

Child- and Parent-reported Outcomes and Experience from Child and Young People's Mental Health Services 2011-2015

Miranda Wolpert, Jenna Jacob, Elisa Napoleone,
Andy Whale, Ana Calderon and Julian Edbrooke-Childs

November 2016

Child- and Parent-reported Outcomes and Experience from Child and Young People's Mental Health Services 2011–2015

Considering data from services who took part in the Children and Young People's Improving Access to Psychological Therapies (CYP IAPT) programme

Referencing this report

This report should be referenced:

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.

Press
CAMHS

ISBN: 978-0-9933436-2-9
© CAMHSPress 2016

CORC has made every effort to ensure accuracy within this report. However, due to the processing required to collect, collate and analyse the data, we cannot be held liable for omissions or inaccuracies.

Authorship

Miranda Wolpert, director of the Child Outcomes Research Consortium (CORC) and of the Evidence Based Practice Unit (EBPU), provided strategic oversight and led the writing of the report. Jenna Jacob (CORC research lead) led the data collection, support and analysis that underpin this report. Elisa Napoleone (CORC research officer) undertook the analysis with support and input from Dr Andy Whale (CORC statistician).

Dr Ana Calderon (EBPU post-doc research fellow) provided additional analytic input and Dr Julian Edbrooke-Childs (EBPU senior research fellow) provided analytic and research support.

by NHS England to support the use of outcome measures as part of the Child and Young People's Improving Access to Psychological Therapies (CYP IAPT) programme, as well as to analyse the data. Miranda Wolpert is national informatics lead for CYP IAPT at NHS England and chaired the outcomes and evaluation group (OEG) within CYP IAPT. CORC has a learning partnership with the Anna Freud National Centre for Children and Families and all authors are employed by the Centre which is involved in providing aspects of the CYP IAPT training, and whose chief executive Peter Fonagy is the clinical lead for CYP IAPT.

Potential conflicts of interest

While every effort has been made to analyse and interpret the data in an objective way, the authors would like to note the following for transparency: CORC was commissioned first by the Department of Health and then

Note on terms

Child/children is used in the report as shorthand for children and young people.

Parent is used in the report as shorthand for parent or carer.

Reliable change refers to amount of change in scores on a scale; it tells us whether change reflects more than the fluctuations of an imprecise measuring instrument. (Jacobson & Truax, 1991).

“Recovery” refers to scores having moved from above the threshold on that scale to below the threshold on that scale. It should be noted that different scales in this report have different thresholds, determined in different ways. Please note that “recovery” is referred to in inverted commas in this report when used to refer to this movement across threshold/s, as opposed to the broader concept of recovery as reflected in the lived experience of the individual concerned (e.g. Leamy, Bird, Le Boutillier, Williams & Slade, 2011).

Reliable “recovery” refers to change in scores on a scale that indicates both reliable change and “recovery” have occurred.

In this report these terms have been expanded to take account of multiple scales being used. Full details are provided in the text.



Contents

Acknowledgements	6
Aims	6
Using this report	7
Use of “flawed, uncertain, proximate and sparse” (FUPS) data	7
Reporting on FUPS data	7
Considering FUPS data	8
Key findings	9
Child-reported outcomes and experience	10
Parent-reported outcomes and experience	11
Executive summary	12
Chapter 1: Introduction	16
Considering outcomes in child health services	17
Considering outcomes in adult mental health services	17
Considering outcomes in child mental health services	17
Context of this report	20
Service transformation context	20
Service context	21
Chapter 2: Measures	22
Parent and child report measures	23
Measure of case characteristics	23
Case-mix algorithm	23
Chapter 3: Method	27
Data capture	28
Duplicates in the data	28
Data completeness and missing data	28
Possible reasons for missing data	29
Chapter 4: Samples of children	31
The seven samples	32
Chapter 5: Demographics and service provision data	34
Number of cases and characteristics of young people seen	35
Types of difficulties seen	36
Case mix (needs-based groupings)	39
Treatment received	40
Length of treatment	41



Contents

Chapter 6: Strengths and limitations of different analytic approaches	42
Raw service experience data	43
Raw change scores	43
“Recovery”	43
Reliable change	43
Reliable “recovery”	44
Standardised effect sizes	44
Added value score	44
Chapter 7: Considering experience and goals	46
Experience of service	48
Achievement of goals	50
Chapter 8: Considering “recovery”, reliable change and effect sizes	51
Sample characteristics	52
Number of scales completed	53
“Recovery”	53
Reliable change	53
“Recovery” or reliable change	55
Reliable “recovery”	55
Standardised pre-post effect sizes	55
Estimating added value	59
Chapter 9: Concluding remarks	60
References	62
Appendices	67
Appendix A: CYP IAPT: Service transformation programme	68
Appendix B: Outcomes and evaluation group	72
Appendix C: KPI summary table	76
Appendix D: Thresholds and reliable change indices for child- and parent-reported outcome measures	78
Appendix E: Current View: development and algorithm	80



Acknowledgements

The authors would like to thank all the young people, parents, clinicians and managers who shared their data to support this report. We would especially like to acknowledge the tireless efforts of the data leads across services, who worked extremely hard to collect the data and flow it centrally on a quarterly basis for collation and analysis in the face of many challenges and constraints.

Many thanks are due to the child mental health policy team at NHS England, in particular Kathryn Pugh, Margaret Oates, Anne O’Herlihy and Faye Henney, for their leadership and continuous support throughout the data collection and analysis process and provision of information about the Children and Young People’s Improving Access to Psychological Therapies (CYP IAPT) programme. Particular thanks are due to Dr Jacqueline Cornish who chaired the expert reference group overseeing the programme and championed the use of data across child physical and mental health.

We would like to thank all members of CORC and its Board at the time of writing: Ashley Wyatt, Isobel Fleming Duncan Law, Julie Elliott, Ann York, Mick Atkinson, Alan Ovenden and Kate Martin, and the CORC team: Matt Barnard, Victoria Zamperoni, Lily Levy, Kate Dalzell, Jenny Bloxham, Amy MacDougall, Ben Ritchie, Craig Hamilton, Sally Marriott, Alison Ford, Lee Atkins, Danielle Antha, Claudia Kanow, Rebecca Neale, Emma Davies and Alice Kedge.

We would also like to thank Tom Patterson at Click Databases for his input and support on this project, and the project team at MegaNexus: Lee Murray, Ailin Tarbinian and Dan Brown, for secure data upload and storage.

We would like to thank Jenny Bloxham for oversight of design, and Helen Waller for her excellent work designing and typesetting the report.

We would like to thank Martin Utley for his central and equal role in developing the idea of FUPS data and principles for analysis of FUPS data – all flaws in the use of this concept are the authors’ alone.

We would like to thank Duncan Law, David Trickey, Ben Ritchie, Rosa Town, Judith Shipman, Jess Deighton, Peter Fonagy, Matt Barnard and Peter Martin for comments on earlier versions, and Terence Stephenson for coining the phrase: “parity of esteem requires parity of data”. We note that all decisions on content were from the authors and may not be in line with the views of those whose helpful comments shaped the report.

Aims

The aim of this report is to share analysis of the routinely collected data related to outcomes for children seen across services taking part in CYP IAPT between April 2011 and December 2015.

The objectives of this report are to:

- 1. Advance the understanding of outcomes and experience of children accessing services**
- 2. Highlight the challenges encountered and suggest ways to address where possible**
- 3. Consider the best ways to measure and capture outcomes in the future.**

It should be noted that this report is not an evaluation of the programme nor is it the final analysis of the dataset. It has been agreed that CORC will continue to hold the collated dataset and will look at it in more depth than could be done within the remit of the current commissioned work.

Initial ideas for future analysis are shared in the report but we welcome thoughts from readers of this report.

Please get in touch:

CORC@annafreud.org

Using this report

Use of “flawed, uncertain, proximate and sparse” (FUPS) data

It is increasingly argued that improving quality services should involve more scrutiny of routinely collected data (Keogh, 2013). The aspiration is that data can be used to establish benchmarks for quality assurance and underpin the evaluation of quality improvement initiatives (Coulter, Locock, Ziebland & Calabrese, 2014). Such data across a wide range of datasets are, however, frequently flawed, uncertain, proximate and sparse; hereto referred to as FUPS (flawed, uncertain, proximate and sparse) data. They are flawed, due to missing or erroneously recorded data; uncertain, due to differences in how data items are rated and/or variation in case mix; proximate, in that they are always a proxy for an indication of the impact of the service provided; and sparse, in that even within complete datasets the low volume of cases within a given subgroup often limits the applicability of statistical inference. Use of such data for quality assurance and improvement is frequently contentious, with debates typically centring on the correctness of data and the analysis or statistical interpretation employed, rather than the utility of data analysis (Lilford, Mohammed, Spiegelhalter & Thomson, 2004; Wolpert et al., 2014b).

Child mental health service experience and outcome data in this report meet criteria for FUPS:

- **Flawed:** there is a high level of missing data.
- **Uncertain:** measurement is by child and parent report with all the uncertainty that involves.
- **Proximate:** metrics used are all proxies for impact and can't be causally linked to provision.
- **Sparse:** data for some important groups involve very small numbers.

CORC believes that data, at least in the field of child mental health, may remain flawed, uncertain, proximate and sparse for some time – arguably long enough to warrant coining the acronym FUPS.

Reporting on FUPS data

CORC is aware that this report could be (and very likely will be) criticised for reporting on data where there are so many questions about the quality and such a high degree of missing data. Respected colleagues might well argue that it is inappropriate to even report findings as they may lead to fallacious conclusions based on flawed data and be used for unhelpful ends. There is much to support such an argument. However, CORC's view is that analysis and sharing of such data is useful intelligence which can inform dialogue amongst key stakeholders, and is vital to advancing the field. We recognise this is not an uncontentious position.

In order to support best use of such FUPS data we have followed best practice principles suggested in relation to the use of FUPS data. These principles were developed by Professor Miranda Wolpert in collaboration with Professor Martin Utley but all errors or infelicities of expression are the current authors' alone.

As data analysts of FUPS data CORC sees their role as being to:

- Help build a conversation around the data rather than providing definitive answers.
- Provide accessible descriptive analyses first and foremost, and only to undertake statistical tests where there is a clear reason to do so.
- Present data in such a way as to convey any limitations to the interpretation of data, stemming from the following: small volumes of cases, rare events and the intrinsically partial nature of any risk adjustment.

In trying to follow best practice advised by Wolpert and Utley in relation to analysis of FUPS data, CORC has tried to:

- Use precise and neutral language. For example, axis labels are factual (what was measured) rather than interpretive (performance or quality of care). When referring to comparisons between groups, we have avoided the use of the terms “significance” or “performance data”.
- Provide full and precise definitions for metrics used in all cases.

- Include in displays and reports the raw numbers that analyses are based on, not just percentages and ratios in isolation.
- Make explicit where cases have been removed from analysis due to issues of data completeness or quality.
- Remind ourselves and readers that analysis may be limited, may not account for subtle clinical points and may contain mistakes.
- Respect and abide by agreed processes.
- Avoid “black boxes”; for example, complex stats on very limited data.
- Ensure those considering the data have time to reflect and absorb the information.

CORC believes that routinely collected outcome and experience data in child mental health are likely to remain FUPS for some time. CORC argues that if such data are to act as a form of intelligence to support thinking and decision-making, and as a spur to improved data collection, it is essential to start to examine what data we have as well as argue for improved data – to walk the fine line between scientific rigour and scientific rigor mortis.

Of course it is important to call for more and better data collection and higher quality data. However, it is only through examination of such FUPS data that can we start to have more informed debates about what outcomes should be expected to be achieved by those seeking help from child mental health services. It is in this spirit that CORC and collaborators present the findings in this report.

Considering FUPS data

CORC recommends that this report is used to inform facilitated stakeholder discussions involving practitioners, funders, service users, policy makers and others, along the lines outlined elsewhere and described as the MINDFUL approach (Wolpert et al., 2014b).

It is recommended that the facilitator of such conversations should seek to help those present to:

- Challenge their own and colleagues’ confirmatory biases.
- Maintain curiosity.
- Apply the same standards of scrutiny to analytic findings that support prior beliefs as to analytic findings that are uncomfortable or not wished for.
- Consider if any actions need to be taken in terms of quality assurance.
- Consider possible initiatives that even if not definitively indicated may do more good than harm.
- Challenge the assumption that change is always more risky than status quo.
- Help ensure agreed rules of engagement are adhered to.

To do this the facilitator should:

- Determine which groups are best brought together in which combinations: e.g. commissioners of services, service users, members of the public, practitioners, policy makers and researchers.
- Set clear ground rules for conversations (e.g. no point scoring, atmosphere of general interest, welcome critical thinking, focus on possible next steps and options that can aid best practice) with an agreed process for making a decision; however imperfect that process is.



Key findings

Key findings

Context: The analysis is of routine data related to outcomes and experience for children and young people (0–25 years old) seen across 75 services taking part in the service transformation initiative: Children and Young People’s Improving Access to Psychological Therapies (CYP IAPT) (April 2011 to Dec 2015).

Caution needs to be taken in interpretation, given data here are flawed, uncertain, proximate, and sparse. In particular, the unknown bias due to missing values means these data are not necessarily applicable to the wider child mental health service population. Better quality data are needed in the future, but in the meantime it is hoped that these data may help inform dialogue on outcomes.

Challenges: Measuring outcomes and experience in child mental health from the perspective of child and parent is challenging due to the diversity of population, perspectives, measures, metrics and lack of counterfactuals, comparisons and quality data.

Cases: 96,325 cases accessed services (mean age 12 years, 52% female, 82% white). 42,798 (44% of all cases) had practitioner ratings. Of these, approximately:

- **1 in 2¹ had family relationship difficulties**
- **1 in 3 had self-harmed**
- **1 in 8 had experience of abuse.**

Case mix: 31,037 cases (32% of all cases) had data allowing use of a case-mix algorithm based on practitioner rating. Of these, approximately:

- **1 in 2 had problems not readily assignable to NICE-recommended treatment.**

Closed cases: 23,373 treatment cases were closed (mean age 12 years, 57% female, 80% white).

Child-reported outcomes and experience

Children who provided outcome data were more likely to be older and female than children in the overall dataset.

Child report: 21 possible child report scales were used in this report (19 with thresholds, each with different thresholds).

Child-reported experience: 3,196 cases (14% of all closed treatment cases) reported on experience of service (mean age 14 years, 65% female, 84% white). Of these, approximately:

- **4 in 5 endorsed receiving good help**
- **3 in 5 endorsed convenient appointments.**

Child-reported movement towards goals: 2,784 cases (12% of all closed treatment cases) had paired data on goals (mean age 13 years, 65% female, 88% white). Of these, approximately:

- **9 in 10 reported movement towards the agreed goals.**

Child-reported outcomes: 5,918 cases (25% of all closed treatment cases) had paired child report data where one or more scales were above threshold at outset (on average four scales completed per child). (Mean age 14 years, 72% female, 87% white). Of these, approximately:

- **1 in 2 reliably improved** (amount of change in score more than likely due to measurement error and no reliable change in the opposite direction on any scale)
- **1 in 3 “recovered”** (no scale above threshold at end of treatment)
- **1 in 4 reliably “recovered”** (all scales below threshold at end of treatment with reliable improvement in scores on at least one measure which also shows “recovery” and no measure showing reliable deterioration)
- **1 in 10 reliably deteriorated** (counted as deteriorated if happened on any scale).

¹ Please note that numbers are presented as fractions in the key findings and executive summary, in an attempt to make the figures more accessible and easier to digest for a general audience.

Parent-reported outcomes and experience

Parents who provided outcome data tended to be providing data in relation to younger and male children than children in the overall dataset.

Parent report: 15 possible parent report scales were used in this report (13 with thresholds, each with different thresholds).

Parent-reported experience: 2,698 cases (12% of all closed treatment cases) reported on experience of service (mean age 11 years, 52% female, 82% white). Of these, approximately:

- **9 in 10 endorsed receiving good help**
- **7 in 10 endorsed convenient appointments.**

Parent-reported movement towards goals: 686 cases (3% closed treatment cases) had paired data on goals (mean age 8 years, 37% female, 78% white). Of these, approximately:

- **9 in 10 reported movement towards the agreed goals.**

Parent-reported outcomes: 3,699 cases (16% of all closed treatment cases) had paired parent report data where one or more scales were above threshold at outset (on average four scales completed per child). (Mean age 11 years, 54% female, 85% white). Of these, approximately:

- **2 in 5 indicated child problems reliably improved** (amount of change in score more than likely due to measurement error and no reliable change in the opposite direction on any scale)
- **3 in 10 indicated child problems “recovered”** (no scale above threshold at end treatment)
- **1 in 6 indicated child problems reliably “recovered”** (all scales below threshold at end of treatment with reliable improvement in scores on at least one measure which also shows “recovery” and no measure showing reliable deterioration)
- **1 in 10 indicated child problems reliably deteriorated** (counted as deteriorated if happened on any scale).



Executive Summary

The following pages give an overview of the findings from an analysis of child- and parent-reported outcomes and experience from child and young people's mental health services in the UK. The analysis is of routine data related to outcomes for children collected between April 2011 and December 2015 from services taking part in the Children and Young People's Improving Access to Psychological Therapies (CYP IAPT) programme.

Context

- **Some areas of physical health care have been able to use data to drive service improvement.** Historically there has been a lack of data in child mental health.
- **Measuring outcomes and experience in child mental health from the perspective of child and parent is challenging** due to the diversity of population, perspectives, measures, metrics and lack of counterfactuals, comparisons and quality data.
- **Data are from cases seen in services which were part of a service transformation initiative.** Child and young people's mental health services sought to embed evidence-based treatments, user participation and use of routine child- and parent-reported outcomes and feedback systems.
- **There were challenges to implementation and data collection.** Historically underfunded services and lack of developed infrastructure combined with increased demand of around 11% per annum, alongside service cuts of up to 75% and major IT difficulties, as well as practitioner concerns about use of measures, resulted in significant challenges for services and for data collection.

Data quality

- **Limitations of data quality and quantity mean all findings have to be treated with caution.** In particular, given the unknown (unmeasured) bias due to missing values and the unknown selection mechanisms into different outcome measures, these data are not necessarily applicable to the wider child mental health service population.
- **This dataset is made up of data that are flawed, uncertain, proximate and sparse (FUPS).** Many might argue that, in the light of this, such data are not safe to use. CORC argues that whilst recognising

the limitations and cautions necessary above, FUPS data are likely to be all we have for some time and, with care, can be used to support dialogue about how to set realistic targets for child mental health outcomes in the future, which can inform service improvement and development.

Measures

- **21 children and 15 parent-reported scales were used in this report (19 child and 13 parent scales included thresholds).**
- **Different scales use different approaches to calculate thresholds.** This presents challenges for comparing across scales and populations.

Data completion

- **Target for data completion in relation to key outcome metrics was not met.** A target of 90% of all closed treatment cases (at least one paired child- or parent-reported scale with thresholds, and a measure of attainment and attendance in relation to employment education or training) was set.
 - By December 2015, 5,240 (22% of closed treatment cases) had a paired scale with thresholds, as well as information on attainment and attendance.

