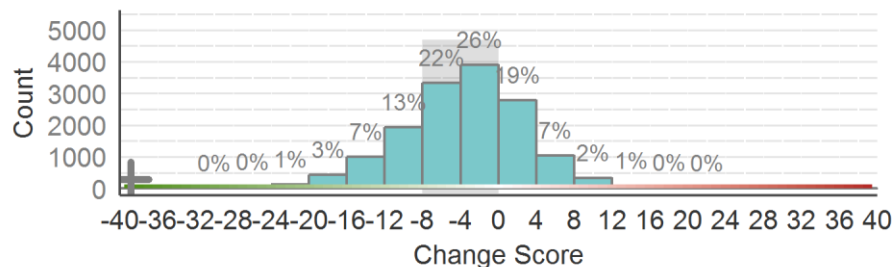
## Conclusion

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.

\* 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: <https://www.sdqinfo.org/py/sdqinfo/c0.py>.



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



Follow Up Rate

Service=67%

Follow Up Rate

RoC=17%

Average Change

-4.2

Average T1 Scores

16.8

-2.6

19.8

## Results included

Any paired Parent SDQ Total Difficulties scores: Service n = 15012, Rest of CORC n = 11612

## How representative is this sample?

67% (15012/22489) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 17% (11612/69269).

## What do the plots show?

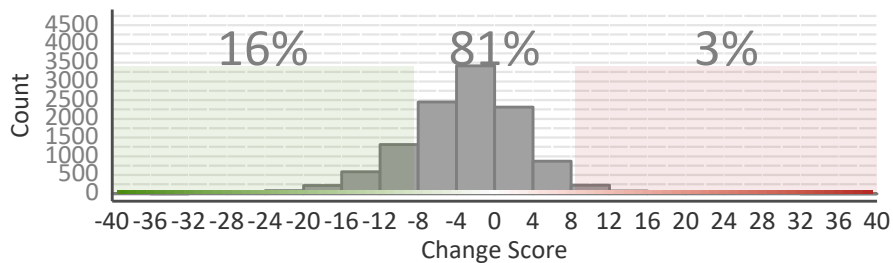
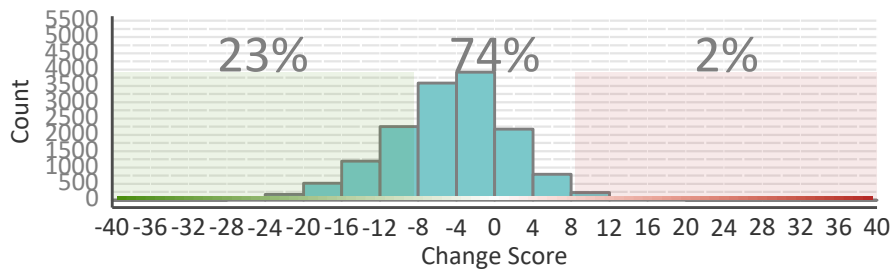
On average, scores changed by -4.2 points, with a margin of error between -4.4 and -4.1 in your service.

In RoC services, the average change score was -2.6, between -2.7 and -2.4.

Looking at the change of individual scores, 72% (10743/15012) of scores improved in your service, and 61% (7138/11612) 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 -6 and 1 for the RoC sample.

# Have Parent SDQ Total Difficulties scores Reliably Changed compared to the rest of CORC (RoC) sample?



Follow Up Rate

Service=67%

Follow Up Rate

RoC=17%

Average Change

-4.2

-4.0 -3.5 -3.0 -2.5

Average T1 Scores

16.8

19.8

17 18 19 20

## Results included

Any paired Parent SDQ Total Difficulties scores, with scores categorised by whether they fall above or below the threshold for reliable change. Change scores higher/lower than +/- 8.47 were classified as reliable change. Service n = 15012, Rest of CORC n = 11612

## How representative is this sample?

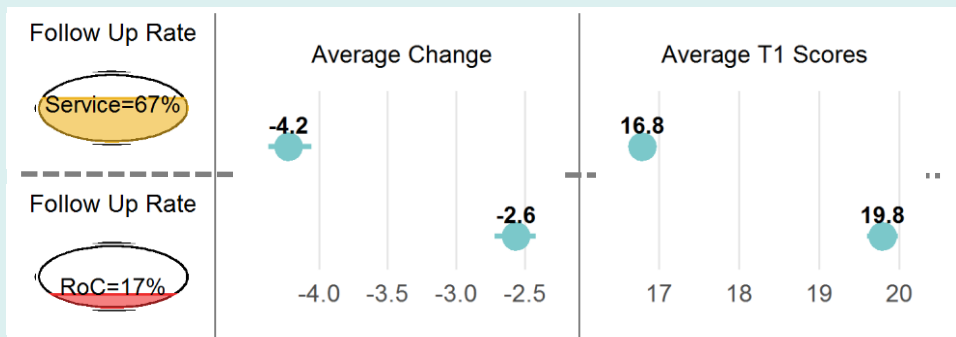
67% (15012/22489) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 17% (11612/69269).

## What do the plots show?

A higher percentage of children in the Place2Be sample reliably improved compared to the rest of CORC sample (23% compared to 16%), a smaller percentage of children showed no reliable change compared to the rest of CORC (74% compared to 81%) and a similar percentage of children reliably deteriorated (2% compared to 3%).

# How have Parent SDQ Total Difficulties scores in relation to a cut-off, compared to the Rest of CORC?

	Place2Be	Rest of CORC
Stayed Above	22.2% (3397)	47.5% (5518)
Improved (above to below)	27.7% (4165)	20.2% (2347)
Deteriorated (below to above)	4.6% (691)	5.7% (663)
Stayed Below	45.0% (6759)	26.6% (3084)



## Results included

Any paired Parent SDQ Total Difficulties score: Service n = 15012, Rest of CORC n = 11612. 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?

67% (15012/22489) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 17% (11612/69269).

## What does the table show?

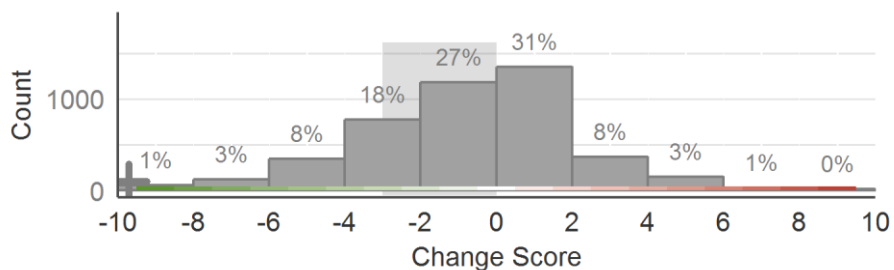
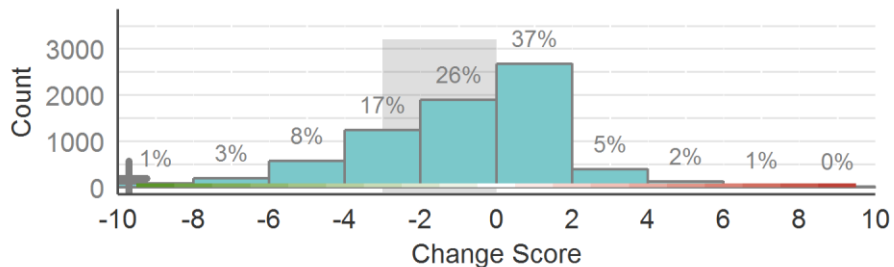
45% children and young people seen by your service had scores that were below cut-off at both time points (compared to 27% in the rest of CORC). 28% of children started above cut-off at time 1 and moved to below cut-off at time 2 (compared to 20% in the rest of CORC), while 22% remained above cut-off at both time points (compared to 48% 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.