Demographics

- **96,325 cases seen:** age range 0–25, mean age 12 years, 82% white, 52% female.
- Of the 96,325 cases seen, 42,798 (44%) had practitioner ratings of at least some case characteristics. Of these, approximately:
 - **1 in 2 had family relationships difficulties**
 - **1 in 3 had self-harmed**
 - **1 in 5 reported parental health issues**
 - **1 in 8 had experience of abuse.**

Completed cases

- **23,373 closed treatment cases** (seen for more than just assessment). Of these, 17,055 (73%) had completed a child- or parent-reported scale with thresholds.
- Of the 17,055 closed treatment cases with child- or parent-reported scale with thresholds, 15,536 had scores on one or more scales at outset. Of these, approximately:
 - **9 in 10 of those with completed scales with threshold at outset were above threshold on one or more child- or parent-reported scales at the start of treatment.**
- Of 31,037 cases where a case-mix algorithm based on practitioner ratings was applied, approximately:
 - **1 in 2 had problems not readily assignable to a NICE-recommended treatment.**

Outcome and experience

Child-reported outcomes were more likely to be from older and female respondents than being representative of the full sample. Parent-reported outcomes were more likely to relate to younger and male children than being representative of the full sample.

A range of approaches was used to consider outcomes and experience for the 23,373 closed treatment cases.

Experience of service

- Of the 3,196 cases (14% of all closed treatment cases) where children reported on their experience of the service, approximately:
 - **4 in 5 endorsed receiving good help**
 - **3 in 5 endorsed convenient appointments.**
- Of the 2,698 cases who had completed a course of treatment and where parents reported on their experience of the service (12% of all closed treatment cases), approximately:
 - **9 in 10 endorsed receiving good help**
 - **7 in 10 endorsed convenient appointments.**

Movement towards achieving goals

- Of the 2,784 cases (12% of all closed treatment cases) who had paired data on child goals, approximately:

- **9 in 10 reported movement towards the agreed goals.**
- Of the 686 cases (3% of all closed treatment cases) who had paired data on parent goals, approximately:
 - **9 in 10 reported movement towards the agreed goals.**

Considering outcomes for closed treatment cases with at least one child or parent report scale above threshold at outset:

“Recovery” rates were calculated using an adaptation of the Adult IAPT approach.

- Of the 5,918 cases (25% of all closed treatment cases) with paired child report data where one or more scales were above threshold at outset (on average four scales completed per child), approximately:
 - **1 in 3 children “recovered”** (no scale above threshold at end of treatment).
- Of the 3,699 cases (16% of all closed treatment cases) with paired parent report data where one or more scales were above threshold at outset (on average four scales completed per child), approximately:
 - **3 in 10 parents indicated child problems “recovered”** (no scale above threshold at end of treatment).

Reliable improvement rates were calculated using an adaptation of the Adult IAPT approach.

- Of the 5,918 cases (25% of all closed treatment cases) with paired child report data where one or more scales were above threshold at outset (on average four scales completed per child), approximately:
 - **1 in 2 children reliably improved** (amount of change in score more than likely due to measurement error and no reliable change in the opposite direction on any scale)
 - **1 in 10 children reliably deteriorated** (counted as deteriorated if happened on any scale).
- Of the 3,699 cases (16% of all closed treatment cases) with paired parent report data where one or more scales were above threshold at outset (on average four scales completed per child), approximately:

- **2 in 5 indicated child problems reliably improved** (amount of change more than likely due to measurement error and no reliable change in the opposite direction on any scale)
- **1 in 10 indicated child problems reliably deteriorated** (counted as deteriorated if happened on any scale).

Reliable “recovery” rates were calculated using an adaptation of the Adult IAPT approach too.

- Of the 5,918 cases (25% of all closed treatment cases) with paired child report data where one or more scales were above threshold at outset (on average four scales completed per child), approximately:
 - **1 in 4 reliably “recovered”** (all scales below threshold at the end of treatment with reliable improvement in scores on at least one

measure which also shows “recovery” and no measure showing reliable deterioration).

- Of the 3,699 cases (16% of all closed treatment cases) with paired parent report data where one or more scales were above threshold at outset (on average four scales completed per child), approximately:
 - **1 in 6 indicated child problems reliably “recovered”** (all scales below threshold at the end of treatment with reliable improvement in scores on at least one measure which also shows “recovery” and no measure showing reliable deterioration).

Other analyses were also undertaken, including pre-post effect sizes for each measure with paired data available, and the added value score metric for paired data from the parent-reported Strengths and Difficulties Questionnaire. These are reported in full in the text.

Concluding remarks

- This report is the first analysis of routinely collected child- and parent-reported outcomes and experience data collated from 75 mental health services in England.
- A range of outcome measures was used for the primary purpose of informing clinical decision-making with individual children and parents.
- In the absence of a commonly agreed method for analysing the data at a national level, several approaches to estimating outcomes were applied. Each method indicated different levels of change in the mental health, wellbeing or personal goals of a selected sample of children and young people using services.
- The findings are hindered by poor data completeness and the lack of a comparator group and should therefore be interpreted with caution.
- Although the findings are based on data that are flawed in a range of ways, and with

a high degree of missing data (which we have termed FUPS), they are the best currently available and should be used carefully to support dialogue about outcomes.

- CORC recommends this report is used to inform facilitated stakeholder discussions involving practitioners, funders, service users, policy makers and others. Key topics for consideration in such facilitated discussions might include the appropriate metrics for considering outcomes and appropriate targets for given metrics.
- Better quality data on outcomes and experience must be facilitated and incentivised to aid review and development of services. To achieve this, leadership focus, improved IT systems, better staff training and stronger incentives may need to be in place.

Achieving parity of esteem between physical and mental health requires parity of data.



Chapter 1:
Introduction

This report sets out the findings from an analysis of child- and parent-reported outcomes and experience from child and young people’s mental health services in the UK. The report analyses routine data related to outcomes for children collected between April 2011 and December 2015 from services taking part in the Children and Young People’s Improving Access to Psychological Therapies programme (CYP IAPT). This introduction describes the context for considering outcomes in child mental health services, the context of measure and scale selection, as well as data collection and analysis for this initiative.

Considering outcomes in child health services

Some areas of health care related to physical health for children have been able to use outcome data powerfully to drive service improvement. For example, since routine review of outcomes was inaugurated in paediatric diabetes, positive recorded outcome rates (based on control of HbA1c to under a threshold of 7.5) have risen from 14.5% in 2009–10 to 23.5% in 2014–15 (National Paediatrics Diabetes Audit, 2015). However, it is important to recognise that these data are drawn from hard biological markers that are used in clinical practice which allow ready capture and clear consideration (and even then there are issues with how best to compare data). Other areas of child health, such as asthma and epilepsy, have no clear indicators of outcome or experience and therefore no national way of considering outcomes.

Considering outcomes in adult mental health services

There are many challenges to finding the best way forward in relation to considering mental health outcomes for children. Some of these challenges are held in common with other areas of mental health and some are more specific to the child context. In common with adult mental health is the issue of what constitutes a positive outcome and whether problems should be considered as categorical or dimensional. For example, there is much debate as to how far a positive outcome should be defined in terms of symptom change or more broadly in terms of functioning, relationships or wellbeing, with a particular focus on the idea of recovery related to personal lived experience (e.g. Leamy et al., 2011). Even within a focus on symptoms, there is increasing interest in considering symptoms as reflecting a more general underlying factor of vulnerability to mental ill-health (Caspi et al., 2014; Patalay et al., 2015).

Within adult mental health, the most prominent protocol for considering patient-reported outcomes in the UK is that developed by the Improving Access to Psychological

Therapies (IAPT) programme (Clark & Oates, 2014). This approach was taken as the starting point for the analysis in this report although it was recognised that there were many challenges with adapting this approach in the more complex environment of child mental health and additional approaches were also explored (see Chapter 6).

Considering outcomes in child mental health services

The lack of outcome data in child mental health has been identified as a central problem for the development of services (Department of Health, 2015). Collaborating members of the Child Outcomes Research Consortium (CORC), which as a learning collaboration constitutes the majority of child mental health providers across England, have for some years been leading the way in the ambitious task of trying to find the best way to collect mental health outcome and experience data from the perspective of children and parents and the challenges have been well documented (Fleming, Jones, Bradley & Wolpert, 2016). All agree that child mental health desperately needs to have some understanding of outcomes achieved in order to start to consider appropriate baselines and benchmarking. As has been argued elsewhere, the absence of such data may foster an unchallenged assumption that services can help everyone if they just try hard enough. This can lead to both practitioners and service users feeling blamed for poor outcomes and act as a disincentive to considering what outcome data there are and to genuine service improvement (Wolpert, 2016). It is for this reason that CORC commends the use even of data that are flawed, uncertain, proximal and sparse (FUPS) (as discussed on pages 7–8).

Collecting and using outcome data in the context of child mental health has particular complexities and challenges. Seven key challenges in particular can be identified: diversity of population, diversity of perspectives, diversity of measures, diversity of metrics, lack of counterfactuals, lack of comparison data and lack of quality outcome data. The implications of these challenges for the current report are considered in turn below.

Diversity of population

Child mental health services see a great diversity of difficulties, ranging from sleep problems in pre-school children to psychosis in adolescents, and include children and young people ranging from infants to 24 year olds. Problems include, but are not limited to, young people struggling with responding to trauma and abuse, children with neurological difficulties and a wide range of family relationship difficulties alongside depression and anxiety (see Chapter 5 for a discussion of the demographics of

the current sample). There is no clear agreed way to parse the population seen in services and much debate still rages over whether diagnostic categories are the best way of considering differences in this population and whether categorical or dimensional approaches are best suited to understanding the nature of child mental health difficulties. A parallel project for developing a case-mix analysis to underpin more targeted payment systems in child mental health services led to the development of needs-based groupings (see Chapter 5, Appendix E and Wolpert et al., 2015b). Whilst this report considers generic outcomes for the diversity of those seen in services (though effect sizes for different scales are presented separately), the proposal is to analyse outcomes for different groups in the future, in order to explore the impact of case characteristics.

Diversity of perspectives on outcomes and experience

There are no objective measures of mental health comparable to some areas of physical health, such as the Hba1c levels in diabetes. Research must therefore rely on proxies of outcomes including self-report questionnaire responses by those accessing services. Research has consistently found poor correlation between child, parent, teacher and practitioner views on the nature of difficulties and outcomes (Yeh & Weisz, 2001). Whilst there have been concerns raised about the reliability of data from younger children in particular, there is also a wish to prioritise the voice of the child wherever possible and some evidence that children as young as eight years old can comment on their own mental states (Deighton et al., 2013, 2014; Patalay, Deighton, Fonagy, Vostanis & Wolpert, 2014). As there is not a clear answer about which perspective to take and all have value in terms of understanding the impact of services (Wolpert et al., 2015a), different perspectives are analysed separately and with a particular focus on the view of the child seeking help, wherever possible.

It is important to note, however, that the parent report sample and the child report sample are likely to be different in a range of ways. For example, parents are more likely to be reporters for younger children (and male because more younger males attend services) and children are more likely to be reporters when they are older (and female because more older females attend services). This needs to be taken into account in considering differences of perspectives in this report.

The issue of practitioner perspective is also an important area. Whilst practitioner-rated outcomes are not considered in this report, they will be analysed in the future.

Diversity of measures of child- and parent-reported outcomes and experience

Due to the breadth of problems children and young people’s mental health services deal with, the different domains that may be relevant and important for those accessing services (including symptom change, general wellbeing, functioning and relationships), and the variety of perspectives noted above, there is a wide range of possible measures to support appropriate use in given clinical contexts. Some 54 different scales were included in the CYP IAPT programme as options for practitioners to use, 36 of which were used in one or more analyses in this report. It was recognised that in prioritising diversity of measures to allow clinical choice, there would be complexity in overall analysis of data. Moreover, there was no protocol that specified which measure should be used at which points, and it was left to clinical judgement to select when and which measure to use in order to support shared decision-making and collaborative working (Law & Wolpert, 2014; Section 5, British Psychological Society, 2015, Appendix 4).

Different scales use different approaches to calculate thresholds (see Appendix D for details). Determining the best method of aggregating scales in some way, so there is a single judgement about outcome in an individual case with multiple scales, and how to avoid conflating measure differences with differences in treatment outcome, was a major challenge which was considered in relation to a range of metrics outlined below.

Diversity of metrics to consider child- and parent-reported outcomes and experience

The range of possible metrics to consider outcomes and experience in child mental health has been reviewed elsewhere (Wolpert et al., 2015a, Deighton et al., 2014) and are summarised in a recent technical briefing by CORC. Key current indices of change and experience have been identified: raw experience of service data, raw score changes, “recovery”, reliable change, reliable “recovery” (a combination of “recovery” and reliable change), effect size calculations, and the added value metric. Each has been used in this report. The strength and limitations of each and their application are outlined in Chapter 6.

It is important to note that all the above possible metrics of outcome and experience share a common limitation in terms of difficulties of interpretation where there is a high degree of missing data, as in this report, and a lack of existing comparison data (see the sections below on the lack of comparison data and the lack of quality outcome and experience data). Moreover all of them, apart from the added value metric, share a common limitation of

lacking counterfactual controls which will be addressed in the next section below.

Lack of counterfactuals

It is not possible to determine the impact of service provision without an appropriate counterfactual. Child mental health problems often follow a fluctuating course. Determining if someone would have got better or worse without the treatment being offered is necessary to be able to interpret what an outcome means in terms of the impact of treatment, but this is challenging in routine care where there are no control groups, and in the absence of randomisation. One approach to considering counterfactuals in the context of a lack of control groups is to use propensity score matching whereby scores of children in a community sample who have not had treatment are compared with those seen for treatment. One metric that attempts to draw on a community sample as a quasi-control group (the added value score of the Parent Strengths and Difficulties Questionnaire) is used in this report and the results discussed in Chapter 6, but there is a lack of counterfactual information to draw on for the wider range of difficulties and perspectives.

Lack of comparison data

A review of the literature reveals few substantive reported findings of outcomes from routine practice in child mental health. What studies there are all show lower “recovery” and reliable improvement rates than for adults and higher rates of deterioration (Warren, Nelson, Mondragon, Baldwin & Burlingame, 2010; Nilsen, Handegard, Eisemann & Kvernmo, 2015). All existing studies draw on FUPS data; they report on small datasets which represent small percentages of overall samples with a high degree of missing data.

One American study comparing outcomes for children and young people (aged 4–17) considered change scores on the parent-completed Y-OQ in routine care in community services ($n=363$, mean age 12) versus those in managed care ($n=1,947$, mean age 13). These represented 27% and 70% of the sample with a Time 1 measure completed in each setting respectively; it was not reported how large the original sample was for either setting (Warren et al., 2010). The outcomes across the two settings (community/managed care) were as follows: 13%/19% showed reliable deterioration, 27%/33% showed no reliable change, 29%/32% showed reliable improvement, 15%/30% showed reliable “recovery”, and 1% showed subclinical deterioration. The authors comment that “the high percentage of cases in each setting showing negative outcome (deterioration or no significant change) is sobering” (p.151). They note this is in line with the limited

extant research in this area.

A more recent Norwegian study considered self- and parent-reported SDQ scores and clinician ratings (HoNSOCA and CGAS) for a small sample of children seen with emotional difficulties across seven services ($n=82$). The relationship of this sample to the larger sample was not reported and not all children had all measures, e.g. 13 (15%) had paired self-report SDQ (mean age 14, 70% female); 16 had paired maternal SDQ (mean age 12, 52% female) (Nilsen et al., 2015). The authors concluded that “the observation that relatively few children and youth with emotional disorders experience clinical significant and statistical reliable change, and that a considerable number had worsened emotional problems (according to parents) while in CAMHS treatment are not uplifting.”

The most recently published British-reported findings come from CORC’s analysis of data from collaborating services considering change in total difficulties scores on parent-completed SDQs from a sample of 9,764 children from across 58 services in England (mean age 11, 45% female) (Wolpert et al., 2015a). How this sample related to the larger sample was not reported in the study but it is known from other CORC analyses that the proportion of cases with paired outcome data from the larger dataset for CORC was around 24% (Wolpert et al., 2012). This study found that 21% moved from above the threshold to below (“recovery”), 5% moved from being below the threshold to above the threshold (37% were not above the threshold on parent-reported SDQ total difficulties scores at Time 1), with the remaining 73% unchanged in terms of movement across thresholds. In terms of analysis of reliable change (as defined by Jacobson & Truax, 1991), 16% of cases reliably improved, 2% reliably deteriorated and 82 % showed no reliable change (Wolpert et al., 2015a).

In the light of the small sample sizes and slightly different criteria used in each of the studies above, and to try, wherever possible, to address the issue of comparison groups, comparisons with pre-CYP IAPT data held by CORC have been considered in this report. These data were collected between 2007 and 2010 and comprised 118,884 episodes of care from children within this time period. The dataset was filtered to match the comparative closed treatment cases sample (3) in the CYP IAPT dataset, yielding a total of 28,955 episodes of care. It is important to note, however, that the nature and structure of the datasets were different, resulting in unavoidable challenges when comparing the datasets. For example, the earlier dataset captured a six-monthly snapshot of time, rather than information on an event level, and comprises data from fewer measures. Moreover, the earlier dataset can also be seen as an example of FUPS data; with high degrees of missing data and questions about quality of data in many fields.

If one looks wider than child mental health, some data from a specific adult mental health population are available from Adult Improving Access to Psychological Therapies (IAPT). IAPT works with adults who are below the threshold for mainstream mental health services and who have anxiety or depression to offer them early intervention to aim to address their symptoms and improve their employability. This project has set itself the target of achieving 50% “recovery”. Given the much greater diversity of difficulties of young people being seen in child mental health services, and the findings to date quoted above, it would be anticipated that “recovery” rates in child mental health may be lower than this.

Lack of quality outcome and experience data

As noted above, data quality and quantity have, historically, been poor in child mental health services, particularly in relation to information about outcomes (Fleming et al., 2016). A random case note audit (61 cases) in the East Midlands found only 6% of cases with paired child- or parent-reported outcome measures (Batty et al., 2013). The reasons for this may include the lack of robust tools or consensus on the best ones to use, the complexities of data collection in a multiagency environment with limited IT support (child mental health has been described as the Cinderella service of the Cinderella service that is mental health), the burden of data collection and/or a lack of practitioner investment or belief in the tools. A 2005 national survey of UK child mental health service leads found that less than 30% were implementing outcome measures ($n=186$) (Johnston & Gowers, 2005). These reasons are explored in greater detail in Chapter 3 which considers possible reasons for missing data in the current dataset.

Context of this report

In 2011 CORC led a collaboration (involving an academic partner: the Evidence Based Practice Unit at UCL and the Anna Freud National Centre for Children and Families and a secure data storage partner: MegaNexus Ltd) which successfully won a Department of Health tender to support the collection and use of routinely collected data from child and young people’s mental health services involved in CYP IAPT (see Appendix A for an overview of the nature of this service transformation programme), pending mainstreaming of data collection by the Health and Social Care Information Centre (since renamed NHS Digital). The project was initially for two years but was re-tendered in 2013 for a further two years by NHS England, as mainstreaming of data flow to NHS Digital was delayed.

CORC and collaborators were successful in gaining this further tender.

CORC’s role was to support services in collecting routine use of outcome and experience measures for children and families, to centrally collate and analyse these data and to help support the mainstreaming of the child mental health dataset being used by the project (which drew on the dataset originally developed by CORC) into NHS Digital.

It was recognised from the outset that this was a very ambitious endeavour. CORC had many years of experience of trying to collect and use child- and parent-reported outcome data to inform service improvement and the complexities and challenges have been highlighted (Fleming et al., 2016). The CYP IAPT programme drew on CORC’s experience and set those involved the ambitious task of achieving 90% data completion for paired outcome measures. This was defined as: paired child or parent report data (from specified scales – see Appendix B), completed for a closed treatment case involving at least three events (and excluding assessment) and where there was also practitioner report of educational, training or work attainment or attendance. This was one of a range for Key Performance Indicators agreed in relation to data – the full list of KPIs and the results obtained are included in Appendix C.

It was agreed that as part of this work that CORC would produce a report sharing key analyses that might help inform and advance understanding of outcome and experience measurement in the future. This is that report. It should be noted that this work was never intended as an evaluation of the programme nor is this the final analysis of the dataset. It was agreed that CORC would continue to hold the collated dataset and could continue to explore the data to advance understanding in this area.

Service transformation context

Data from this report are drawn from children and families seen in services across England who took part in CYP IAPT. This service transformation programme launched in 2011 involved geographical partnerships between NHS local authorities and voluntary sector providers (termed “partnerships”). Groups of partnerships were linked with specific higher education institutions (HEIs) across five areas (termed “collaboratives”) and the programme was rolled out over four years in four waves of implementation (see Appendix A).

The programme drew on a model of implementation based on one devised by Fixsen and colleagues (2005, 2009, 2011). The premise was to train a selection of practitioners, supervisors and managers, alongside

providing additional resources for infrastructure, and building regional and national collaborations to support best practice. In this way the aim was to maximise limited resources and to make the programme sustainable. In each wave of implementation a new group of practitioners was trained in evidence-based approaches and service managers in leadership and service development. Over the course of the four years just over 1,000 practitioners and managers were trained. During each wave, new partnerships collected data primarily from practitioner trainees for the first year of implementation and then, in the second year of implementation, collected data from all practitioners in the partnership (see Appendix A).

A central part of the CYP IAPT service transformation initiative was an emphasis on collecting child and parent outcome and experience data to help inform decision-making in clinical practice. Practitioners were encouraged to use “session-by-session” monitoring approaches, which would generally involve using one measure or briefer scales from longer measures that related to particular types of problem (Law & Wolpert, 2014). The intention was that service-user outcomes and feedback were considered frequently using scales relevant to their particular situation (see Appendix B).

Miranda Wolpert (director of CORC 1.5 days a week) was appointed the national informatics lead for CYP IAPT within NHS England (1 day a week). In this role she chaired the outcomes and evaluation group (OEG) which selected outcome and experience measures for use in services taking part in CYP IAPT and the development of the CYP IAPT dataset and reporting processes. (See Appendix B for membership and processes).

The OEG agreed the following principles for choosing outcome measures and embedding patient feedback and routine monitoring: clinically useful; able to cover a wide range of contexts and difficulties to allow shared decision-making between service users and practitioners in choosing appropriate measures for individuals; low burden; free (at least in paper form); chosen by stakeholder consensus; iterative and involving public consultations. A central principle adopted by the OEG was that scale selection should prioritise potential clinical utility over use for service review or evaluation. This meant that many scales were included that were not intended for summative use. For example, some feedback scales about experience in treatment were specifically intended to empower children and parents to honestly share any concerns or issues to allow the practitioner to address these during a course of treatment (see list of tools in Appendix B). The training of practitioners to encourage


and respond to negative feedback is also seen as a core component of good practice (Sapyta, Riemer & Bickman, 2005).

The OEG decided after much debate not to mandate any one scale for use in all cases. This was partly because of the wish to allow practitioners to choose collaboratively with service users the right scale for their particular circumstances, but also because any scale identified as potentially possible to use in this way (e.g. Weisz et al.’s “Top Problems scale”, 2011) required payment for use and was thus excluded from consideration.

All suggested measures were reviewed by the OEG and stakeholders were consulted by both targeted and public consultation. Feedback was collated by CORC and brought to the OEG for consideration. The results of detailed national consultations from 2012 and 2014 can be found on the CORC website: www.corc.uk.net. Through this process several measures were added to the initial data specification following suggestions from sites (see details in Appendix B).

Service context

The CYP IAPT programme was implemented in historically underfunded and fragmented services and implementation coincided with a recession with up to 75% cuts in some areas (YoungMinds, 2014). Data from the national benchmarking that took place during the programme indicated an increase in demand for services of at least 11% per annum (NHS Benchmarking Network, 2013). A Health Select Committee investigation also heard from providers of increased complexity and severity of cases (Hindley, 2014), all of which provided a challenging context for implementation.



Chapter 2: Measures

This chapter describes the tools that were used to measure outcomes for this analysis. It sets out the different scales that were used and the sample size for children and parents, before describing the Current View tool, which provided key contextual information.

Parent and child report measures

21 child report scales and 15 parent report scales are drawn on for one or more analyses in this report. Table 1 outlines all the child and parent report measure drawn on in this report including which scales within measures were and were not included (Appendix B outlines the wider list of measures used clinically). The table also indicates which scales had thresholds – which were used in some of the key analyses undertaken (see Chapter 6). 19 child report scales and 13 parent report scales had thresholds. It is important to note that different scales use different approaches to calculate thresholds. Scales with thresholds based on sensitivity and specificity analysis compared with clinical diagnoses provided by practitioners include GAD7 (Spitzer, Kroenke, Williams & Lowe, 2006) and PHQ9 (Kroenke, Spitzer & Williams, 2001). Scales with statistically derived thresholds based on highest scores within a general community population include SDQ measure and subscales (the top 10% of scores based on international samples) (Goodman, 2001). Scales with thresholds based on how far away the score is from the mean, based on the norm of the sample, include the RCADS measure and subscales (assuming scores are normally distributed this should equate to top 6% of scores based on a sample in Hawaii replicated in Australia and Denmark) (Chorpita, Yim, Moffitt, Umemoto & Francis, 2000). The fact that different scales use different thresholds presents challenges for comparing across scales and populations. Appendix D provides details of the thresholds and reliable change indices for the normed scales analysed in this report. All measures can be found on the CORC website (www.corc.uk.net). Note practitioner-rated measures (CGAS and HoNOSCA) are not included but will be considered in the future.

Measure of case characteristics

The practitioner-completed Current View tool was used to capture key characteristics of children accessing services (see Appendix E). The Current View considers 30 different types of presenting problem, 14 complexity factors, four contextual factors and educational, employment and training attendance and attainment (Jones et al., 2013) (see Appendix E).

The tool is completed by practitioners as close to their first meeting with a child or family as possible and updated

as their understanding of the issues change over time. It was developed and used as part of the payment systems development work for CYP mental health services (Wolpert et al., 2015b). This work was undertaken concurrently with the CYP IAPT programme of system transformation.

Data from the full sample of Current View were available for 42,798 cases. Data to inform case-mix algorithm (which included the stricter criteria of having been completed within 56 days of start date) consist of 31,038 cases.

Case-mix algorithm

The Current View tool was used as part of the payment systems work to develop an algorithm that could provisionally assign children and young people to “needs-based groupings”. The algorithm drew on the practitioner ratings of the 30 presenting problems listed in the tool together with the rating of the complexity factor, “pervasive developmental disorders”, and the age of the child (see Wolpert et al., 2014a; Wolpert et al., 2015b and Appendix E).

The algorithm assigns children to one of 20 potential needs-based groupings, some of which relate directly to NICE-based treatment guidelines. For the algorithm to suggest a NICE-informed grouping, the child must have the relevant index problem rated as at least moderate. The index problem is the main symptom associated with a particular NICE clinical guideline and the child must NOT have high ratings on a selection of other presenting problems (“exclusion criteria” are different for each grouping based on clinical judgement regarding which kinds of symptoms may mean that the group indicated by the index problem may not be appropriate). Details of the algorithm are found in Appendix E and full details in Wolpert et al. (2015b).

Since the development of this algorithm was based on an assumption that the assessment of the patients’ presentation of problems takes place within 56 days of the recorded start of their therapy, this stricter criteria was applied to the sample to consider needs-based grouping allocations and resulted in a sample of 31,038 cases.

Details on the Current View tool including its development, reliability and validity, along with full details of the algorithm can be found in Wolpert et al., 2015b (a brief summary can also be found in Appendix E).

Table 1: Child and parent scales used in this report

Domain/ Focus	Child scale	Parent scale	Includes thresholds?	Included in one or more analysis in this report?	Largest sample size of completed scales drawn on for one or more analysis*	
					Child	Parent
Depression	RCADS depression subscale	RCADS depression subscale	Yes	Yes	4,016	1,817
	PHQ9		Yes	Yes	438	–
Obsessive compulsive disorder (OCD)	RCADS OCD subscale	RCADS OCD subscale	Yes	Yes	3,461	1,761
Generalised anxiety	RCADS generalised anxiety subscale	RCADS generalised anxiety subscale	Yes	Yes	3,451	1,784
Social phobia	RCADS generalised anxiety subscale	RCADS generalised anxiety subscale	Yes	Yes	3,543	1,771
Panic	RCADS panic subscale	RCADS panic subscale	Yes	Yes	3,549	1,718
Separation Anxiety	RCADS separation anxiety subscale	RCADS separation anxiety subscale	Yes	Yes	3,450	1,808
Trauma	Impact of events scale (IES)		Yes	Yes	183	–
ADHD	SDQ subscale on hyperactivity	SDQ subscale on hyperactivity	Yes	Yes	2,446	2,556
Peer problems	SDQ subscale on peer problems	SDQ subscale on peer problems	Yes	No, as not clear that it relates to a treatable mental health issue	–	–
Prosocial behaviour	SDQ subscale on prosocial problems	SDQ subscale on prosocial problems	Yes	No, as not clear that it relates to a treatable mental health issue	–	–


Table 1: Child and parent scales used in this report (continued)

				Included in one or more analysis in this report?	Largest sample size of completed scales drawn on for one or more analysis*	
					Child	Parent
Behaviour difficulties	Me and My School (renamed Me and My School feelings) behavioural subscale		Yes	Yes	67	–
	SDQ subscale on conduct disorder	SDQ subscale on conduct disorder	Yes	Yes	2,459	2,591
Impact on functioning	Routine monitoring questionnaire (SxS)	Routine monitoring questionnaire (SxS)	Yes	Yes	491	309
	SDQ subscale on total impact	SDQ subscale on total impact	Yes	Yes	2,006	2,106
Overall anxiety	RCADS anxiety subscales combined	RCADS anxiety subscales combined	Yes	No, as analysed subscales so overall scale not used to prevent double counting	–	–
	GAD-7			Yes	194	–
Overall anxiety and/or depression	RCADS full scale	RCADS full scale	Yes	No, as analysed subscales so overall scale not used to prevent double counting	–	–
Overall emotional problems	SDQ emotional subscale	SDQ emotional subscale	Yes	Yes	2,456	2,579
	CORE-10		Yes	Yes	53	–
	YP CORE		Yes	Yes	69	–
Overall psychological problems	SDQ total difficulties	SDQ total difficulties	Yes	No, as analysed subscales so overall scale not used to prevent double counting	–	–

Table 1: Child and parent scales used in this report (continued)

Domain/ Focus	Child scale	Parent scale	Includes thresholds?	Included in one or more analysis in this report?	Largest sample size of completed scales drawn on for one or more analysis*	
					Child	Parent
General wellbeing	ORS	ORS	Yes	Yes	532	118
	CORS		Yes	Yes	446	–
	Short/Warwick Edinburgh Mental Wellbeing Scale		No	No, as no cut- offs available	–	–
Eating disorders	EDE-Q/A		No	No, as no cut- offs available	–	–
Family functioning	SCORE-15	SCORE-15	No	No, as no cut- offs available	–	–
Oppositional defiant disorder		ODDp	Yes	Yes	–	139
Parental self-efficacy		BPSES	No	No, as no cut- offs available	–	–
Learning disability		SLDOM	No	No, as no cut- offs available	–	–
Achievement of goals	GBO	GBO	No	Yes	2,784	686
Service satisfaction	ESQ	ESQ	No	Yes	3,196	2,698

* Analyses done on paired outcomes for above treatment threshold sample in all cases except GBO and ESQ, where analyses were done on closed treatment cases sample



Chapter 3: Method

Chapter 3 sets out the approach to analysis taken in this report. It describes how the data were captured, the challenges of collecting the data and a range of explanations for the difficulties encountered.

Data capture

Patient-level data from staff across all of the participating partnerships were collated and submitted quarterly to the central team based at CORC according to the CYP IAPT data specification (CYP IAPT programme, 2013) via secure data handling and storage by a secure data storage provider. In the first year of involvement in the initiative, data were

largely sent from those involved directly in the training, but from the second year of involvement onwards data were sent from all practitioners across the partnership. In line with the CYP IAPT protocol, staff routinely collect demographic, outcome, and experience measures completed by the therapist, young person, and/or carer at assessment (Time 1 or T1), on a session-by-session basis, and at a review point 4-8 months later or, if sooner, case closure (Time 2 or T2) (Law & Wolpert, 2014). The data comprised of “periods of contact” for each child seen, which consists of a number of “events” (see Figure 1).

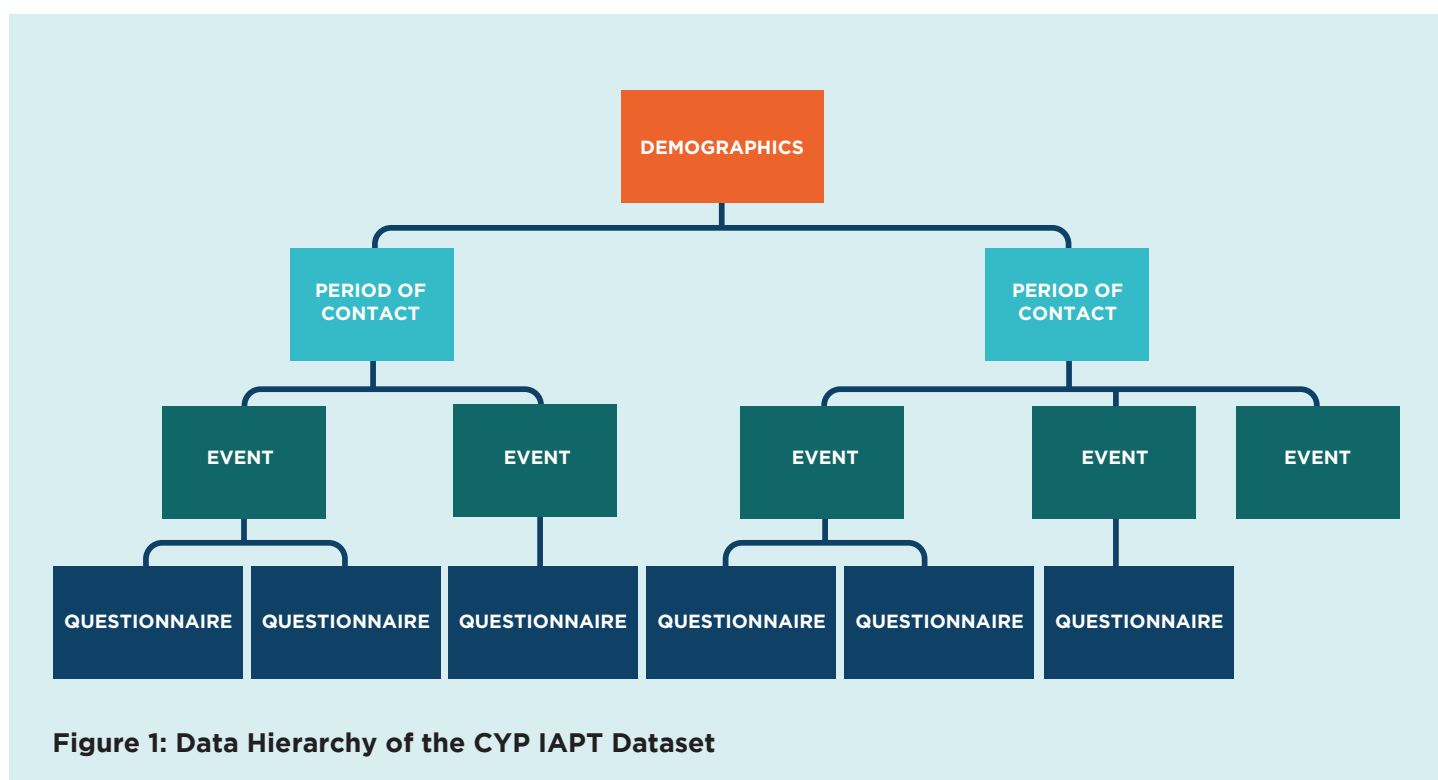


Figure 1: Data Hierarchy of the CYP IAPT Dataset

Duplicates in the data

A significant issue that arose was that duplicate periods of contact (POC) were identified in the dataset. These took the form of identical information in every field of a given measure, with the exception of the recorded POC start date. It therefore appeared that the same individual completed the same measure, at the same time, on the same date, and at the same service, and gave identical responses to every single question. The extent of duplicated POCs ranged from 1% to 5% across the different outcome measures. It should be noted that duplication issues such as these are not unique to this dataset and other international datasets have encountered similar issues (McCoy et al., 2013). Investigations into this issue identified that some of the duplicate POCs within the

dataset were the result of the original data specification. However, other duplicates were genuine examples of where people were seen for multiple periods of contact. The issue was dealt with by applying criteria at the level of the individual to determine which recorded POCs could reasonably be judged as being legitimate and excluding the others. Checks on the analyses undertaken for the report indicate that the presence of duplicates did not have a substantial impact on the findings.

Data completeness and missing data

At the outset a target was set for data completion of 90% of all closed treatment cases (defined as at least three

recorded events, not just for assessment and the case recorded as closed) having paired parent- or child-reported outcome data on at least one scale with thresholds, along with recorded data on any issues with education, employment or training (EET). Scales were considered to be paired if they had been recorded as taking place on different days, and from the same perspective. Only 22% of closed treatment cases had both EET information at least once, and at least one set of paired scales with thresholds from the same perspective.

It is important to acknowledge the impact of the high degree of missing data. Given the unknown (unmeasured) bias due to missing values and the unknown selection mechanisms into different outcome measures, these data are not necessarily applicable to the wider child mental health service population and it is not possible to hypothesise with any certainty the likely direction of a skew. It could be that children and parents who are likely to complete scales are likely to be those most positively disposed towards services (e.g. they may be the most likely population for practitioners to remember to ask to complete a scale and may be the most likely to comply with completing). On the other hand, it may be that those who quickly improved stopped coming and thus were the most likely not to complete a Time 2 measure. The issue of appropriate use of a dataset with such a high degree of missing data has already been discussed extensively and CORC's potentially controversial position laid out. It is fully anticipated that esteemed colleagues may take a different view on these contentious issues.

In terms of addressing missing data, a variety of approaches are advised in the literature, including the last measure being carried forward and imputation which would not be appropriate in this instance. Every effort has been made to be transparent throughout about the large degree of missing data in all samples, which introduces unknown biases and makes it vital that all findings are treated with caution. For all findings in this report, the percentage of data the analysis is based on from a specified sample is highlighted and any known differences in sample characteristics noted.

Possible reasons for missing data

Below are possible factors influencing the high degree of missing data with some ideas of how these might be addressed in the future.