\* 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: <https://www.sdqinfo.org/py/sdqinfo/c0.py>.

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



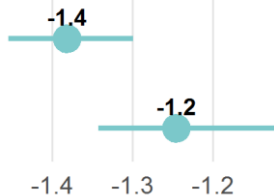
Follow Up Rate

Service=76%

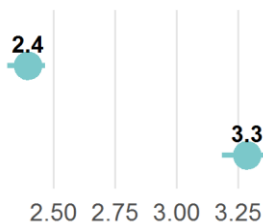
Follow Up Rate

RoC=13%

Average Change



Average T1 Scores



## Results included

Any paired Child SDQ Impact scores: Service n = 7242, Rest of CORC n = 4393

## How representative is this sample?

76% (7242/9569) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 13% (4393/33306).

## What do the plots show?

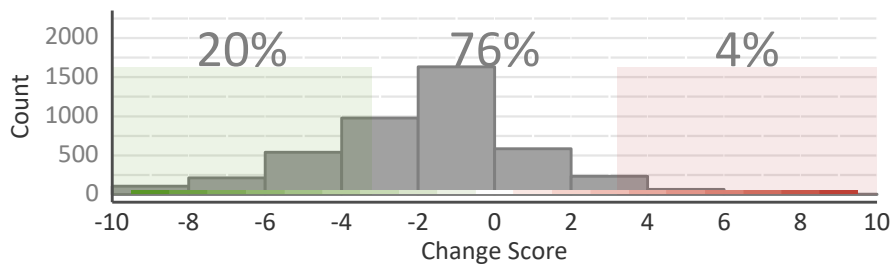
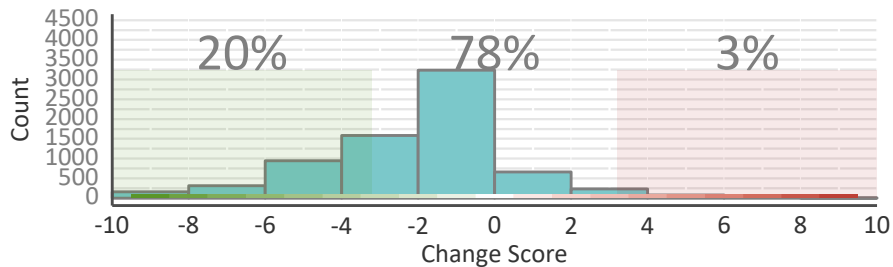
On average, scores changed by -1.4 points, with a margin of error between -1.5 and -1.3 in your service.

In RoC services, the average change score was -1.2, between -1.3 and -1.1.

Looking at the change of individual scores, 55% (3980/7242) of scores improved in your service, and 56% (2475/4393) in the RoC sample.

The middle 50% of children and young people with paired scores changed by between -3 and 0 points for your service, and between -3 and 0 for the RoC sample.

# Have Child SDQ Impact scores Reliably Changed compared to the rest of CORC (RoC) sample?



Follow Up Rate

Service=76%

Follow Up Rate

RoC=13%

Average Change

-1.4

-1.4 -1.3 -1.2

Average T1 Scores

2.4

3.3

2.50 2.75 3.00 3.25

## Results included

Any paired Child SDQ Impact scores, with scores categorised by whether they fall above or below the threshold for reliable change. Change scores higher/lower than +/- 3.21 were classified as reliable change. Service n = 7242, Rest of CORC n = 4393.

## How representative is this sample?

76% (7242/9569) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 13% (4393/33306).

## What do the plots show?

A similar proportion of children showed reliable improvement at Place2Be and in the Rest of CORC (20% compared to 20%), no change (78% compared to 76%) and reliable deterioration (3% vs. 4%).

# How have Child SDQ Impact scores changed in relation to a cut-off, compared to the Rest of CORC?

	Place2Be	Rest of CORC
Stayed Above	22.2% (3397)	47.5% (5518)
Improved (above to below)	27.7% (4165)	20.2% (2347)
Deteriorated (below to above)	4.6% (691)	5.7% (663)
Stayed Below	45.0% (6759)	26.6% (3084)

## Results included

Any paired Child SDQ Impact score: Service n = 15012, Rest of CORC n = 11612. 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?

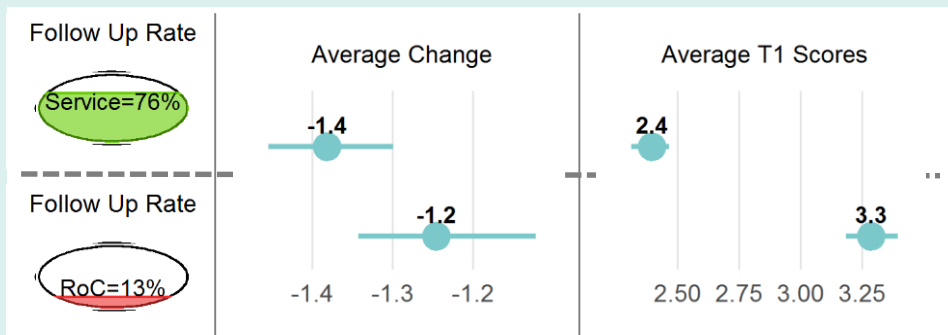
76% (7242/9569) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 13% (4393/33306).

## 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 27% 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 20% in the rest of CORC), while 22% remained above cut-off at both time points (compared to 48% in the rest of CORC). A small proportion (5-6%) of children in both groups declined from time 1 to time 2.

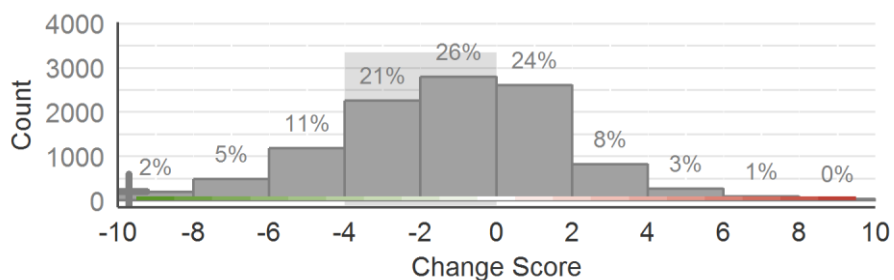
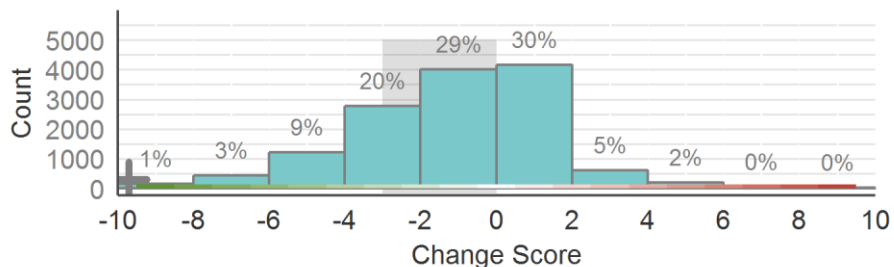
## Conclusion

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 Child SDQ 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: <https://www.sdqinfo.org/py/sdqinfo/c0.py>.