Infrastructure challenges

Most services involved in CYP IAPT were small parts of larger organisations (e.g. child mental health teams within larger mental health trusts). This meant they were

frequently using patient information systems adapted from the needs of the larger trust and not necessarily responsive to the requests for adaptation for child mental health purposes. The absence of an existing nationally mandated collection and collation protocol meant there was no one agreed dataset already in place and there was no one agreed system common across sites, nor the possibility to mandate such a system given the governance structures within the NHS and the relationship between NHS England and services.

Inter-agency working

The fact that partnerships were collaborations between different sectors, including voluntary sector and NHS provision, was a key part of the programme and will be important for service provision in the future. However, coordinating and centralising the flow of data from different agencies presented many challenges, including complexities of governance arrangements between providers, as well as IT inter-operability or the lack of it. This is likely to be an essential issue to address in the future if the ambition of Future in Mind is to be achieved.

Burden of collection

Given the issues with infrastructure and IT systems, for the majority of services manual data entry, and frequently dual data entry, was required (the same information having to be put on two different systems by hand), leading to an increased burden on services and problems with both the capture and transfer of data. The burden of data entry, especially in already pressurised services, is likely to have been a major barrier.

Practitioner lack of confidence

Use of routine outcome measures is still a new practice in child mental health services. As has been noted, the international literature has identified practitioner concerns and how best to implement them (e.g. Wolpert, 2014a; Moran, Kelesidi, Guglani, Davidson & Ford, 2012; Batty et al., 2013). The CYP IAPT programme trained a cohort of practitioners in the use of routine outcome measurement and there was evidence of this impacting on those involved in terms of their sense of self efficacy (Edbrooke-Childs, Wolpert & Deighton, 2014). However only a small number of practitioners were trained relative to the overall workforce within services from which data were derived. Whilst there was anecdotal evidence of a wider range of practitioners reporting finding the measures helpful in informing practice, it is not clear how far the full range of practitioners had started to use this approach with the children and families they were working with.

Inadequate measures available

It is widely acknowledged that the measures of outcome and experience in child mental health are all flawed in one way or another. Many of the measures used were developed as epidemiological tools and, as has been outlined, there is variation in what thresholds are used for different measures and for some key measures (such as CORS and RCADS) a lack of UK norms (Law & Wolpert, 2014). All this may contribute to appropriate practitioner concern about how best to use measures clinically. The suggestion from the OEG was to use goals-based outcomes where practitioners were not sure of the appropriateness of other indicators but this may have been more of a challenge for some to build into their ways of working than for others. A decision made by the OEG in light of NHS England policy to consider only free measures may have excluded key measures that practitioners would have found more clinically robust, such as the Beck Inventories. Moreover, the lack of scales coproduced with children may have reduced acceptability or relevance of scales. Having noted this, however, it is worth also noting that 79% (18,411/23,373) of cases had service feedback information at case closure, indicating clinicians were finding these measures of clinical utility (which were not part of the data completion target) easier to administer. Additionally, there was high engagement in the consultations on scales and measures and particular calls by practitioners to use the CORS, ORS and CORE measures (all of which were incorporated in the dataset) which many reported finding clinically useful. Many also reported finding RCADS and other measures useful.


Too many measures

It is possible that deciding to allow for the use of a wide variety of scales in clinically diverse populations, and building on the principle of collaborative working and choice between providers and users of services, may have reduced the likelihood of collecting enough data on any one scale. Diversion of energies in considering and collecting multiple measures may have limited concentrated effort on the collection of a single scale that could be used to consider outcomes across populations. A focus on a goals-based metric may be a step forward in this regard though this approach also has limitations in terms of its robustness and rigour.

Unrealistic expectations about outcomes

As has been stated elsewhere, it can be argued that the lack of data in child mental health and the emphasis on a need for greater access have stifled debate about the appropriate level of improvement that can be expected for

those accessing services. As has been written elsewhere, the largely unchallenged shared belief that everyone can potentially be helped is a powerful disincentive to genuine learning and service development. Where is the incentive for therapists to rigorously examine who they do or don't help if any data that show less than 100% success is going to be worse than current internal and external expectations? Moreover, the assumption conveyed to service users that treatment failure rarely occurs may lead to unhelpfully raised expectations, deprives therapists of a language to honestly talk about the limits of help for anyone, and may leave those accessing services who find they are not "better" blaming themselves (Wolpert, 2016).



Chapter 4:
**Samples
of children**

In total, seven nested samples were derived in successive steps from the submitted data (see Figure 2 on the following page). This chapter describes each of the samples in detail and how they related to one another. The characteristics of key samples are described in Appendix C.

The seven samples

1. Full sample which comprised 96,325 records of care. These data were provided by 81 (out of 82) partnerships. In the CYP IAPT dataset, these records represent 91,503 young people because each separate episode of care creates a different case record. However, all the analyses in this report were focused on the period of contact level. This sample was used to examine key characteristics about the children and families that were seen by sites taking part in the programme. It is not known how large a percentage of the actual number of children seen these data represent. There were concerns that, given the multiple systems, data might have gone missing in this process and it was not possible to identify how much, if any, data were lost in this way. Children and parents were asked to give consent for questionnaire returns, together with their identifiable data, to be held by a third party in a secure data warehouse to allow for data linkage across time and across partnerships. Feedback from service leads indicated that only a minority of children and families did not agree to have their data held in this way, but the exact number of non-consenting cases is not known, nor is the number of cases that were not recorded on the system for others reasons. These factors may have led to skews in the data.

2. Ended referrals which comprised 50,465 records of care from 76 partnerships. This sample includes closed cases and those with a period of contact end date (indicating that treatment had finished even if the case had not been officially closed). This sample was used to look at the length of treatment for closed cases.

3. Closed treatment cases which included cases where the young person was seen for a course of treatment (of at least three events, excluding assessment only). This sample comprised 23,373 records of care from 75 partnerships and was used to examine goals and feedback/satisfaction information.

4. Measured closed treatment cases which included closed cases where the young person was seen for a course of treatment (of at least three events, excluding assessment only) and had completed a scale at at least one time point. This sample comprised 17,055, which had a child- or parent-reported measure, from 75 partnerships (15,490 cases from the child-reported perspective and 11,909 cases from the parent-reported perspective).

5. Above threshold closed treatment sample

which included closed cases where the young person was seen for a course of treatment, and where they had at least one scale where the score was above the threshold set within that scale at the start of treatment. This sample comprised 15,536 with a child- and/or parent-reported measure, from 75 partnerships (12,128 cases from the child-reported perspective and 10,438 cases from the parent-reported perspective).

6. Paired outcomes for above threshold closed treatment sample

which included closed cases where the young person was seen for a course of treatment and where they completed the same scale that was above the threshold at outset, at a second time point. Scales were considered to be paired if they had been recorded as taking place on different days but by the same respondent. This sample comprised 7,808 with a child- and/or parent-reported scale, from 74 partnerships (5,918 cases from the child-reported perspective and 3,699 cases from the parent-reported perspective). For the analysis, only scales that captured issues that can be specifically targeted in child and young people mental health services, and that have thresholds and reliability information, were included. This means that some were excluded from key analyses; namely, the SDQ peer and prosocial subscales (although potentially relevant to assessing difficulties, it was felt that these behaviours might not be a particular focus of clinical change) and Goal Based Outcomes (which were analysed separately), BPSES, ODDp, S/WEMWBS, SLDOM and SCORE-15 (since they did not have thresholds and/or reliability information). Others were excluded due to a lack of Time 2 data – for example, EDE-Q/A and Kessler-10. Effect sizes, “recovery” and reliable change were derived from this sample.

7. Added value score sample which included closed cases where the young person was seen for a course of treatment and where they had paired parent SDQ emotional, total difficulties and total impact scores above threshold at a first time point. This sample comprised 1,010 from 59 partnerships. The added value score was calculated for this sample.

Sample representativeness

Due to the significant extent of missing data across the dataset, the findings in this report need to be treated with caution. In the chapters that follow we consider representativeness of data samples wherever possible. However, it is important to note that there are likely to be additional systematic differences between the children for whom there are data and those for whom data are missing, related to characteristics that were not measured, such as levels of motivation.

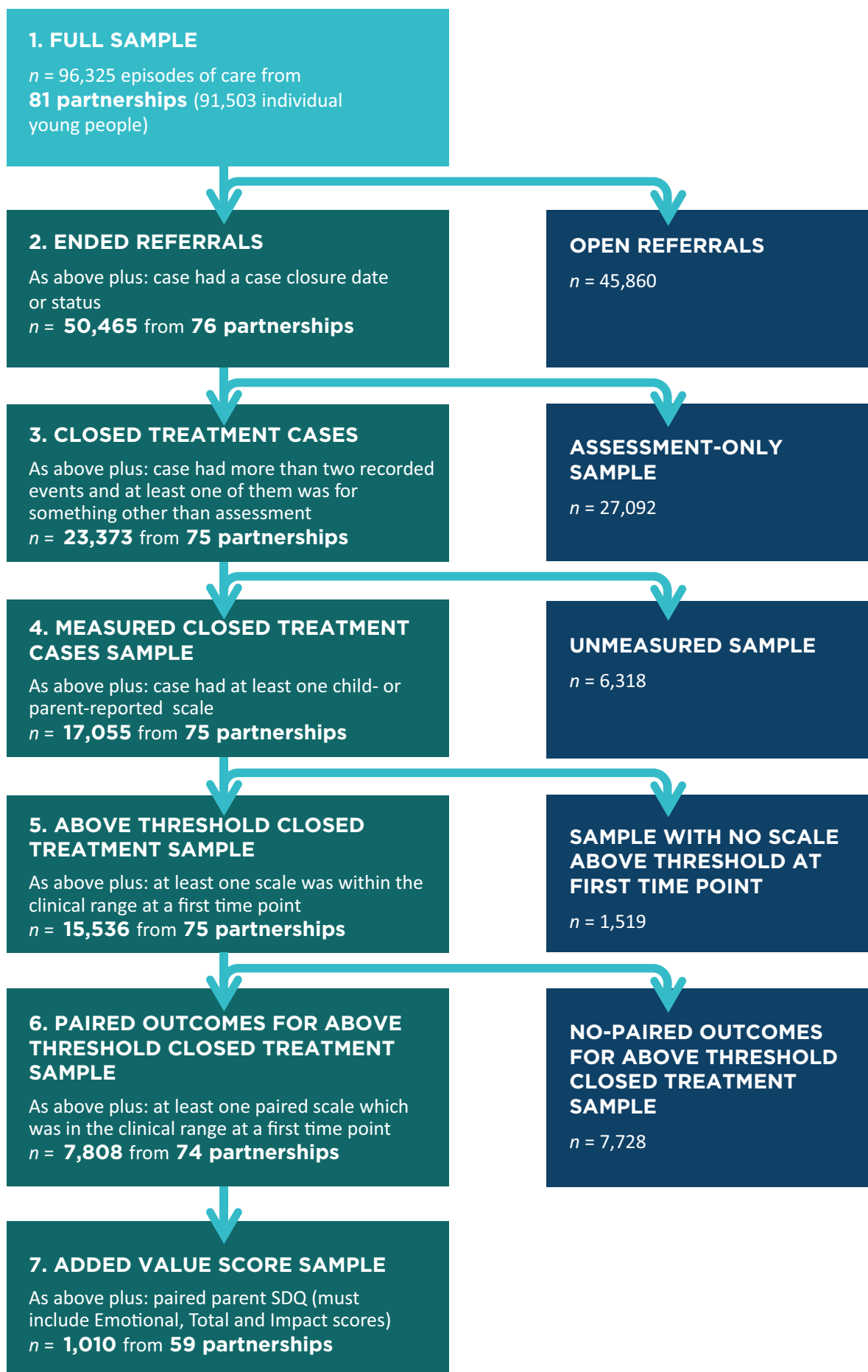


Figure 2: Diagram of samples used in the analyses



Chapter 5:

**Demographics
and service
provision data**

This chapter describes young people’s demographic and case characteristics and types of treatment they received. All analyses in this chapter relate to the full sample (see Figure 2).

Number of cases and characteristics of young people seen

The number of cases for which we have data arising from this period was 96,325. Service and data leads indicated that the number of parents and children not giving consent for their data to be held was small. However, 96,325 is clearly a subset (of unknown size) of all the

children seen by these partnerships over the relevant periods as the partnerships joined in waves, and data from each partnership were only collected for the small number of actual trainees in year one of their wave of joining. In addition, as discussed above, there were challenges with collecting and uploading data.

In terms of the demographic and referral source characteristics of young people within this sample, the average age was 12 years, there were slightly more females (52%) than males (48%) and the most common ethnicity was white (82%). The mean age was 11.98 years (SD = 3.78).

Table 2: Age and gender distribution in the whole sample

	0–4 years count (%)	5–9 years count (%)	10–14 years count (%)	15–19 years count (%)	20–25 years count (%)	Total Gender count (%)
Male	2,111 (65.7%)	15,602 (69.8%)	18,560 (46%)	9,736 (32.4%)	56 (36.1%)	46,065 (47.9%)
Female	1,102 (34.2%)	6,759 (30.2%)	21,739 (53.9%)	20,283 (67.5%)	99 (63.9%)	49,982 (52%)
Indeterminate	0 –	3 (<0.1%)	9 (<0.1%)	16 (0.1%)	0 –	28 (<0.1%)
Total Count (%)	3,213 (3.3%)	22,364 (23.3%)	40,308 (42%)	30,035 (31.3%)	155 (0.2%)	96,075 (100%)

Notes: n = 163 incorrectly recorded age (either negative or higher than 25 years; 0.2% of the sample); n = 14 missing age (<0.1% of the sample); Notes: n = 73 missing gender (0.1% of the sample)

Table 3: Ethnicity distribution in the whole sample

White count (%)	Black count (%)	Asian count (%)	Mixed count (%)	Other count (%)
56,370	3,569	3,593	3,410	1,721
(82.1%)	(5.2%)	(5.2%)	(5.0%)	(2.5%)

Notes: n = 17,970 not stated (18.6% of the sample); n = 9,692 missing (10% of the sample); percentages are out of cases with completed information.

In comparison to the pre-CYP IAPT dataset held by CORC (2007–10), the present sample is broadly similar in terms of age and gender, although slightly more female and slightly older. The pre-CYP IAPT dataset had a sample of 27,065 closed episodes of care seen for at least three events across 34 children and young people’s mental health services, and there were data on age and gender (representing 94% data completion from all closed

treatment cases). In the pre-CYP IAPT sample, the mean age was 11.38 years (SD = 3.81) and 56.3% were male. In terms of ethnicity, the current sample is more white than the pre-CYP IAPT sample in which the 23,551 cases with information about ethnic categorisation (representing 81% data completion of all closed treatment cases) indicated 79% of cases recorded as white and the rest as coming from an ethnic minority.

Referral source

Children were most often referred from primary care (51%). This is in line with the earlier CYP IAPT dataset which found the majority of cases (55%) for whom this data was available ($n=24,993$; 86% of the sample) was referred from primary healthcare, with 13% from education, 13% from child health, and the remaining 19% from other sources, including self-referral (1.6%).

Out of all the cases in the dataset that had information on referral source ($n=30,030$), 4% (95% CI 3.88% – 4.12%) were recorded as being self-referrals. This is a larger proportion than the 1.6% of children who were recorded as being self-referred in the pre-CYP IAPT dataset held by CORC (2007–2010). Children were most often referred from primary care (51%).

Table 4: Sources of referral in the whole sample

Referral source	N	%
Primary health	33,811	51
Local authority	9,968	15
Other	6,203	9.4
Child health	5,299	8
Internal referrals from community	4,141	6.2
Acute secondary care	2,934	4.4
Self-referral	2,709	4.1
Internal referrals from inpatient	546	0.8
Independent/voluntary sector	220	0.3
Justice system	165	0.2
Transfer by graduation	162	0.2
Other mental health NHS trust	134	0.2
Employer	3	<0.1
Total	66,295	100

Notes: $n = 30,030$ missing (31.2% of the sample); percentages are out of cases with completed information

Types of difficulties seen

Data from the full sample of Current View were available for 42,798 cases. Of these, the most common presenting problem was family relationship difficulties (52%), followed by depression/low mood (50%), and/or generalised anxiety disorder (49%) (not mutually exclusive issues). The most prevalent complexity and contextual factors were parental health issues (19%), experience of abuse (13%), and home (59%) and school issues (57%).

In addition, 31% of cases indicated that they had mild, moderate or severe attendance difficulties with education, employment and training, and 42% of cases indicated that they had attainment difficulties in these contexts. Figure 3 to Figure 5 below show the types of presenting problems, complexity and contextual factors (as captured by practitioner report on the Current View) at treatment outset for all cases seen.

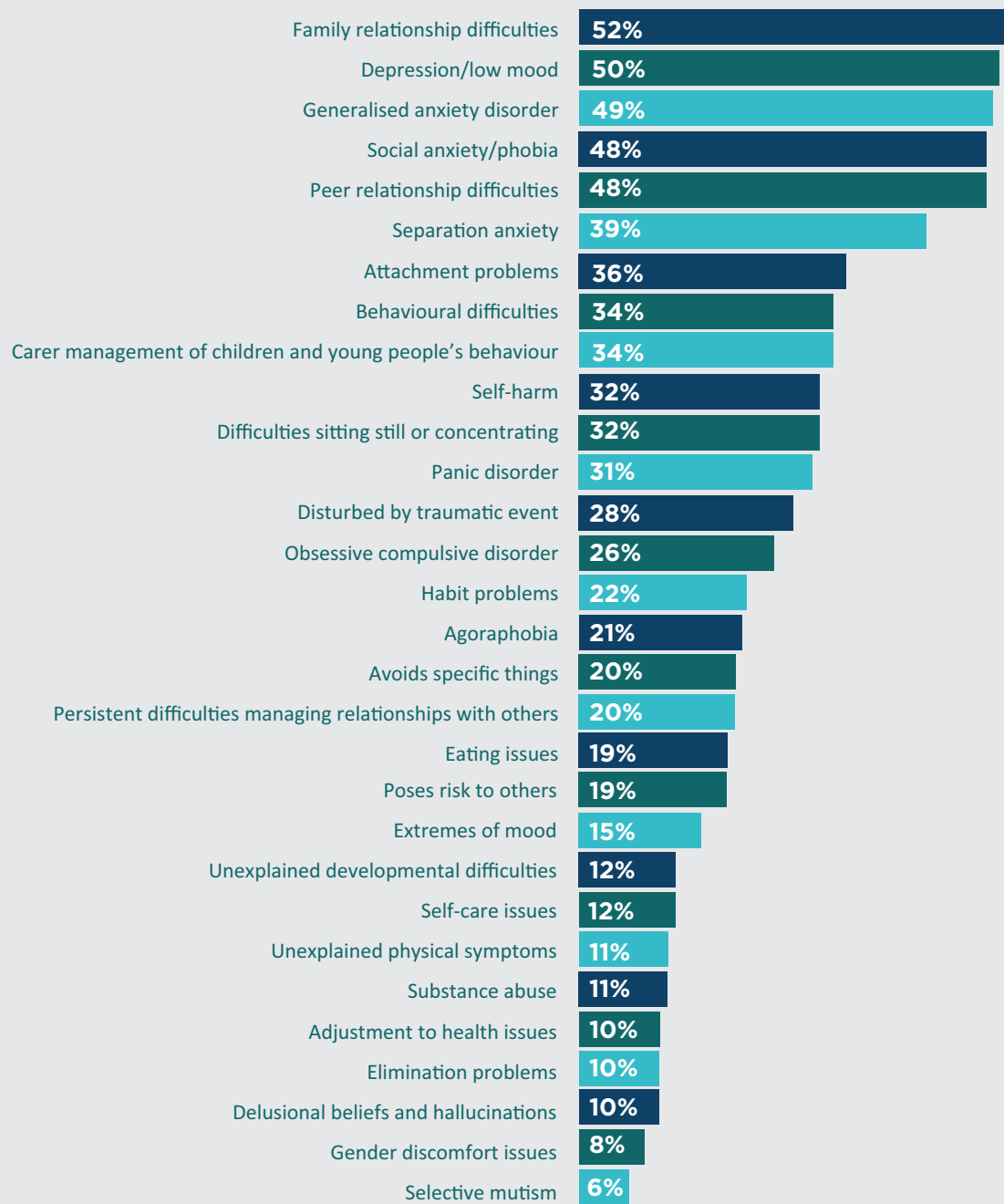


Figure 3: Percentage of cases with a provisional problem descriptor endorsed as mild or above in the first recorded Current View.