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



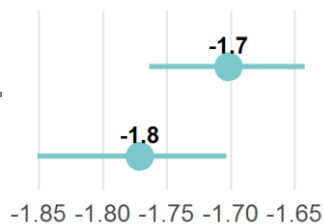
Follow Up Rate

Service=65%

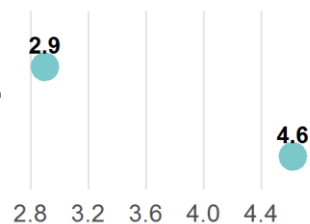
Follow Up Rate

RoC=16%

Average Change



Average T1 Scores



## Results included

Any paired Parent SDQ Impact scores: Service n = 13648, Rest of CORC n = 10713

## How representative is this sample?

65% (13648/21149) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 16% (10713/65792).

## What do the plots show?

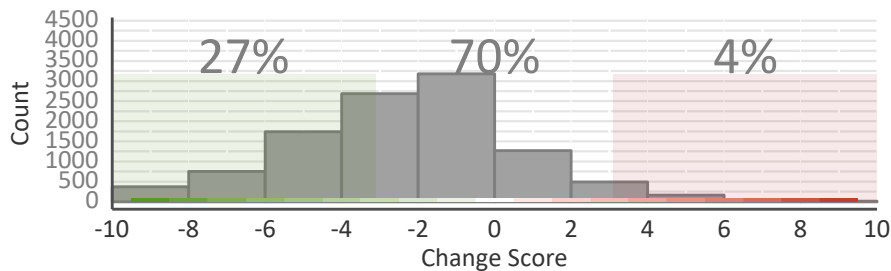
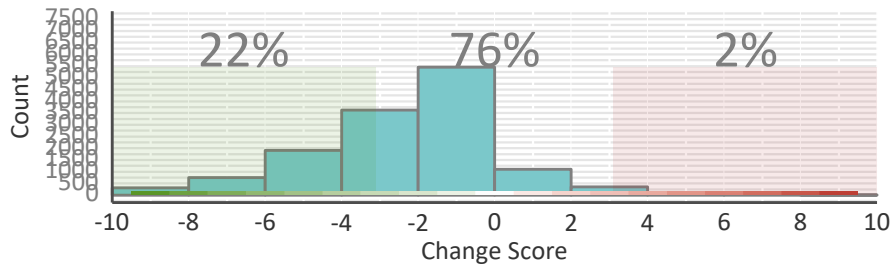
On average, scores changed by -1.7 points, with a margin of error between -1.8 and -1.6 in your service.

In RoC services, the average change score was -1.8, between -1.9 and -1.7.

Looking at the change of individual scores, 63% (8595/13648) of scores improved in your service, and 64% (6904/10713) in the RoC sample.

The middle 50% of children and young people with paired scores changed by between -3 and 0 points for your service, and between -4 and 0 for the RoC sample.

# Have Parent SDQ Impact scores Reliably Changed compared to the rest of CORC (RoC) sample?



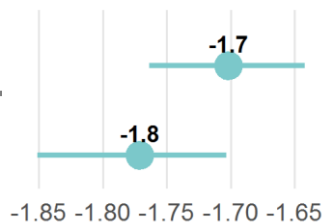
Follow Up Rate

Service=65%

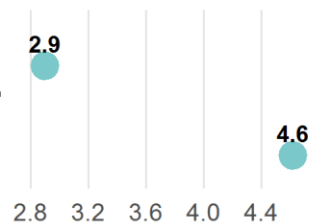
Follow Up Rate

RoC=16%

Average Change



Average T1 Scores



## Results included

Any paired Parent SDQ Impact scores, with scores categorised by whether they fall above or below the threshold for reliable change. Change scores higher/lower than +/- 3.12 were classified as reliable change. Service n = 13648, Rest of CORC n = 10713.

## How representative is this sample?

65% (13648/21149) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 16% (10713/65792).

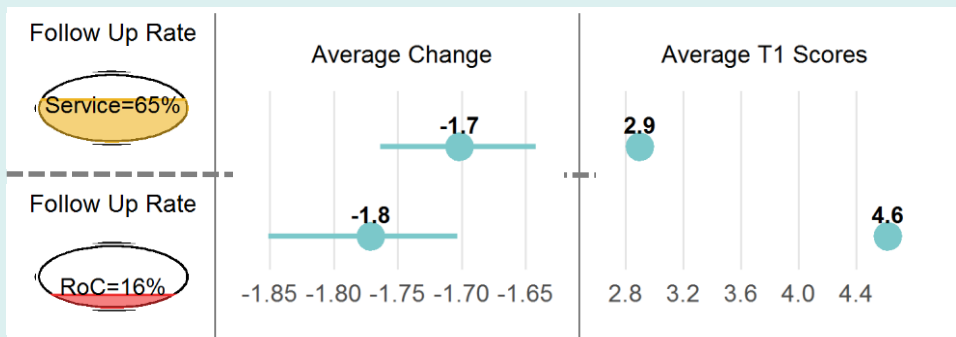
## What do the plots show?

A smaller proportion of children showed reliable improvement on the Parent SDQ Impact scores at Place2Be compared to the Rest of CORC (22% compared to 27%), a higher percentage of children showed no change (76% compared to 70%) and a similar proportion showed reliable deterioration (2% vs. 4%).



# How have Parent SDQ Impact scores changed in relation to a cut-off, compared to the rest of CORC (RoC) sample?

	Place2Be	Rest of CORC
Stayed Above	23.8% (3250)	52.6% (5638)
Improved (above to below)	41.2% (5629)	31.7% (3400)
Deteriorated (below to above)	4.1% (565)	4.0% (432)
Stayed Below	30.8% (4204)	11.6% (1243)



## Results included

Any paired Parent SDQ Impact score: Service n = 13648, Rest of CORC n = 10713. 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?

65% (13648/21149) of those with a time 1 score, had a corresponding time 2 score at your service. For the RoC sample this figure was 16% (10713/65792).

## 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 27% 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 20% in the rest of CORC), while 22% remained above cut-off at both time points (compared to 48% 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 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: <https://www.sdqinfo.org/py/sdqinfo/c0.py>.

# Appendix A

## 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.

Analysis will be conducted on an outcome scale if:

1. there are at least 10 paired cases for the outcome scale, and
2. the follow up rate with respect to the outcome scale is at least 25%.

# 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	YPCORE
Service	1	10049	22489	9569	21149	362
Service	2	7773	15012	7242	13648	164
Rest of CORC	1	35401	69269	33306	65792	0
Rest of CORC	2	5025	11612	4393	10713	0

*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	YPCORE
Service	77%	67%	76%	65%	45%
Rest of CORC	14%	17%	13%	16%	NA

*Corresponds to slide: What are the Follow Up Rates?*

## Cases with any paired outcome measure

Sample	Percentage	Count
Service	69.46%	17362
RoC	26.46%	32903

*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	16.8	16.6	16.9
Child SDQ Total Difficulties	2	12.8	12.6	13.0
Parent SDQ Total Difficulties	1	16.8	16.7	16.9
Parent SDQ Total Difficulties	2	12.6	12.4	12.7
Child SDQ Impact	1	2.4	2.3	2.5
Child SDQ Impact	2	1.0	1.0	1.1
Parent SDQ Impact	1	2.9	2.8	2.9
Parent SDQ Impact	2	1.2	1.1	1.2
YPCORE	1	16.6	14.8	18.1
YPCORE	2	11.4	9.7	13.1

*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	18.0	17.7	18.2
Child SDQ Total Difficulties	2	15.5	15.2	15.8
Parent SDQ Total Difficulties	1	19.8	19.6	20.0
Parent SDQ Total Difficulties	2	17.2	17.0	17.4
Child SDQ Impact	1	3.3	3.2	3.4
Child SDQ Impact	2	2.0	1.9	2.1
Parent SDQ Impact	1	4.6	4.5	4.7
Parent SDQ Impact	2	2.8	2.8	2.9
YPCORE	1	0.0	0.0	0.0
YPCORE	2	0.0	0.0	0.0

*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	-3.9	-4.1	-3.8
Parent SDQ Total Difficulties	-4.2	-4.4	-4.1
Child SDQ Impact	-1.4	-1.5	-1.3
Parent SDQ Impact	-1.7	-1.8	-1.6
YPCORE	-5.1	-7.1	-3.3

*Corresponds to slides on individual outcome measures*

## Means and confidence intervals: Rest of CORC

Measure	Mean	Lower limit	Upper limit
Child SDQ Total Difficulties	-2.5	-2.7	-2.3
Parent SDQ Total Difficulties	-2.6	-2.7	-2.4
Child SDQ Impact	-1.2	-1.3	-1.1
Parent SDQ Impact	-1.8	-1.9	-1.7

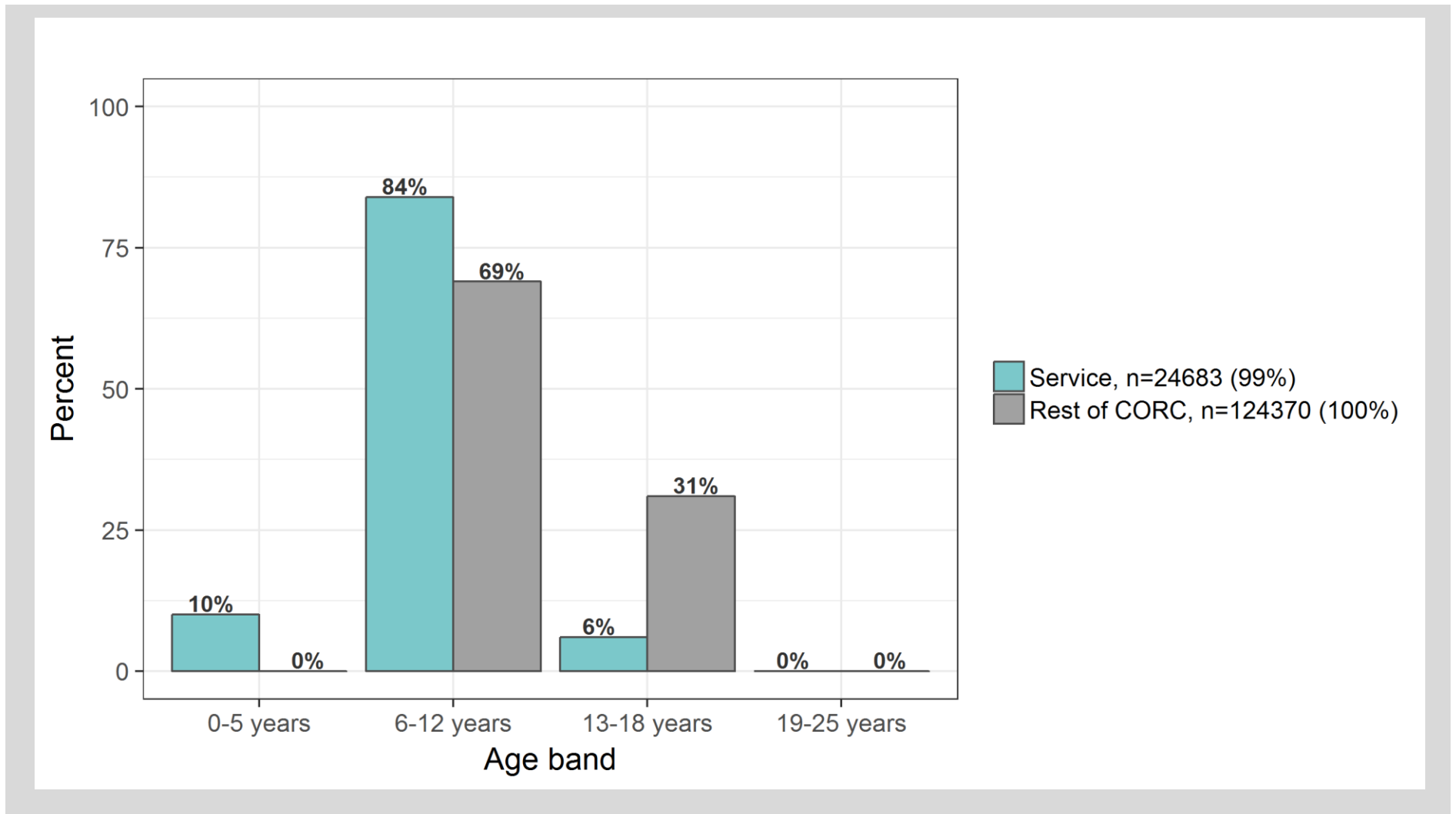
*Corresponds to slides on individual outcome measures*