Notes: $n = 42,798$ (44% of the sample); percentages are out of those with a completed Current View form; categories are not mutually exclusive.

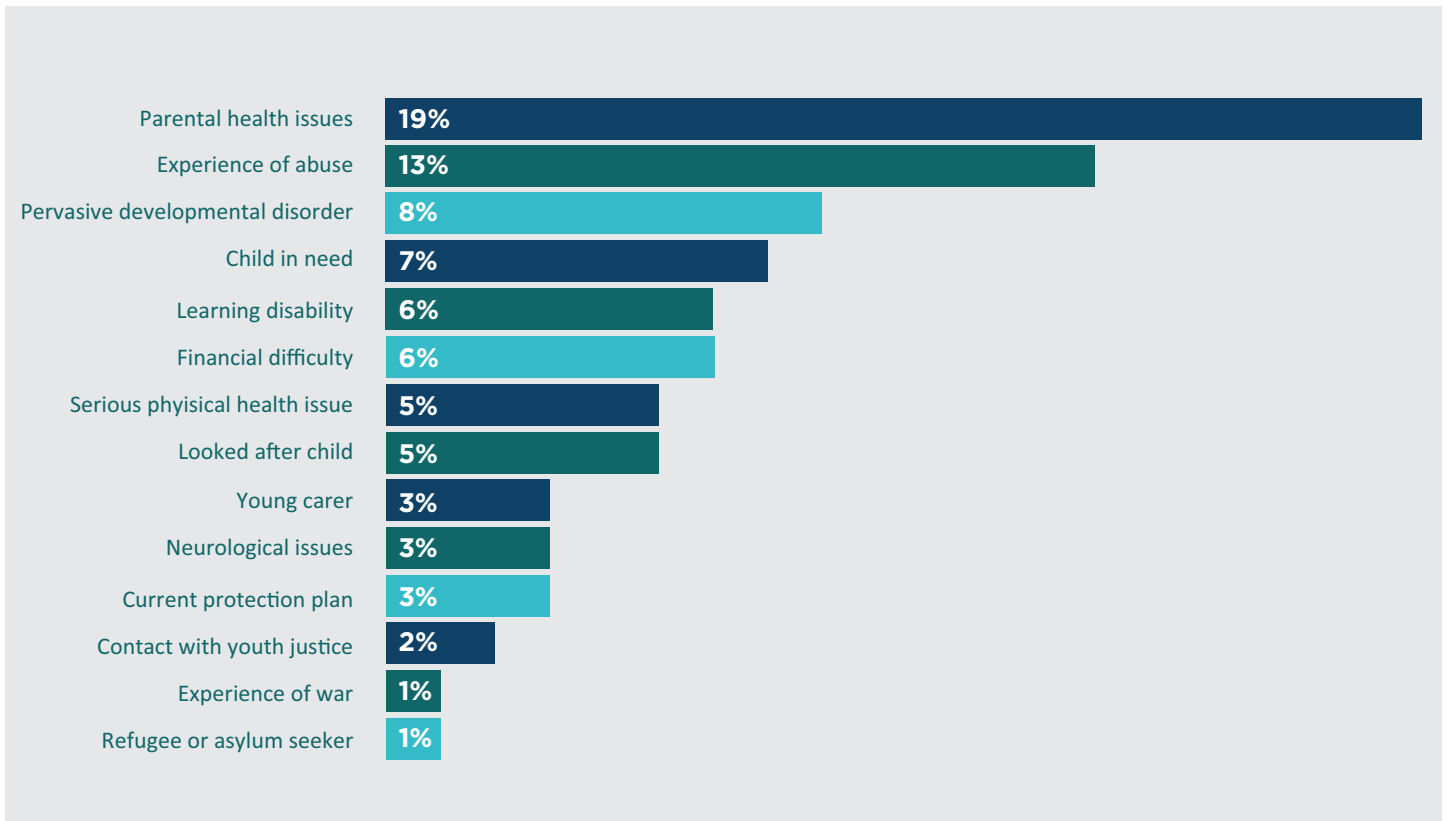


Figure 4: Percentage of cases with a complexity factor endorsed in the first recorded Current View.

Notes: $n = 42,798$ (44% of the sample); percentages are out of those with a completed Current View form; categories are not mutually exclusive.

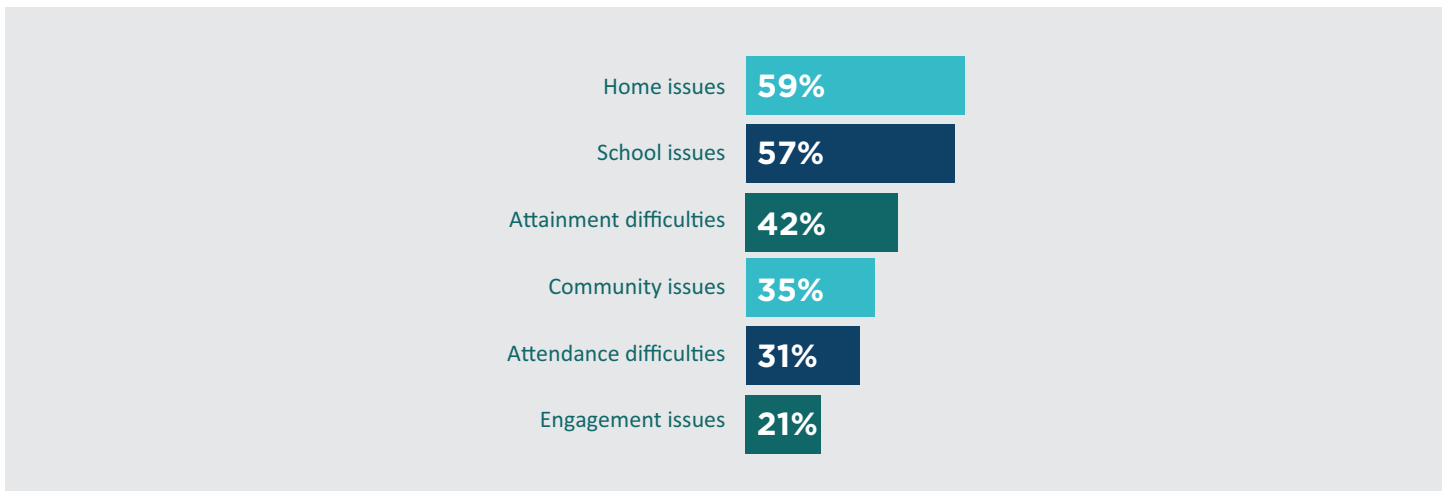


Figure 5: Percentage of cases with a contextual or attainment factor endorsed as mild or above in the first recorded Current View

Notes: $n = 42,798$ (44% of the sample); percentages are out of those with a completed Current View form; categories are not mutually exclusive.

No comparable data are available from the pre-CYP IAPT dataset as the Current View tool was not used.

Case mix (needs-based groupings)

The algorithm used to provisionally assign children to “needs-based groupings” (see Appendix E) was based on an assumption that the assessment of the patients’ presentation of problems takes place within 56 days of the recorded start of their therapy. The sample presented below is therefore based on stricter criteria, and consists of 31,038 cases.

Based on this algorithm, 37% of cases ($n=11,484$) fell into one of the needs-based categories where it was likely that a particular NICE-based guideline may be able to inform

treatment and a further 12% of cases ($n=3,724$) fell into the category of having co-occurring difficulties where more than one NICE guideline may be relevant (see Table 5 below). For the algorithm to suggest a NICE-informed grouping, the child must have the relevant “index problem” rated at least moderate. The index problem is the main symptom associated with a particular NICE clinical guideline and the child must NOT have high ratings on a selection of other presenting problems (“exclusion criteria” are different for each grouping based on clinical judgement regarding which kinds of symptoms may mean that the group indicated by the index problem may not be appropriate). Full details of the algorithm are found in Appendix E.

Table 5: Potential allocation to support guided by NICE guidelines

Index difficulties as indicated on current view	Percentage in CYP IAPT dataset	Relevant NICE guideline (at time of development of algorithm 2014)
Difficulties sitting still or concentrating (ADHD)	6%	ADHD NICE guideline 72
Pervasive developmental disorder (autism)	2%	Autism spectrum NICE guideline 170
Behavioural difficulties (conduct disorder or oppositional defiant disorder)	4%	Anti-social behaviour and conduct disorders NICE guideline 158
Extremes of mood (bipolar disorder)	1%	Bipolar disorder NICE guideline 185
Depression/low mood (depression)	5%	Depression NICE guideline 28
Anxious generally (generalised anxiety disorder, GAD) and/or panics (panic disorder)	5%	GAD and/or panic disorder NICE guideline 113
Compelled to do or think things (obsessive compulsive disorder, OCD)	1%	OCD NICE guideline 31
Disturbed by traumatic event (post traumatic stress disorder)	2%	PTSD NICE guideline 26
Self-harm (self-injury or self-harm)	6%	Self-harm NICE guidelines 16 and/or 133
Anxious in social situations (social anxiety or phobia)	2%	Social anxiety disorder by NICE guideline 159
Eating difficulties (anorexia, bulimia)	2%	Eating disorders NICE Guideline 9
Delusional beliefs and hallucinations (psychosis)	1%	Psychosis NICE guidelines 155 and/or 185
Co-occurring emotional problems	10%	One or more of NICE guidelines above
Co-occurring emotional and behavioural difficulties	2%	One or more of NICE guidelines above

The remaining 51% of cases ($n=15,829$) had difficulties not clearly relatable to any NICE guidance. This may be because they have difficulties not covered by NICE guidance, or a combination of difficulties that could not be captured in “co-occurring emotional” or “co-occurring emotional and behavioural” categories. For example, a young person who is experiencing family relationship difficulties, gender dysphoria or peer relationship difficulties, or any combination of these with, for instance, low mood, anxiety and substance abuse would fall within this category.

Treatment received

Over half of the total sample (65%, $n=62,611$) did not have a type of therapy recorded; excluding those cases from the analysis. Figure 6 shows that the most frequent type

of defined therapies were cognitive behavioural therapy (CBT; 17%, $n=41,247$ events) followed by multimodal integrated therapy (11%, $n=25,859$ events), family systemic therapy (8%, $n=19,442$ events) and parent training (7%, $n=17,173$ events). However, the most commonly recorded type overall was “other therapy” (19%, $n=46,438$ events). This could refer to any other type of work carried out that practitioners did not feel fitted into the categories available. One possible reason for the prevalence of the “other therapy” category was the absence of an assessment category. However, analysing the data just using events that were categorised as “during therapy” resulted in very similar findings, with the most commonly completed type of therapy also “other” therapy (17%), followed by CBT (14%) and multimodal integrated therapy (11%).

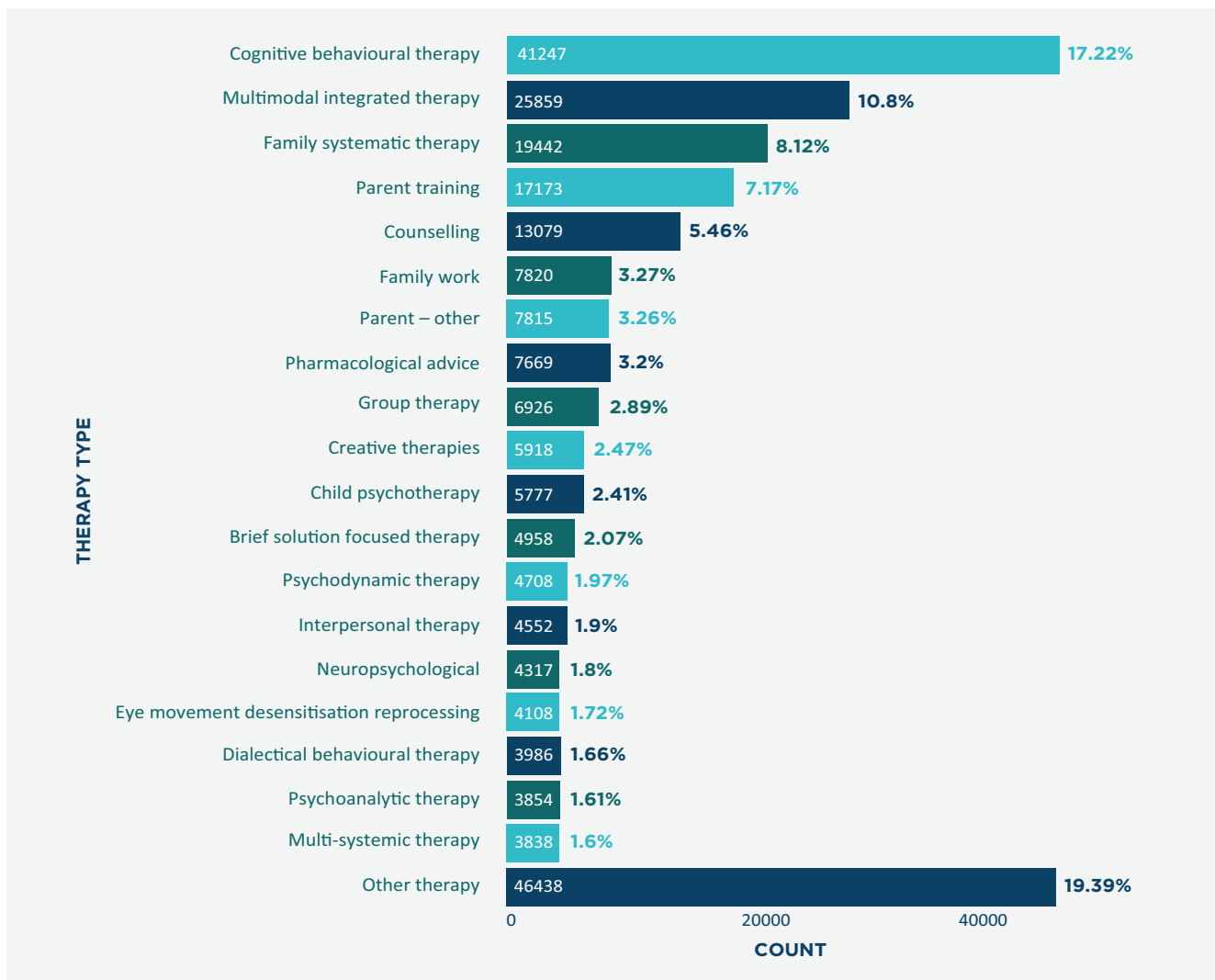


Figure 6: Therapy types for events attended for all cases

Notes: Categories are not mutually exclusive.

Length of treatment

Out of the 50,465 cases recorded as being closed, 452 were either recorded as being opened earlier than 1 January 2010 (two years before the beginning of CYP-IAPT), or were recorded as being closed later than 1 January 2016 (after data collection had ended), or were recorded as being closed before they were opened. Since the date records for these cases are clearly suspect, they are excluded from analyses about the length of contact and number of sessions. Of the remaining 50,013 cases, the mean length of contact was 195 days (SD = 207 days, median = 140 days), with a range of 0 to 1905 days (approximately 5 years, 2 months). 41,068 cases had at least 1 recorded event. The mean number of recorded events in this sample was 5.3 events (SD = 7.8 events), median = 3 events, range 1 – 268 events.



Chapter 6:

**Strengths
and limitations
of different
analytic
approaches**

As was alluded to in Chapter 1, key indices of change and experience have been identified to consider outcomes and experience in child mental health (Wolpert et al., 2015a; Brown, Ford, Deighton & Wolpert, 2014): raw experience of service data, raw change scores, “recovery”, reliable change, reliable “recovery” (a combination of recovery and reliable change), effect size calculation and added value score. Key strengths and limitations of each are considered below, followed by a brief summary of how they have been used in this report.

Raw service experience data

Considering child and parent service satisfaction and experience data may yield some insights into their experience of provision. There are many limitations to this approach and it cannot be taken as a proxy of outcomes as such. The first clear limitation is that many feedback and experience scales are specifically designed to empower children and parents to honestly share any concerns or issues to allow the practitioner to address these during a course of treatment (such as the Session Rating Scale and the Session Feedback Questionnaire – see Appendix B). The training of practitioners to encourage and respond to negative feedback is also seen as a core component of good practice (Lambert et al., 2001; Saptya et al., 2005). These feedback measures are not included in this analysis. Other scales are designed as summative evaluation at the end of treatment and thus more appropriate to use to review experience (e.g. CHI ESQ) and this is included in terms of raw score analysis in this report. However, using such data has limitations. It cannot be used simply as a proxy indicator of outcomes, since an individual may be appreciative of help without feeling things have improved significantly (Brown et al., 2014). Moreover, they may be particularly affected by the context, with people not wanting to be seen to report negative feedback if they know this will be shared in an un-anonymised form with their service provider.

Raw change scores

Raw change in scores is an appealingly simple metric to use and has been suggested for use particularly for those scales without thresholds. Limitations of using raw change in scores alone when considering use with psychometrically robust scales – where measurement error is known and where thresholds might exist – is that these psychometric attributes provide more robust indicators of meaningful change (such as “recovery” and reliable changes, as discussed below). For those scales without proven psychometric properties such as validity and reliability, it is hard to know how much change should be treated as important or how to interpret any changes.

This approach is used in this report to analyse movement towards goals as it was felt this measure was particularly important in capturing what was important to those accessing services. It therefore addressed, in various ways, some of the limitations of potential meaningfulness raised by other measures.

“Recovery”

“Recovery” considers movement across a threshold on scale/s and generally only considers those who at least have some problems rated above the threshold at the outset. One strength is that it is conceptually simple to explain and intuitive to understand. Disadvantages of “recovery” as a metric include the fact that it does not consider those with no problems above the threshold at outset. By only considering children who were above the threshold at the outset, there is a danger that the recovery measure creates findings which are an artefact of “regression to the mean”, because cases are selected with higher initial scores (where the likelihood is greater that they move down), and cases excluded with lower initial scores (where the likelihood is greater that they move up). Moreover, any change does not necessarily reflect the individuals’ lived experience nor take into account the magnitude of change experienced. A very small change in one person’s score could be counted towards the proportion of cases that recovered, but a very large change in another person’s score that did not cross the threshold would count as “no change”.

As discussed above, a particular challenge in child mental health is that different measures use different criteria for determining thresholds and their robustness as a measure of clinically significant difficulties is open to challenge. An additional challenge is how to address the issue of different scales for the same children (where some may show “recovery”, some no change and some a move across the threshold to “caseness”). Different approaches to considering “recovery” in the light of these complexities need to be considered in the future.