# Appendix C

## Demographics & Case Characteristics

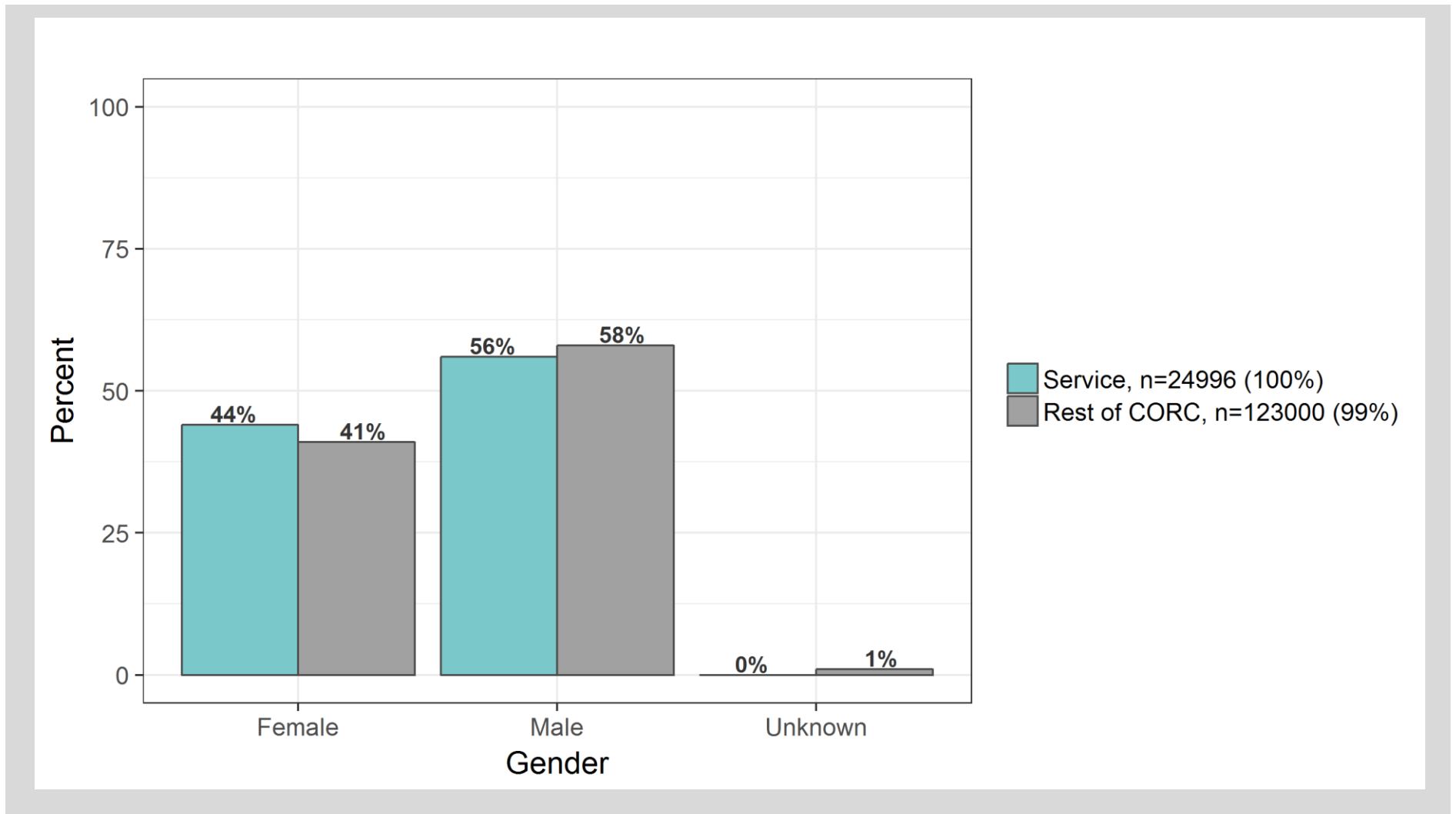


# Age Bands



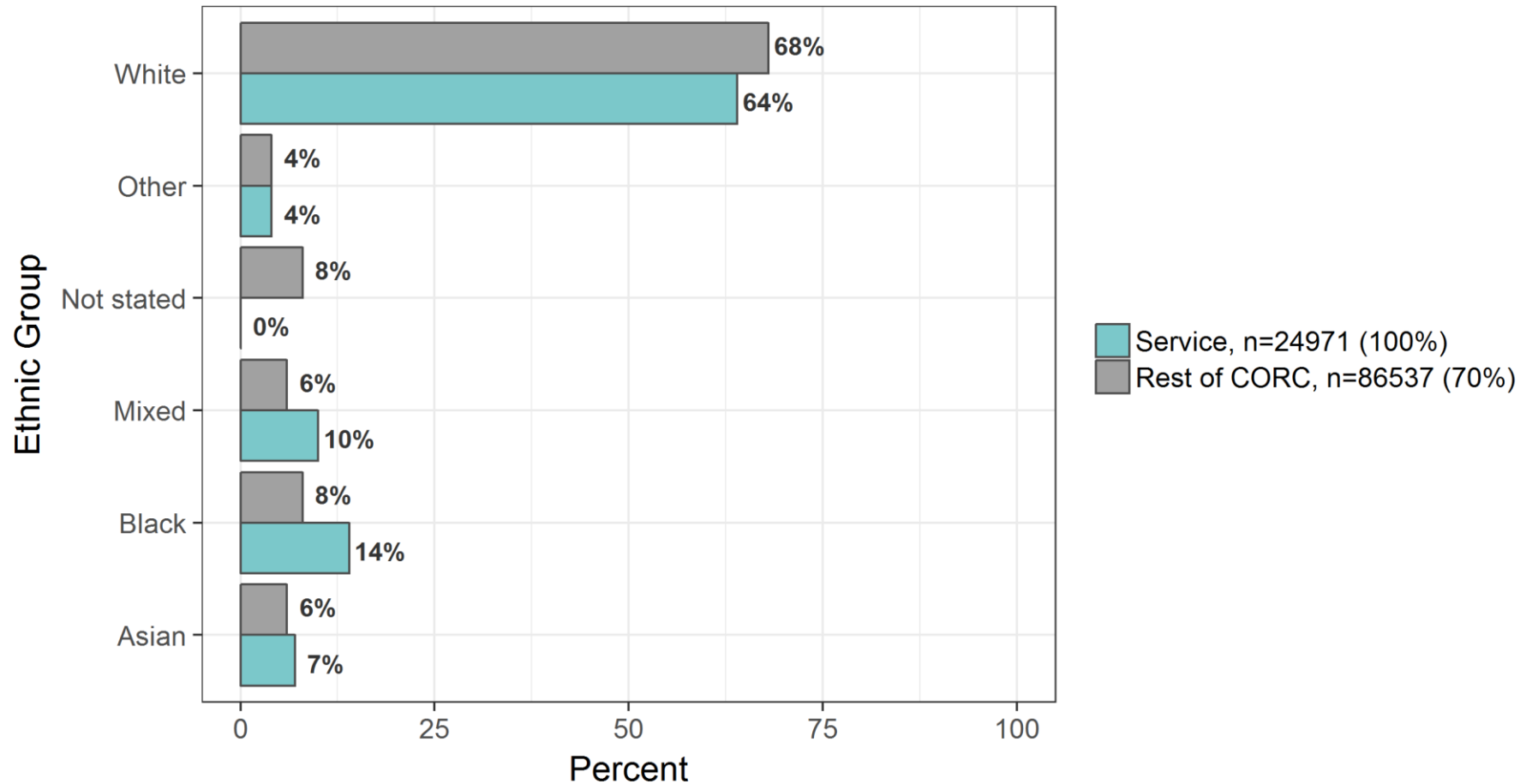
The samples' distribution for age bands for Place2Be and the Rest of CORC were significantly different (Chi-square  $X^2(\text{degrees of freedom: } 3, N = 149053) = 17979, p < .001$ ).

# Gender



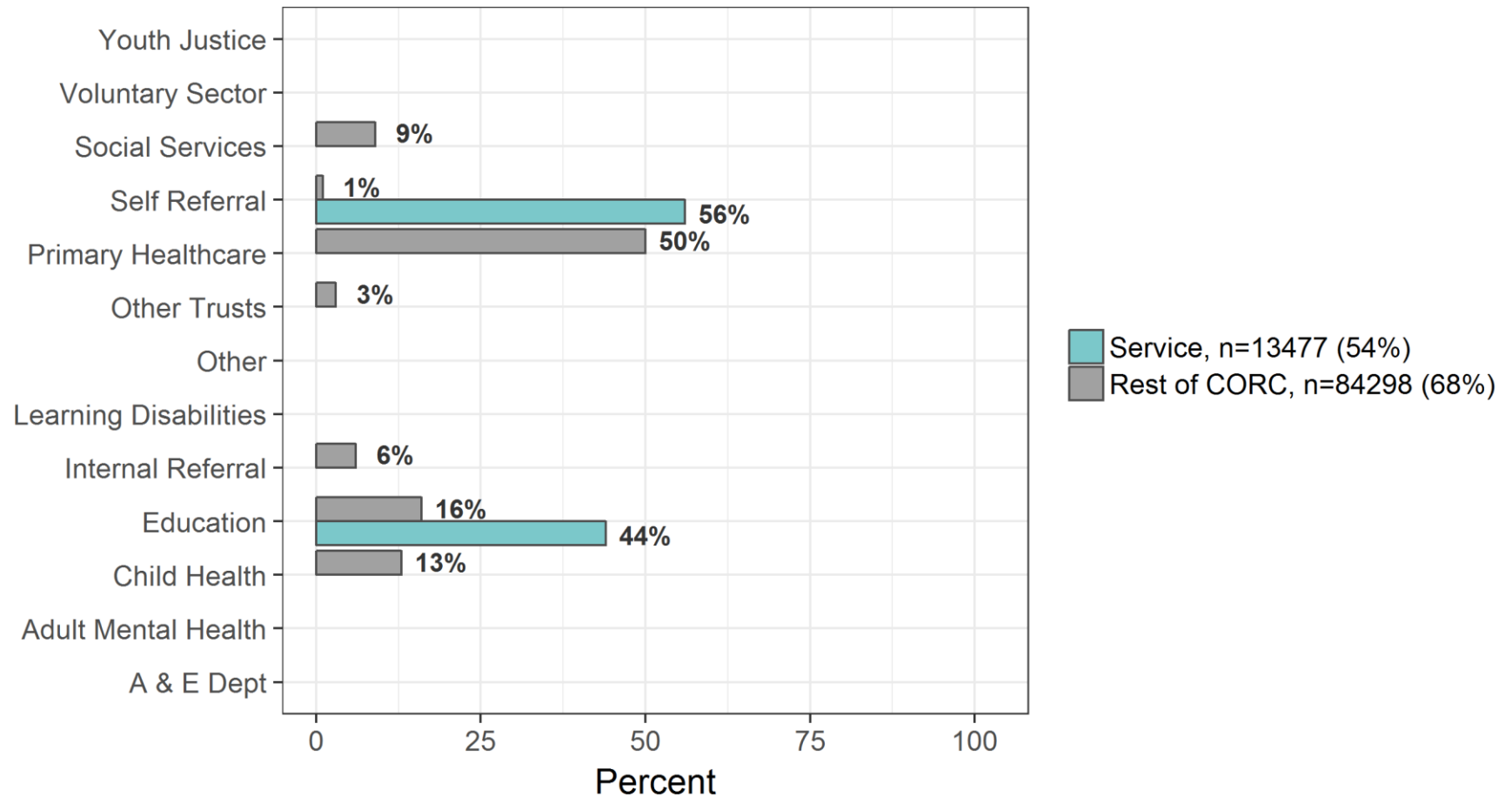
The samples' distribution for gender for Place2Be and the Rest of CORC were significantly different (Chi-square  $X^2(\text{degrees of freedom: } 2, N = 147996) = 182.25, p < .001$ ).

# Ethnicity



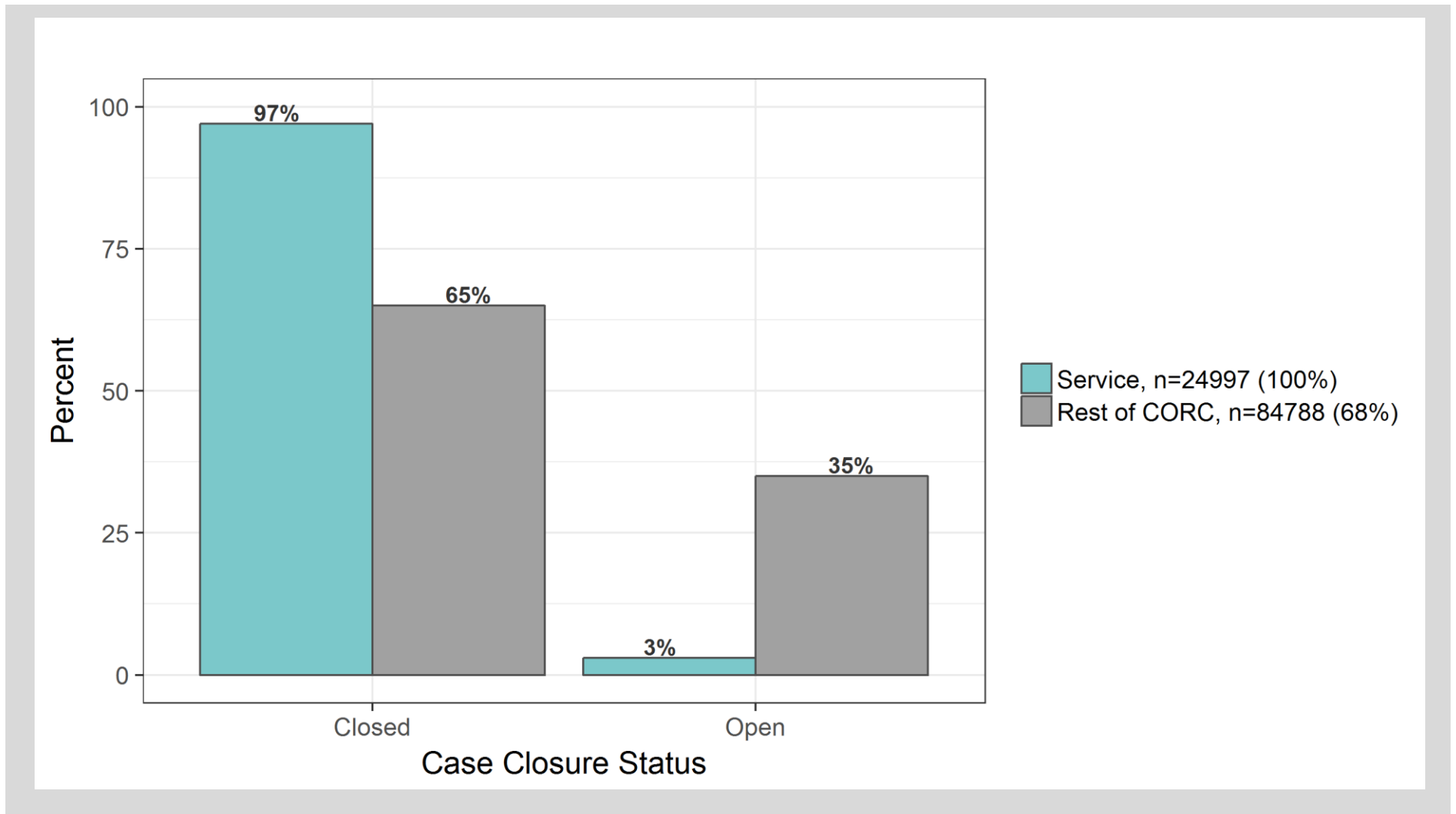
The samples' distribution for ethnicity for Place2Be and the Rest of CORC were significantly different (Chi-square  $X^2(\text{degrees of freedom: } 5, N = 111508) = 3378.1, p < .001$ ).