Reliable change

Reliable change is a statistical method of assessing how likely it is that changes in scores could be attributed to random fluctuations or measurement error. The proportion of children whose change in scores indicates an amount of change greater than likely to be due to measurement error alone is calculated. The advantage of this method is that it gives greater confidence that genuine change has occurred. However, it does not indicate whether that amount of change is clinically relevant. Therefore the scores of some young people may reliably change but that amount of change may not be enough

to be clinically significant; conversely, the scores of other young people may change less than is needed for it to be counted as reliable – yet that amount of change is clinically significant. A particular issue for the child context is how to handle the situation where one scale has reliably changed and another reliably deteriorated or not changed reliably. Different approaches to this are used.

Reliable “recovery”

Reliable “recovery” refers to those cases that had both “recovery” and reliable improvement. Its strength is that it is seen as a more stringent definition of “recovery” as it requires both reliable change and movement across a threshold to have occurred. However, it presents challenges when applied to a case with multiple measures, as is generally the case for children. If applied in any straightforward way, it would mean a case could be categorised as reliably “recovered” when one scale shows “recovery” and a completely different scale shows reliable improvement. One method to address this involves considering if there is any scale showing reliable “recovery”, in the absence of any scale showing reliable deterioration and all scales being below threshold at case closure.

Standardised effect sizes

Effect sizes offer a standardised way to quantify the average amount of change in a sample (as opposed to in an individual) that has occurred during treatment by comparing it to how spread out individual change scores are (i.e. using the standard deviation). A strength is that this addresses the issue of giving a sense of the magnitude of change. However, it does not indicate whether the change is clinically significant. Moreover, different measures are likely to have different effect sizes so aggregating measures across a child who had completed multiple scales, or comparing children who have completed different scales, is likely to be problematic. Another weakness of this approach is that we don’t know which population we are generalising to with each measure, since we don’t know how children were selected into each measure. See below for more detailed comment on the added value score.

In this report effect sizes are calculated for different measures separately because of this issue.

Added value score

The added value metric uses changes in the parent reported Strengths and Difficulties Questionnaire (SDQ)

to estimate the difference between changes seen for children by services, compared with anticipated changes if they had not. This is based on data collected from a community sample and therefore attempts to introduce a counterfactual data into the metric. There has been some testing of this approach, which has provided some support for its validity (Ford, Hutchings, Bywater, Goodman & Goodman, 2009), but its applicability to those accessing specialist services has not been fully explored. In addition, the added value score can currently only be used with the parent-reported Strengths and Difficulties Questionnaire and not other measures. A major weakness of this approach in the current data is the unknown path of selection into having an SDQ. The SDQ was just one measure among many; there was no requirement to fill in an SDQ if a different outcome measure was deemed more appropriate; it is not known how clinicians decided whether the SDQ was the appropriate outcome measure or not, and it is likely that different clinicians used different criteria to decide this. The added value score is calculated in this report using the limited data available for this purpose.

The seven approaches listed above have been used to analyse the data in this report.


Chapter 7 considers experience and outcomes for closed treatment cases (of at least three events, excluding assessment only – see Chapter 4). The following analyses are shown:

- **Raw scores on experience of service questionnaire at end of treatment**
- **Raw change in relation to paired goals**

Chapter 8 considers outcomes for paired above threshold closed treatment cases (of at least three events, excluding assessment only, with at least one child or parent report scale above threshold at outset and at least one paired child or parent report scale). The following analyses were undertaken:

- **“Recovery” rates** were calculated using an adaptation of the Adult IAPT approach i.e. no completed scale above threshold at the second time point.
- **Reliable improvement rates** were calculated using an adaptation of the Adult IAPT approach i.e. at least one completed scale improved more than what would be likely to be attributable to measurement error and no completed scale deteriorated more than what would be likely to be attributable to measurement error.

- **Reliable deterioration** was considered to have happened where scores on any one scale had deteriorated more than what would be likely to be attributable to measurement error even if other scales had reliably improved.
- **Reliable “recovery” rates** were calculated using a combination of reliable improvement and “recovery” rate calculations, where both must have occurred in at least one scale; and where none indicated reliable deterioration.
- **Standardised effect sizes** were calculated for each scale considered separately.
- **Added value score** was calculated using comparison of parent-rated scores on a measure (SDQ) where an algorithm has been developed to consider the degree of difference between these change scores and those found in a population largely not seen by services.



Chapter 7:
**Considering
experience
and goals**

This chapter sets out the findings in terms of young people's and parents' experiences of the services they received and the degree to which young people moved towards their agreed goals.

All analyses in this chapter relate to the closed treatment cases sample (see Figure 2). Characteristics of those who provided experience and goals data and how they differed from the overall sample of closed treatment cases are outlined in Table 6 below.

Table 6: Key characteristics of the closed treatment cases sample and those from the closed treatment cases (sample with child/parent ESQ/Goals)

	Closed treatment cases sample (n=23,373)	Closed treatment cases sample with child ESQ (n=3,196)	Closed treatment cases sample with parent ESQ (n=2,698)	Closed treatment cases sample with child Goals (n=2,784)	Closed treatment cases sample with parent Goals (n=686)
Definition	Closed treatment cases with three or more recorded events, of which at least one was for something other than assessment	Closed treatment cases with child-reported ESQ	Closed treatment cases with parent-reported ESQ	Closed treatment cases with paired child-reported goals	Closed treatment cases with paired parent-reported goals
Age: M (SD)	12.4 (3.8)	13.8 (2.8)	11.1 (4)	13.4 (3.1)	7.7 (3.7)
Missing: % (N)	0.1 (31)	0.2 (7)	0.3 (9)	0.1 (4)	0
Females: % (N)	57.1 (13,343)	64.8 (2,070)	52.3 (1,410)	65.2 (1,815)	37.2 (255)
Missing: % (N)	0.1 (16)	0.1 (4)	0.1 (2)	0.1 (2)	0.1 (1)
White: % (N)	80.4 (14,081)	83.9 (2,094)	82.3 (1,761)	88.2 (1,781)	77.6 (434)
Asian: % (N)	5.4 (938)	4.9 (123)	5.1 (110)	2.5 (51)	5.4 (30)
Black: % (N)	6.2 (1,085)	4.6 (116)	5.1 (110)	3.2 (65)	8.8 (49)
Mixed: % (N)	5.3 (922)	4.3 (107)	5.2 (112)	4 (80)	5.4 (30)
Other: % (N)	2.8 (482)	2.3 (57)	2.2 (48)	2.1 (43)	2.9 (16)
Not stated or missing: % (N)	25.1 (5,865)	21.9 (699)	20.6 (557)	27.5 (764)	18.5 (127)

Note: percentages are calculated on the cases in the relevant sample with valid data for each variable.

Experience of service

The Experience of Service Questionnaire (ESQ; Attride-Stirling, 2002) was used to explore client satisfaction with services. The ESQ is typically administered at review or case closure, and covers questions relating to the experience of the care received (e.g. “I was treated well by the people who saw me”) as well as the environment of the service (e.g. “The facilities here are comfortable” (e.g. waiting area).

Table 7 and Table 8 show child-reported and parent-reported responses about their experience of service.

“Convenient appointments” and “comfortable facilities” were the lowest rated items from the child-reported perspective (62% and 64%) and “treated well” and “listened to” were the highest rated items (87% and 85%). “Convenient appointments” and “convenient location” were the lowest rated items from the parent-reported perspective (69% and 78%) and “treated well” and “views and worries” were the highest rated items in parent reports (91% and 89%).

Table 7: Experience of service child-reported responses

Item	Not true		Partly true		Certainly true	
Listened to	184	5.7%	311	9.6%	2729	84.6%
Easy to talk to	237	7.4%	704	21.8%	2282	70.8%
Treated well	183	5.7%	243	7.5%	2808	86.8%
Views and worries	236	7.4%	275	8.6%	2699	84.1%
Know how to help	206	6.4%	611	19.1%	2378	74.4%
Given enough explanation	187	5.9%	619	19.7%	2343	74.4%
Working together	203	6.6%	500	16.2%	2386	77.2%
Comfortable facilities	233	7.4%	912	29.1%	1993	63.5%
Convenient appointments	309	9.8%	889	28.2%	1953	62.0%
Convenient location	266	8.4%	801	25.3%	2105	66.4%
Recommend to a friend	225	7.2%	565	18.0%	2348	74.8%
Good help	202	6.3%	365	11.4%	2629	82.3%



Figure 7: Child-report ESQ

Table 8: Experience of service parent-reported responses

Item	Not true		Partly true		Certainly true	
Listened to	118	4.4%	201	7.5%	2375	88.1%
Easy to talk to	124	4.6%	227	8.4%	2346	87.0%
Treated well	115	4.3%	124	4.6%	2459	91.1%
Views and worries	137	5.1%	168	6.2%	2399	88.7%
Know how to help	129	4.8%	371	13.8%	2192	81.4%
Given enough explanation	122	4.5%	364	13.6%	2195	81.9%
Working together	143	5.4%	274	10.4%	2228	84.2%
Comfortable facilities	127	4.8%	432	16.3%	2086	78.9%
Convenient appointments	253	9.4%	585	21.8%	1849	68.8%
Convenient location	162	6.1%	434	16.3%	2068	77.6%
Recommend to a friend	127	4.7%	206	7.7%	2349	87.6%
Good help	124	4.6%	192	7.1%	2382	88.3%

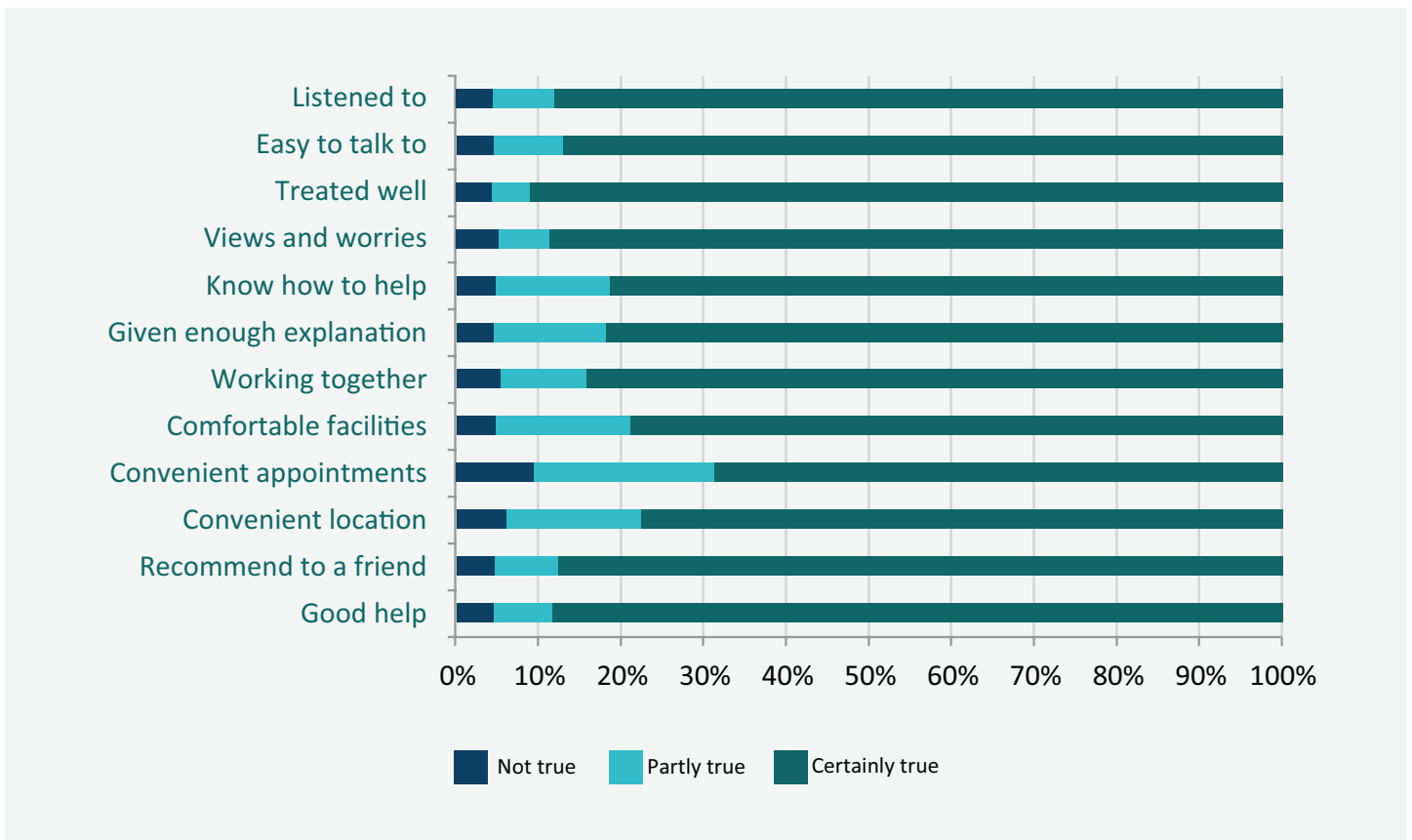


Figure 8: Parent-report ESQ

Achievement of goals

Goal setting and tracking is used in child and young people’s mental health services as a bespoke method of tracking outcomes that are meaningful to the child and their family and are central to shared decision-making (Law & Jacob, 2015). The method of tracking goals for those taking part in the CYP IAPT programme was to set up to three goals at the outset of treatment and for young people to rate how close they were to achieving those goals at frequent intervals on a scale of 0 (not achieved) to 10 (achieved). Only matched goal information was taken into consideration in this analysis, i.e. if a case had three goal scores at Time 1 and two goal scores at Time 2, the average for both time points would be taken from the two goals with information at both time points. Overall, the pre-post effect size for child-reported goals ($n=2,784$, 12% closed treatment cases) was 1.61, with a mean at Time 1 of 3, and a mean at Time 2 of 6.8. The average change in self-reported goals score was 3.73 points; 86% reported movement towards their goals of at least one point; 5% reported movement away from their goals; 10% reported no change in their goal scores. The pre-post effect size for parent-reported goals was 1.76, with a mean at Time

1 of 2.5, and a mean at Time 2 of 6.3 ($n=686$, 3% closed treatment cases). The average change in parent-reported goals score was 3.7 points; 87% reported movement towards their goals of at least one point; 3% reported movement away from their goals; 10% reported no change in their goal scores.



Chapter 8:

**Considering
“recovery”,
reliable change
and effect sizes**

Chapter 8 sets out the findings of the analysis of the measures used to track symptoms of young people. It first describes the characteristics of children with paired outcome data before presenting the findings using a range of different analytic approaches.

Sample characteristics

All analyses in this chapter relate to the paired outcomes for above threshold closed treatment sample (see

Figure 2). Characteristics of those who provided outcome data and how they differed from the overall clinical sample are outlined in Table 9 below.

The child report sample was older with more girls and more white than the overall sample; the parent report sample was younger with more boys and more white than the overall sample.

Table 9: Key characteristics of the clinical sample (5), paired outcomes for above threshold closed treatment samples (6) and added value score sample (7)

	Above threshold closed treatment sample (n=15,536)	Paired outcomes for above threshold closed treatment sample with child measures (n=5918)	Paired outcomes for above threshold closed treatment sample with parent measures (n=3,699)	Added value score sample (n=1,010)
Definition	Closed cases with three or more recorded events, of which at least one was for something other than assessment, and with at least one measure in clinical range at T1	Clinical sample cases with at least one paired child-reported measure (which includes the one in clinical range at T1)	Clinical sample cases with at least one paired parent-reported measure (which includes the one in clinical range at T1)	Clinical sample cases with paired parent SDQ (including total difficulties, emotional problems and impact in clinical range at T1)
Age: mean (SD)	12.8 (3.6)	14.3 (2.4)	11.2 (3.7)	11.1 (3.5)
Missing: % (N)	0.1 (20)	0.2 (9)	0.2 (7)	0
Females: % (N)	61.1 (9,495)	72.4 (4,287)	53.6 (1,982)	51.9 (524)
Missing: % (N)	0.1 (12)	0.1 (6)	0.1 (2)	0
White: % (N)	82.3 (9,424))	86.7 (3,662)	84.8 (2,451)	85.4 (696)
Asian: % (N)	4 (459)	3.1 (131)	3.1 (91)	2.2 (18)
Black: % (N)	5.7 (651)	3.6 (151)	4.9 (142)	5.1 (42)
Mixed: % (N)	5.2 (593)	4 (169)	4.9 (141)	5 (41)
Other: % (N)	2.8 (318)	2.7 (113)	2.2 (64)	2.2 (18)
Not stated or missing: % (N)	26.3 (4,091)	28.6 (1,692)	21.9 (810)	19.3 (195)

Note: percentages are calculated on the cases in the relevant sample with valid data for each variable.

Number of scales completed

The child report sample completed a total of 12,128 scales, the mean number of scales completed by any one child was 3.8 (SD = 2.5) the median was 3 and the range was 1–13.

The parent report sample completed a total of 10,438 scales, the mean number of scales completed about any child by their parent was 3.6 (SD = 2.2), the median was 3 and the range was 1–12.

“Recovery”

“Recovery” considers whether the child or young person moved from being above the threshold on a scale at the first recorded time point to below the threshold at the last recorded time point. Following the Adult IAPT approach (Gyani, Shafran, Layard & Clark, 2013; NHS England, 2014, 2016), we considered “recovered” children that moved from being above threshold on at least one measure at the beginning of treatment, to being below threshold on all measures following treatment.

Adaptations to the criteria made in the light of the child context comprised the fact that “recovery” criteria adopted by Adult IAPT were extended to take into account the much wider range of self- and parent-reported measures, including their differing thresholds that were potentially included in the analysis for any one case. To be classed as “recovered”, no completed scale could be within the clinical range at a second time point, and to be classed as reliably improved, at least one scale had to have reliably improved and no completed scale could have reliably deteriorated. As noted previously, the inclusion criteria for treatment included only those cases with three events to take into account the structure of the dataset for children. This was agreed at the outset of the programme because an event could relate to the uploading of a questionnaire and not necessarily contact.

Based on the child-reported measures, 36% (95% CI 34.77% – 37.23%) of children’s scores were considered “recovered”, whereby they were above threshold on at least one scale at the outset of treatment and below threshold on *all measures* at the end of treatment. Based on the parent-reported measures, 28% (95% CI 26.55% – 29.47%) of children’s scores were considered “recovered” at the end of treatment.

Of the 1,758 cases with paired child-report data (6% of the closed treatment cases) in the pre-CYP IAPT dataset, which only included the SDQ subscales, 33% (95% CI 30.78% – 35.24%) indicated “recovery” and for the 3,708 from parent-reported measures (12% of closed treatment cases) “recovery” was indicated for 26% (95% CI 24.58% – 27.43%). When restricting the analysis to look at only the

SDQ subscales in the CYP IAPT dataset ($n=2,347$), these figures are comparable: 32% (95% CI 30.1% – 33.92%) for child-reported measures and 27% (25.27% – 28.75%) for parent-reported measures ($n=2,558$).