# Referral Sources



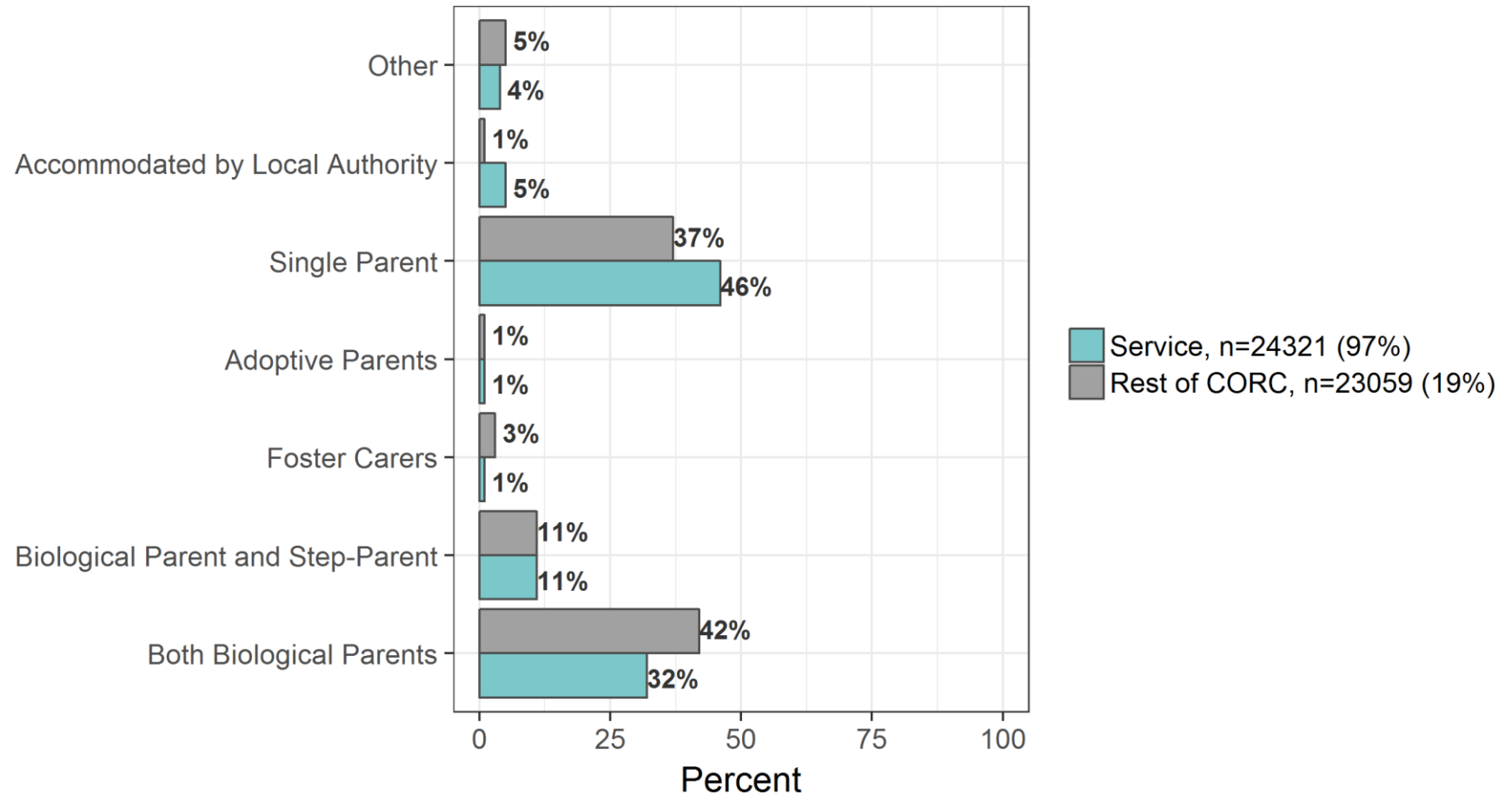
The samples' distribution for referral sources for Place2Be and the Rest of CORC were significantly different (Chi-square  $X^2(\text{degrees of freedom: } 12, N = 97775) = 56736, p < .001$ ).

# Case Closure Status



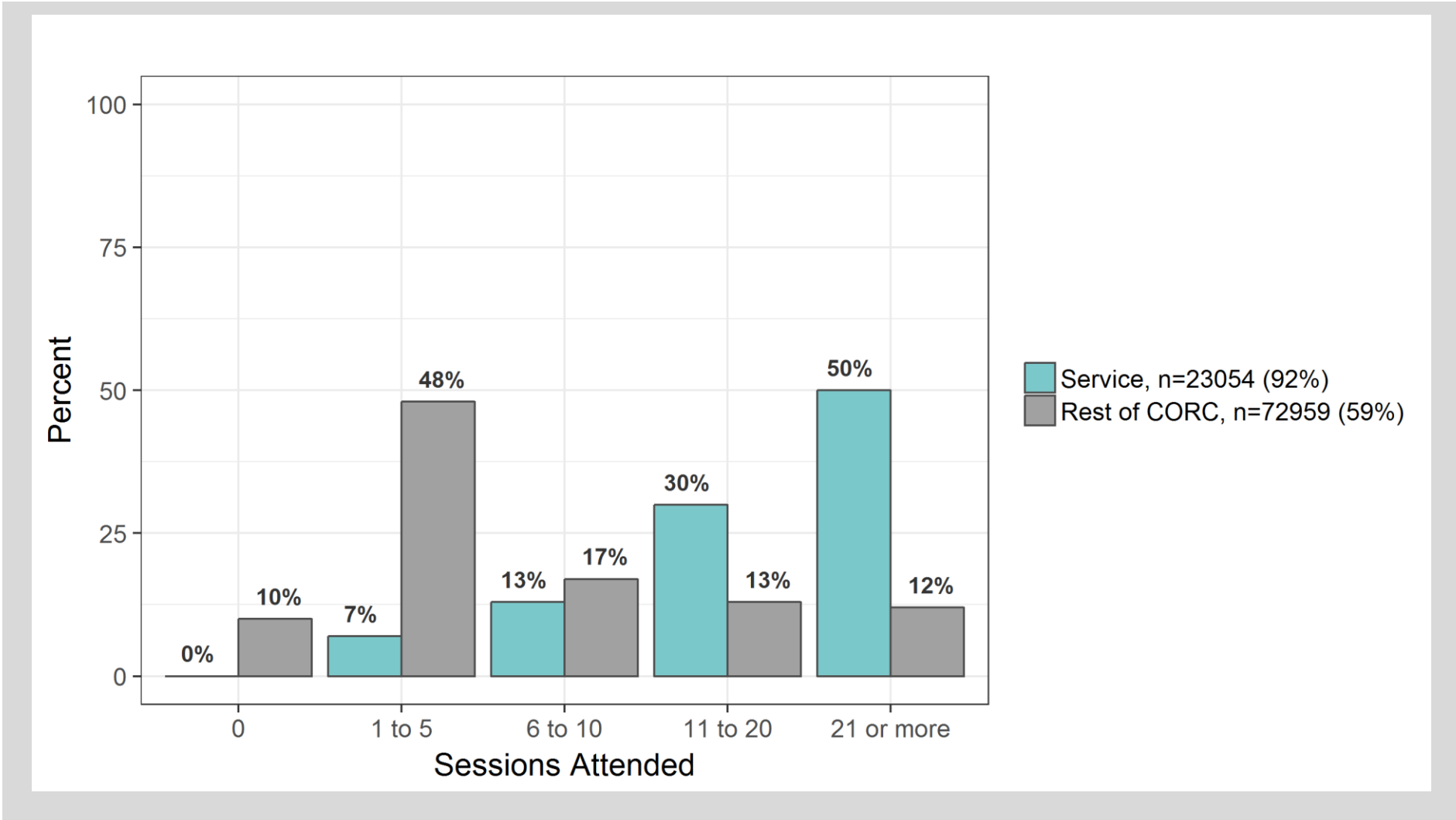
The samples' distribution for case closure status for Place2Be and the Rest of CORC were significantly different (Chi-square  $X^2(\text{degrees of freedom: } 1, N = 109785) = 10242, p < .001$ ).

# Family Composition



The samples' distribution for family composition for Place2Be and the Rest of CORC were significantly different (Chi-square  $X(\text{degrees of freedom: } 6, N = 47380) = 1735.9, p < .001$ ).

# Sessions Attended



The samples' distribution for number of sessions attended for Place2Be and the Rest of CORC were significantly different (Chi-square X(degrees of freedom: 5, N = 149367) = 42368, p < .001).

# Appendix D

Reliable Change for Individual SDQ  
Measures



# Reliable Change on Individual SDQ Measures

Reliable Change was calculated for each SDQ measures based the following values:

SDQ Child Total Difficulties: Cronbach's alpha = .78 (Becker at al. 2004a)

SDQ Parent Total Difficulties: Cronbach's alpha = .83 (Becker at al. 2004b)

SDQ Child Impact: Cronbach's alpha = .81 (Goodman, R. 2001)

SDQ Parent Impact: Cronbach's alpha = .85 (Goodman, R. 2001)

For each measure, we also calculated the standard deviation based on the whole sample (including both Place2Be and Rest of CORC data).

Becker, A., Hagenberg, N., Roessner, V., Woerner, W., & Rothenberger, A. (2004a). Evaluation of the self-reported SDQ in a clinical setting: Do self-reports tell us more than ratings by adult informants?. *European child & adolescent psychiatry*, 13(2), ii17-ii24.

Becker, A., Woerner, W., Hasselhorn, M., Banaschewski, T., & Rothenberger, A. (2004b). Validation of the parent and teacher SDQ in a clinical sample. *European Child & Adolescent Psychiatry*, 13(2), ii11-ii16.

Goodman, R. (2001). Psychometric Properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40(11), 1337-1345.

# Appendix E

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

# “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”.

A large range of outcome measures are included in the reliable change calculation, including sub-scales of the SDQ and the YP-CORE as well as sub-scales of the Revised Children’s Anxiety and Depression Scale, the Patient Health Questionnaire – 9, General Anxiety Disorder – 7 and others.

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

# “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”.

A large range of outcome measures are included in the reliable change calculation, including sub-scales of the SDQ as well as sub-scales of the Revised Children’s Anxiety and Depression Scale, the Oppositional Defiant Disorder Rating Scale and others.

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 F

Outcome measures

# Outcome Measures

## **Strengths and Difficulties Questionnaire (SDQ)**

[Questionnaire Versions and Scoring](#) (Developer's Website)

## **Young Person CORE (YPCORE)**

[Questionnaire Versions](#) (Developer's Website)

# Appendix G

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

# OUR DATA

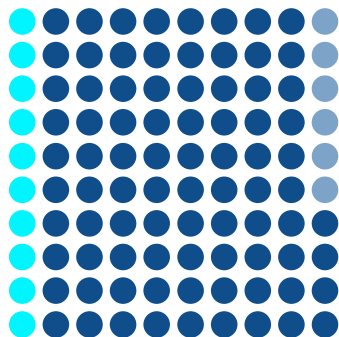
01 May 2008 to 30 November 2019

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.

24997 people were seen by the service

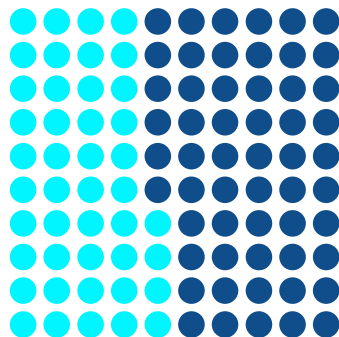
If the service saw 100 people:

Age  
People would be:



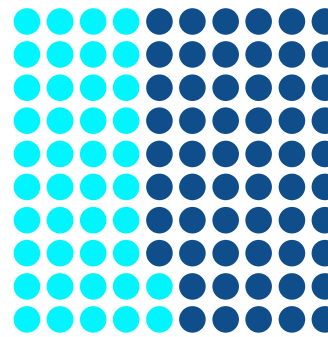
● 0-5 years   ● 6-12 years  
● 13-18 years

Gender  
People would be:



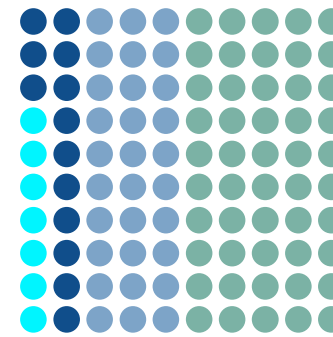
● Female  
● Male

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

69%

## Current completeness

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



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.