Reliable change

Reliable change considers the amount of change from one time point to another, relative to the properties of the measure used, thereby counting as reliable change that is unlikely to be attributable to measurement error alone (Jacobson & Truax, 1991). Following the Adult IAPT approach (Gyani et al., 2013, p. 599), the three categories of change were considered:

1. **Reliable improvement:** where a score on at least one measure changed enough for it to be considered statistically reliable and no other score reliably deteriorated.
2. **No reliable change:** where a score on *all* completed measures did not show any statistically reliable change.
3. **Reliable deterioration:** where a score on one or more measures changed enough in a negative direction for it to be considered statistically reliable.

Adaptations to the criteria made in the light of the child context involved extending the definition of reliable change to encompass all scales; to take into account the much wider range of self- and parent-reported measures. To be classed as reliably improved at least one scale had to have reliably improved and no scale could have reliably deteriorated. As noted previously, the inclusion criteria for treatment included only those cases with three events to take into account the structure of the dataset for children. This was agreed at the outset of the programme as an event could relate to the uploading of a questionnaire and not necessarily contact.

Figure 9 and Figure 10 show the aggregated reliable change results across the sample of cases that were seen for a course of treatment, and who had at least one paired outcome measure that was indicated as being in the clinical range at the start of treatment (the paired outcomes for above threshold closed treatment sample), separately for child- and parent-reported measures. Based on the child-reported measures, 52% (95% CI 51.94 – 52.25%) of children’s scores reliably improved, 37% (95% CI 37.04% – 37.34%) had no reliable change and 11% reliably deteriorated (see Figure 9). On the parent-reported measures, 40% (95% CI 40.34% – 40.7%) of children’s scores showed reliable improvement, 51% indicated no reliable change and 9% (95% CI 8.71% – 9.07%) reliably deteriorated (see Figure 10).

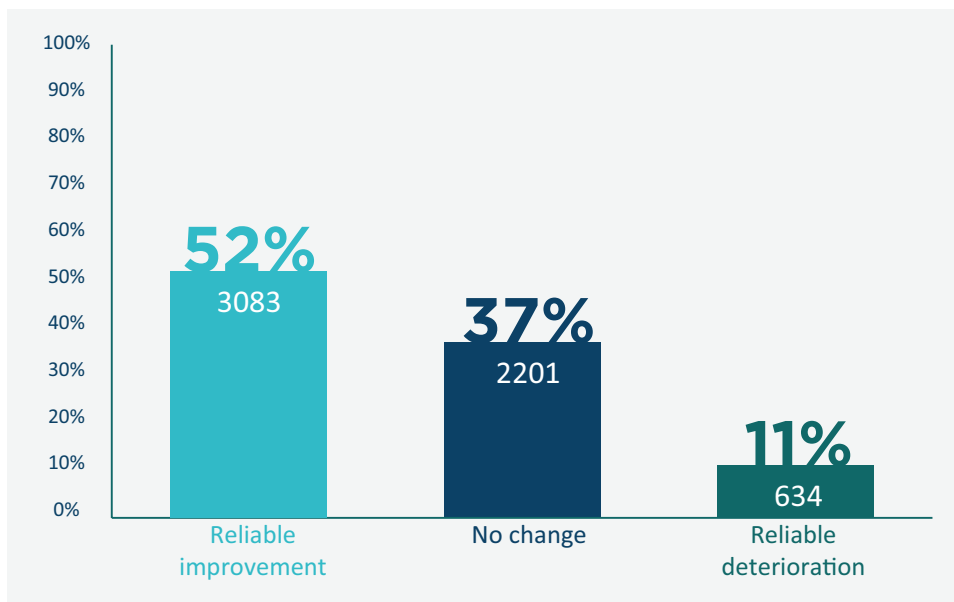


Figure 9: Reliable change: child-reported measures

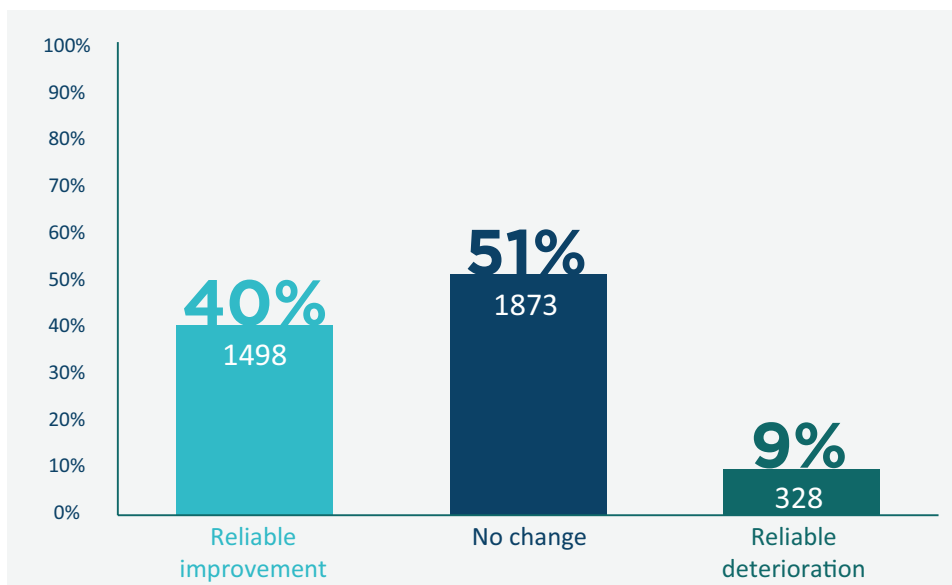


Figure 10: Reliable change: parent-reported measures

In the pre-CYP IAPT dataset, according to child-reported measures, around 36% (95% CI 35.88% – 36.14%) of children’s scores indicated reliable improvement, 57% (95% CI 56.88% – 57.14%) had a degree of change that was not enough for it to be considered reliable and 7% (95% CI 6.88% – 7.14%) became reliably worse ($n=6,767$). When restricting the analysis to look at only the SDQ subscales in the CYP IAPT dataset, these figures are comparable: 36% (95% CI 35.54% – 35.76%) reliable improvement; 59% (95% CI 58.31% – 58.54%) no reliable change; and 6% (95% CI 5.83% – 6.06%) reliable deterioration; $n=8,324$.

According to parent-reported measures in the pre-CYP IAPT dataset, around 35% (95% CI 34.91% – 35.1%) of scores showed reliable improvement, 57% (95% CI 56.91% – 57.1%) had a degree of change that was not enough for it to be considered reliable and 8% (95% CI 7.91% – 8.1%) showed scores that were reliably worse ($n=12,865$; 2007–2010). When restricting the analysis to look at only the SDQ subscales in the CYP IAPT dataset, these figures are comparable (36%, 95% CI 35.89% – 36.12%, reliable improvement, 57%, 95% CI 56.89% – 57.12%, no reliable change and 7%, 95% CI 6.89% – 7.12%, reliable deterioration; $n=8,702$).

“Recovery” or reliable change

Overall, 60% (95% CI 58.37% – 61.63%) of children’s scores ($n=3,526$) “recovered” and/or reliably improved based on the child-reported measures, and 51% (95% CI 48.72% – 53.28%) based on the parent-reported measures ($n=1,891$).

In the pre-CYP IAPT dataset, these figures were 49% (95% CI 47.8% – 50.2%) according to child-report and 45% (95% CI 44.14% – 45.86%) according to parent report; $n=6,767$ and $n=12,865$). When including the same measures in the CYP IAPT dataset, these figures are comparable (47%, 95% CI 45.92% – 48.08%, and 46%, 95% CI 44.95% – 47.05%; $n=8,324$ and $n=8,702$).

Reliable “recovery”

It is not possible to use any simple way to calculate reliable “recovery” (“recovery” and reliable change) because of the larger number of measures involved. This issue was addressed adapting the Adult IAPT approach of considering reliable “recovery” on any one scale in the absence of deterioration and where every scale was below threshold at the end of treatment (Gyani et al., 2013, p. 599).

To estimate reliable “recovery” **on any one scale** in the absence of deterioration and where every scale was below threshold at end of treatment rates, data were considered according to the following criteria: case closed with at least three events, excluding assessment only; at least one measure above threshold at T1; no measures above threshold at T2; at least one score moved from above to below threshold (“recovered”), and reliably improved on the same measure; no measure reliably deteriorated. Using these criteria 1,560 of the 5,918 children with paired measures showed reliable “recovery” (26.4%) and 617 of 3,699 parents with paired measures (16.7%).

An alternative approach to considering the metric of reliable “recovery” over multiple scales might be **random selection of scales**. This may be explored in future analyses.

Standardised pre-post effect sizes

Effect sizes offer a standardised way to quantify the amount of change that has occurred during treatment and provide an estimation that can easily be compared across studies using the same measures. However, because of the variation due to the different sensitivity of measures, they cannot be meaningfully compared or aggregated across measures (Coe, 2002) so this analysis is completed measure by measure as displayed in the forest plots in Figure 11 and Figure 12.

Standardised effect sizes and 95% confidence intervals were estimated using a bootstrapped implementation of the Becker method (1988), using 10,000 bootstrap samples taken using the R package ‘boot’ (Cant & Ripley, 2016).

It should be noted that since some measures represent difficulties (i.e. higher scores represent a higher level of severity) and others measure functioning/wellbeing (i.e. higher scores represent a lower level of severity), the above formula was adjusted for each measure, for consistency, such that a positive effect size is always indicative of an improvement in the patients’ condition (i.e. a decrease in the indicated level of severity).

The most commonly used subscales were those from the Revised Child Anxiety and Depression Scale (RCADS; Chorpita et al., 2000). This is likely to reflect the fact that anxiety and depression were the most prevalent presenting issues among the population in the dataset, apart from family relationship difficulties, which do not imply a particular set of symptoms. RCADS is a validated measure that comprises six subscales, five of which focus on different types of anxiety and the final subscale measures major depression.

As Figure 11 shows, the pre-post effect sizes on the subscales from the child-reported RCADS ranged from 0.73 (95% CI 0.70 – 0.77) for the separation anxiety subscale, to 1.22 (95% CI 1.17 – 1.28) for the generalised anxiety subscale. As Figure 12 shows, the pre-post effect sizes on the parent-reported RCADS subscales ranged from 0.54 (95% CI 0.50 – 0.58) for the depression subscale, to 0.78 (95% CI 0.73 – 0.82) for the generalised anxiety disorder subscale. The degree of change is more than that found in some research trial samples. For example, a pre-post effect size for the intervention group of 0.74 on the GAD subscale has been reported in an American sample of child-reported RCADS, which specifically looked at those who had received CBT (Ishikawa, Okajima, Matsuka & Sakano, 2007). However, the high levels of missing data mean such comparisons need to be treated with caution.

The second most commonly used subscales were from the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997), a brief behavioural screening tool used with 2–17 year olds. It comprises five subscales, four of which focus on different aspects of difficulties (hyperactivity, conduct, emotional and peer relationship difficulties), and scores on these subscales are added together to generate the total difficulties score. In addition, there is one subscale focused on positive behaviour (prosocial behaviour), and an impact subscale that asks questions about areas that could be affected by the young person’s behavioural issues, such as difficulties with home life or in learning (Goodman, 1997). The peer and prosocial subscales were not analysed for this purpose and the total difficulties not derived as it

would duplicate the findings from the subscales.

The SDQ is a well-validated tool that is widely used in practice and research as a general measure of mental health and wellbeing (Goodman, 2001). Children aged 11 years and over can complete the SDQ for themselves in addition to parental report and, as Figure 11 shows, on average their scores improved on all of the individual subscales, with effect sizes ranging from 0.63 (95% CI 0.59 – 0.67) on the impact subscale to 0.77 (95% CI 0.73 – 0.80) on the emotional difficulties subscale.

For children under 11 years a parent or carer will usually complete the SDQ instead of their child. The findings from the parent-reported SDQ are shown in Figure 12. As with child-reported scores, on average their scores improved on all of the individual subscales, ranging from 0.44 (95% CI 0.41 – 0.47) on the conduct subscale and 0.47 (95% CI 0.44 – 0.50) on the hyperactivity subscale to 0.61 (95% CI 0.58 – 0.64) on the emotional difficulties subscale. The degree of change for the parent-reported SDQ is slightly less than for the child-reported one. However, the two sets of data are not directly comparable. Although there will be overlap between the two groups, children in the self-report group had an average age of 14.3 years, whereas children in the parent-report group had an average of 11.2 years.

The pre-post effect sizes for the child- and parent-reported SDQ are in line with the pre-post effect sizes found in the pre-CYP IAPT dataset held by CORC, whereby the child-reported pre-post effect sizes ranged from 0.68 for the hyperactivity subscale to 0.89 for the conduct subscale. The parent-reported pre-post effect sizes ranged from 0.50 for the hyperactivity subscale to 0.77 for the impact subscale.

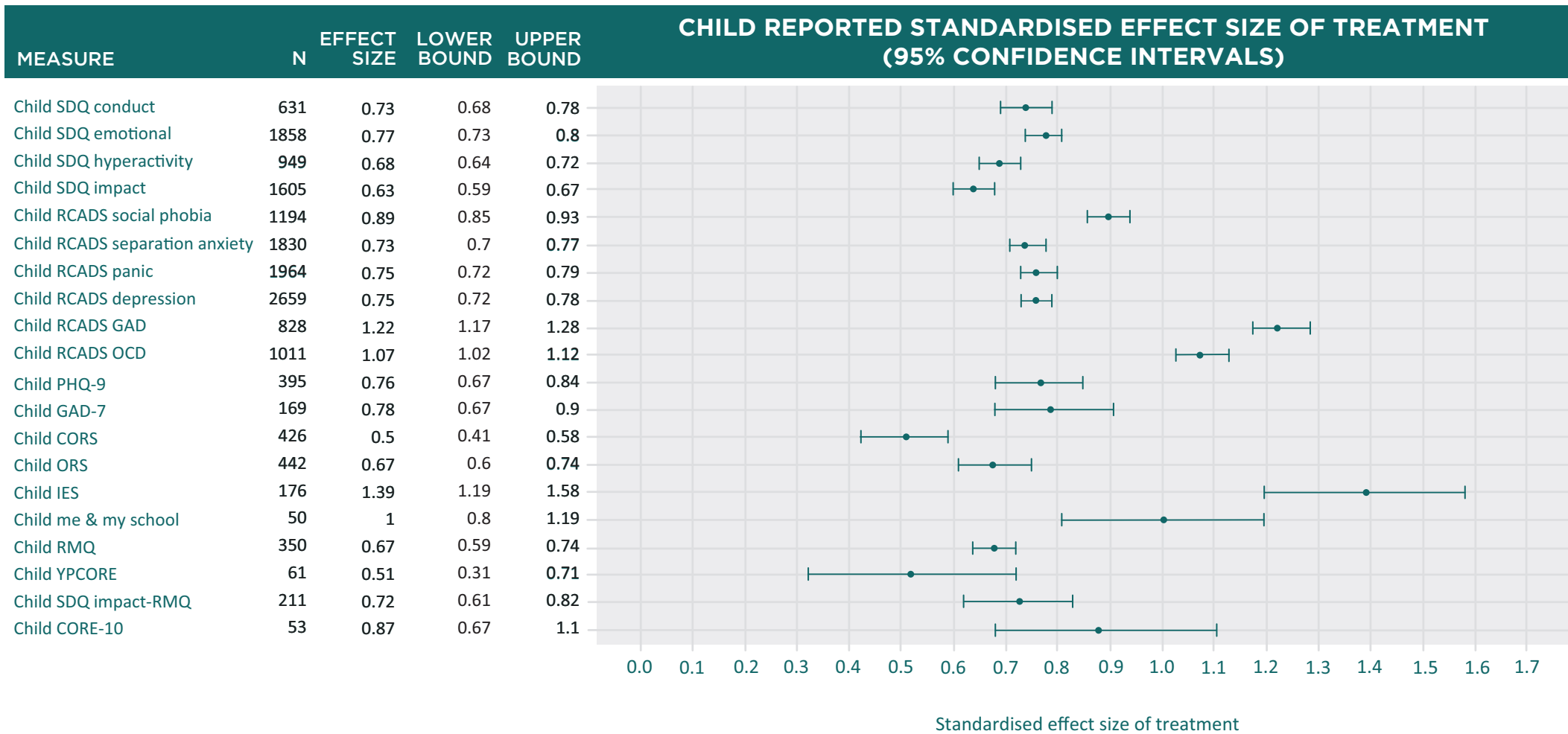


Figure 11: Standardised effect sizes and 95% confidence intervals: child-reported measures

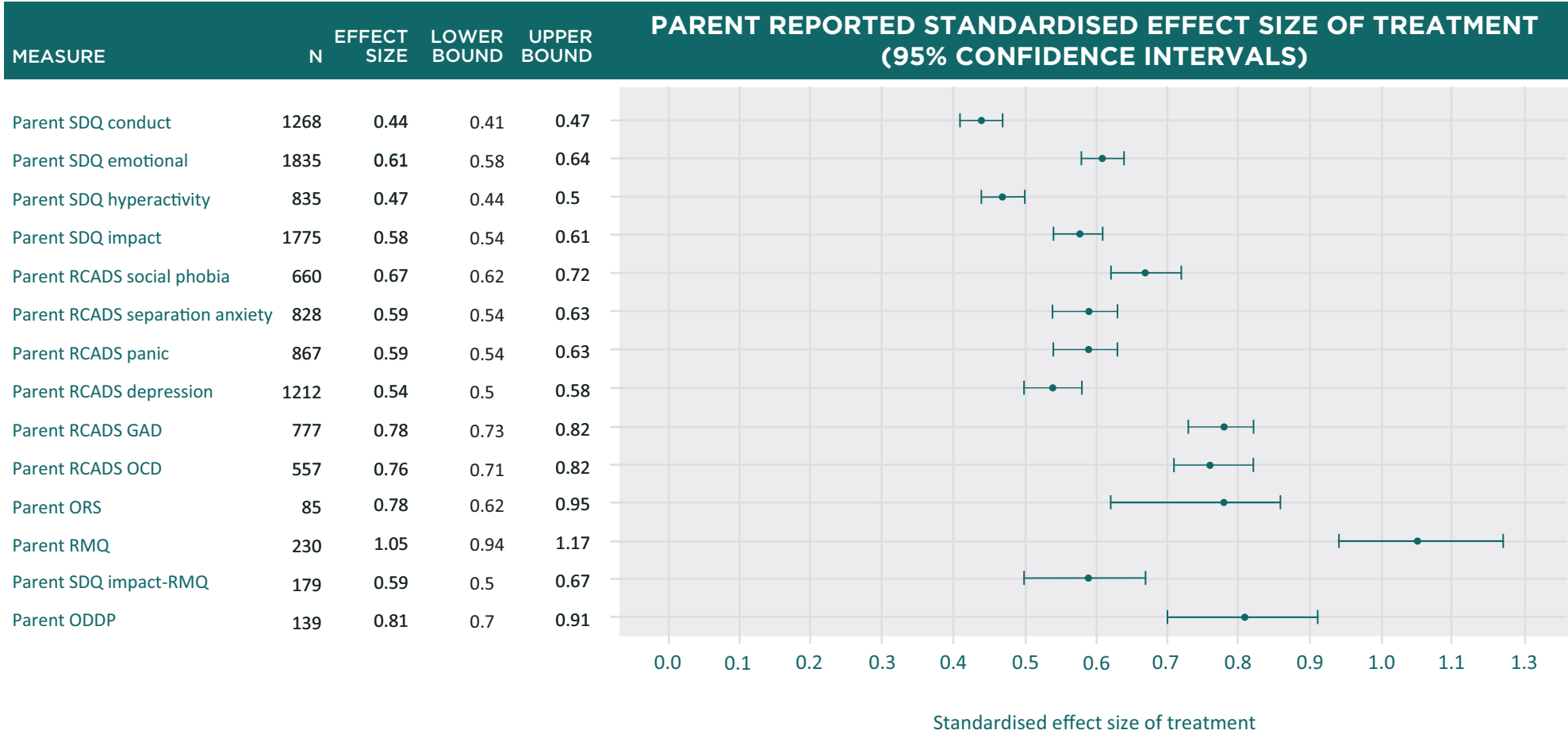


Figure 12: Standardised effect sizes and 95% confidence intervals: parent-reported measures


Estimating added value

The outcomes discussed in the previous sections indicate whether, on average, children and young people have improved, deteriorated or stayed the same (in the case of the overall analysis), and the extent of the change (in the case of the individual measures). These findings cannot on their own indicate whether the improvement would have happened even if children and young people had not received treatment from services. In order to explore whether the changes were likely to have happened in the absence of intervention or not, the average pre-post effect size for change in the parent-reported SDQ scores was compared with a national sample of children in the community who were not receiving treatment, using the formula suggested for an added value score (AVS; Ford et al., 2009).

The AVS allows outcomes in the present sample to be compared with the outcomes of a sample of children and young people with similar issues who were on the whole not seen by services. This gives an indication of the impact of the interventions as compared to spontaneous remission or regression to the mean (i.e. the statistical phenomenon that follow-up scores naturally migrate towards the mean, especially if they were high to start with). This represents the average additional change young people in the CYP IAPT sample experienced over and above what would be expected had they not received treatment.

To calculate the AVS, the filtering criteria for the analysis were such that the case had to have the following: a paired parent-reported SDQ, be indicated as a closed case and attended three or more sessions, and be identified as above the clinical range at the outset of treatment. To try to comply with the timeframe criteria for Time 2 (T2) set out by Ford et al., (2009) we explored the size of the sample when the data were filtered, such that the T2 measure was between four and eight months after the Time 1 (T1) measure; however, this reduced the overall sample to 63 cases, so analysis using this filter was not run.

For those with the relevant parent-completed SDQ information ($n=1,010$, .04% of closed treatment cases), the overall average AVS effect size was 0.26 ($SD= 0.99$, 95% CI = 0.2 – 0.32). The confidence intervals around the average AVS score for the overall sample did not cross zero, suggesting it was unlikely to be due to chance alone. The AVS for the CYP IAPT sample was significantly higher than that of 0.15 from the pre-CYP IAPT dataset held by CORC ($n=1,476$, .05% of closed treatment cases).



Chapter 9:
**Concluding
remarks**

This report is the first analysis of routinely collected child- and parent-reported outcomes and experience data collated from 75 mental health services in England. A range of outcome measures were used for the primary purpose of informing clinical decision-making with individual children and parents. In the absence of a commonly agreed method for analysing the data at a national level, several approaches to estimating outcomes were applied. Each method indicated different levels of change in the mental health, wellbeing or personal goals of a selected sample of children and young people using services. The findings are hindered by poor data completeness and the lack of a comparator group and should therefore be interpreted with caution.

Collecting routine outcome and experience data from children and parents in child mental health is an internationally recognised challenge. The CYP IAPT programme set itself the ambitious challenge of achieving 90% paired outcomes on one or more scales with thresholds for all closed treatment cases and to include information about education employment and training. The fact that this target was not achieved raises major challenges for the analysis and interpretation of data which have been discussed throughout this report. What data were collected were the result of immense efforts on the part of all those involved, given the major stresses on services at the time, challenges of IT systems and other pressures and factors discussed in the preceding chapters.

Although the findings are based on data that are flawed in a range of ways, and with a high degree of missing data (which we have termed FUPS), they are the best currently available and should be carefully used to support dialogue about outcomes. CORC recommends this report be used to inform facilitated stakeholder discussion involving practitioners, funders, service users, policy makers and others, along the lines outlined elsewhere and described as the MINDFUL approach (Wolpert et al., 2014c). In particular, it is recommended that the facilitator of such conversations should help these key stakeholders to: challenge their own and colleagues' confirmatory biases; maintain curiosity; apply the same standards of scrutiny to analytic findings that support prior beliefs as to analytic findings that are uncomfortable or not wished for; consider if any actions need to be taken in terms of quality assurance; consider possible initiatives that, even if not definitively indicated, may do more good than harm; and challenge the assumption that change is always more risky than status quo. Key topics for consideration in such facilitated discussions might include the best metrics for considering outcomes and appropriate targets.

Determining outcomes for those seen by services is clearly complex and multiple difficulties, multiple perspectives and the potential for different metrics of change are challenging. CYP IAPT helped produced consensus on a range of child- and parent-reported indicators of outcomes that are now included in the nationally mandated mental health services dataset. This means that a range of child- and parent-reported indicators of outcomes can now be collated and reported nationally by NHS Digital.

It is hoped that this dataset can continue to be drawn on as a form of intelligence to encourage thinking and decision-making in the complex but vital area of supporting children with mental health difficulties. Further analysis of this dataset and initiatives to support thinking about the findings that emerge will be undertaken by CORC and collaborators. They will consider different outcomes for different types of problems and measures within this dataset and support facilitated stakeholder dialogue about implications for practice in terms of potential targets for achievement and improvement.

CORC will continue to work with NHS Digital to share learning and help shape future practice. High quality data collection on outcomes and experience must be facilitated and incentivised to aid review and development of services. To achieve this, leadership focus, improved IT systems, better staff training and stronger incentives may need to be in place. Achieving parity of esteem between physical and mental health requires parity of data.




Chapter 10:
References

- Attride-Stirling, J. (2002). Development of methods to capture users' views of child and adolescent mental health services in clinical governance reviews: Commission for Health Improvement.
- Batty, M. J., Moldavsky, M., Foroushani, P. S., Pass, S., Marriott, M., Sayal, K., & Hollis, C. (2013). Implementing routine outcome measures in child and adolescent mental health services: from present to future practice. *Child and Adolescent Mental Health, 18*(2), 82–87. doi: 10.1111/j.1475-3588.2012.00658.x
- Becker, B. J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology, 41*(2), 257–278.
- British Psychological Society. (2015). What good looks like in psychological services for children, young people and their families. Retrieved from https://www.bps.org.uk/system/files/user-files/DCP%20Faculty%20for%20Children,%20Young%20People%20and%20their%20Families/public/cfcp_r_3.pdf
- Brown, A., Ford, T., Deighton, J., & Wolpert, M. (2014). Satisfaction in child and adolescent mental health services: Translating users' feedback into measurement. *Administration and Policy in Mental Health and Mental Health Services Research, 41*(4), 434–446.
- CAMHS EBPU – Child and Adolescent Mental Health Services Evidence Based Practice Unit. (2011). Current View. Retrieved from <http://pbrcamhs.org/wp-content/uploads/2012/10/Current-View-Form>.
- Cant, A. & Ripley, B. (2016). Boot: Bootstrap R (S-Plus) Functions. R package version 1. 3–18.
- Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., Meier, M.H., Ramrakha, S., Shalev, I., Poulton, R., & Moffitt, T. E. (2014). The p factor one general psychopathology factor in the structure of psychiatric disorders?. *Clinical Psychological Science, 2*(2), 119–137.
- Children and Young People's Improving Access to Psychological Therapies Programme. (2013). National curriculum for core, cognitive behavioural therapy, parenting training (3–10 year olds), systemic family practice, interpersonal psychotherapy for adolescents, supervision, and transformational service leadership. London: IAPT.
- Chorpita, B. F., Yim, L., Moffitt, C., Umemoto, L. A., & Francis, S. E. (2000). Assessment of symptoms of DSM-IV anxiety and depression in children: a revised child anxiety and depression scale. *Behaviour Research and Therapy, 38*(8), 835–855. doi: 10.1016/S0005-7967(99)00130-8
- Clark, D., & Oates, M. (2014). *Improving Access to Psychological Therapies. Measuring recovery in IAPT services. Version 2*. Retrieved from <http://iapt.nhs.uk/silo/files/measuring-recovery-2014.pdf>
- Coe, R. (2002). *It's the effect size, stupid! What effect size is and why it is important*. Paper presented at the British Educational Research Association annual conference, Exeter, UK.
- Coulter, A., Locock L., Ziebland S., & Calabrese, J. (2014). Collecting data on patient experience is not enough: they must be used to improve care. *BMJ 348*: g2225
- CYP IAPT CORC & CAMHS providers. (2014). CYP IAPT principles in CAMH services – values and standards: “Delivering with and delivering well”. London: CAMHS Press. Retrieved from <https://www.england.nhs.uk/wp-content/uploads/2014/12/delvr-with-delvrng-well.pdf>
- Deighton, J., Tymms, P., Vostanis, P., Belsky, J., Fonagy, P., Brown, A., Martin, A., Patalay, P., Wolpert, M. (2013). The Development of a School-Based Measure of Child Mental Health. *Journal of Psychoeducational Assessment, 31*(3), 247–257. doi: 10.1177/0734282912465570
- Deighton, J., Croudace, T., Fonagy, P., Brown, J., Patalay, P., & Wolpert, M. (2014). Measuring mental health and wellbeing outcomes for children and adolescents to inform practice and policy: a review of child self-report measures. *Child and Adolescent Psychiatry and Mental Health, 8*(1): 14. doi: 10.1186/1753-2000-8-14
- Department of Health. (2015). *Future in Mind: Promoting, protecting and improving our children and young people's mental health and wellbeing*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/414024/Childrens_Mental_Health.pdf
- Edbrooke-Childs, J., Wolpert, M., & Deighton, J. (2014). Using Patient Reported Outcome Measures to Improve Service Effectiveness (UPROMISE): Training clinicians to use outcome measures in child mental health. *Administration and Policy in Mental Health and Mental Health Services Research, 43*(3), 302–308.
- Fixsen, D. L., Naoom, S. F., Blasé, K. A., Friedman, R. M., & Wallace, F. (2005). Implementation Research: Training clinicians to use outcome measures in child mental health. *Administration and Policy in Mental Health and Mental Health Services Research, 43*(3), 302–308.

- Fixsen, D. L., Blasé, K. A., Naoom, S. F., & Wallace, F. (2009). Core Implementation Components. *Research on Social Work Practice, 19*(5), 531–540. doi: 10.1177/1049731509335549
- Fixsen, D. L., Blasé, K. A., & Van Dyke, M. K. (2011). Mobilizing Communities for Implementing Evidence-Based Youth Violence Prevention Programming: A Commentary. *American Journal of Community Psychology, 48*(1), 133–137. doi: 10.1007/s10464-010-9410-1
- Fleming, I., Jones, M., Bradley, J., & Wolpert, M. (2016). Learning from a learning collaboration: The CORC approach to combining research, evaluation and practice in child mental health. *Administration and Policy in Mental Health and Mental Health Services Research, 43*(3), 297–301.
- 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. *British Journal of Psychiatry, 194*(6), 552–558. doi: 10.1192/bjp.bp.108.052373
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581–586. doi: 10.1111/j.1469-7610.1997.tb01545.x
- Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(11), 1337–1345. doi: 10.1097/00004583-200111000-00015
- Gyani, A., Shafran, R., Layard, R., & Clark, D. M. (2013). Enhancing recovery rates: Lessons from year one of IAPT. *Behaviour Research and Therapy, 51*(9), 597–606. doi:10.1016/j.brat.2013.06.004
- Hindley, P. (2014). Written evidence for the House of Commons Select Committee Inquiry into Child and Adolescent Mental Health Services from the Faculty of Child and Adolescent Psychiatrists. London, Royal College of Psychiatrists.
- Ishikawa, S.-i., Okajima, I., Matsuoka, H., & Sakano, Y. (2007). Cognitive Behavioural Therapy for Anxiety Disorders in Children and Adolescents: A Meta-Analysis. *Child and Adolescent Mental Health, 12*(4), 164–172. doi: 10.1111/j.1475-3588.2006.00433.x
- 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. doi: 10.1037/0022-006X.59.1.12
- Johnston, C., & Gowers, S. (2005). Routine Outcome Measurement: A Survey of UK Child and Adolescent Mental Health Services. *Child and Adolescent Mental Health, 10*(3), 133–139. doi: 10.1111/j.1475-3588.2005.00357.x
- Jones, M., Hopkins, K., Kyrke-Smith, R., Davies, R., Vostanis, P., Wolpert, M. (2013). Current View Tool Completion Guide. London: CAMHS Press. Retrieved from http://pbrcamhs.org/wp-content/uploads/2012/10/Current-View-web_option110213.pdf
- Keogh, B. (2013, 16 July). Review into the quality of care and treatment provided by 14 hospital trusts in England: Overview report. Retrieved from <http://www.nhs.uk/NHSEngland/bruce-keogh-review/Documents/outcomes/keogh-review-final-report.pdf>
- Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a Brief Depression Severity Measure. *Journal of General Internal Medicine, 16*(9), 606–613. doi: 10.1046/j.1525-1497.2001.016009606.x
- Lambert, M. J., Whipple, J. L., Smart, D. W., Vermeersch, D. A., Nielsen, S. L., & Hawkins, E. J. (2001). The effects of providing therapists with feedback on patient progress during psychotherapy: Are outcomes enhanced?. *Psychotherapy research, 11*(1), 49–68.
- Law, D., & Jacob, J. (2015). *Goals and Goal Based Outcomes (GBOs): some useful information* (Third ed.). London: CAMHS Press.
- Law, D., & Wolpert, M. (Eds.). (2014). *Guide to Using Outcomes and Feedback Tools With Children, Young People and Families* (Second ed.). London: CAMHS Press.
- Leamy, M., Bird, V., Le Boutillier, C., Williams, J., & Slade, M. (2011). Conceptual framework for personal recovery in mental health: systematic review and narrative synthesis. *The British Journal of Psychiatry, 199*(6), 445–452.
- Lilford, R., Mohammed, M.A., Spiegelhalter, D., & Thomson, R. (2004). Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma, *The Lancet*, Volume 363 (9415), 1147–1154.
- McCoy, A. B., Wright, A., Kahn, M. G., Shapiro, J. S., Bernstam, E. V., & Sittig, D. F. (2013). Matching identifiers in electronic health records: implications for duplicate records and patient safety. *BMJ Quality & Safety, 22*(3), 219–224. doi: 10.1136/bmjqs-2012-001419

- Moran, P., Kelesidi, K., Guglani, S., Davidson, S., & Ford, T. (2012). What do parents and carers think about routine outcome measures and their use? A focus group study of CAMHS attenders. *Clinical Child Psychology and Psychiatry*, 17(1), 65–79. doi: 10.1177/1359104510391859
- NHS Benchmarking Network. (2013). CAMHS Benchmarking report. Retrieved from [http://www.rcpsych.ac.uk/pdf/CAMHS%20Report%20Dec%202013%20v1\(1\).pdf](http://www.rcpsych.ac.uk/pdf/CAMHS%20Report%20Dec%202013%20v1(1).pdf)
- NHS England. (2013). Children and Young People's Improving Access to Psychological Therapies: National Curriculum. Retrieved from http://www.cypiapt.org/docs/CYP_Curriculum_December_2013.pdf
- NHS England. (2014). Improving Access to Psychological Therapies: Measuring Improvement and Recovery, Adult Services, version 2. Retrieved from <http://iapt.nhs.uk/silo/files/measuring-recovery-2014.pdf>
- NHS England. (2016). Good Clinical Outcomes. Retrieved from <http://www.iapt.nhs.uk/good-clinical-outcomes/>
- Nilsen, T. S., Handegård, B. H., Eisemann, M., & Kvernmo, S. (2015). Evaluating change in symptomatic and functional level of children and youth with emotional disorders: a naturalistic observation study. *European Child & Adolescent Psychiatry*, 24(10), 1219–1231.
- Patalay, P., Deighton, J., Fonagy, P., Vostanis, P., & Wolpert, M. (2014). Clinical validity of the Me and My School questionnaire: a self-report mental health measure for children and adolescents. *Child and Adolescent Psychiatry and Mental Health*, 8(17). Doi: 10.1186/1753-2000-8-17
- Patalay, P., Fonagy, P., Deighton, J., Belsky, J., Vostanis, P., & Wolpert, M. (2015). A general psychopathology factor in early adolescence. *The British Journal of Psychiatry*, 207(1), 15–22.
- Royal College of Paediatrics and Child Health. (2015). National Paediatric Diabetes Audit. Retrieved from <http://www.rcpch.ac.uk/system/files/protected/page/NPDA%202014-15%20Lay%20Summary.pdf>
- Sapyta, J., Riemer, M., & Bickman, L. (2005). Feedback to clinicians: Theory, research, and practice. *Journal of Clinical Psychology*, 61(2), 145–153.
- Spitzer, R. L., Kroenke, K., Williams, J. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The gad-7. *Archives of Internal Medicine*, 166(10), 1092–1097. doi: 10.1001/archinte.166.10.1092
- Walter, G., Cleary, M., & Rey, J. (1998). Attitudes of mental health personnel towards rating outcome. *Journal of Quality in Clinical Practice*, 18(2), 109–115.
- Warren, J. S., Nelson, P. L., Mondragon, S. A., Baldwin, S. A., & Burlingame, G. M. (2010). Youth psychotherapy change trajectories and outcomes in usual care: Community mental health versus managed care settings. *Journal of consulting and clinical psychology*, 78(2), 144.
- Weisz, J. R., Chorpita, B. F., Frye, A., Ng, M. Y., Lau, N., Bearman, S. K., Uqueto, A.M., Langer, D.A., & Hoagwood, K. E. (2011). Youth Top Problems: using idiographic, consumer-guided assessment to identify treatment needs and to track change during psychotherapy. *Journal of Consulting and Clinical Psychology*, 79(3), 369–80. Doi: 10.1037/a0023307.
- Wolpert, M., Ford, T., Trustam, E., Law, D., Deighton, J., Flannery, H., & Fugard, R. J. (2012). Patient-reported outcomes in child and adolescent mental health services (CAMHS): Use of idiographic and standardized measures. *Journal of Mental Health*, 21(2), 165–173.
- Wolpert, M. (2014a). Uses and Abuses of Patient Reported Outcome Measures (PROMs): Potential Iatrogenic Impact of PROMs Implementation and How It Can Be Mitigated. *Administration and Policy in Mental Health and Mental Health Services Research*, 41(2), 141–145. doi: 10.1007/s10488-013-0509-1
- Wolpert M, Deighton J, De Francesco D, Martin P, Fonagy P, Ford T. (2014b). From 'reckless' to 'mindful' in the use of outcome data to inform service-level performance management: perspectives from child mental health. *BMJ Quality & Safety*. 23(4), 272–6.
- Wolpert, M., Görzig, A., Deighton, J., Fugard, A. J. B., Newman, R., & Ford, T. (2015a). Comparison of indices of clinically meaningful change in child and adolescent mental health services: difference scores, reliable change, crossing clinical thresholds and 'added value' – an exploration using parent rated scores on the SDQ. *Child and Adolescent Mental Health*, 20(2), 94–101. doi: 10.1111/camh.12080
- Wolpert, M., Vostanis, P., Young, S., Clark, B., Davies, R., Fleming, I., Howey, L., Howley, P., Macdougall, A., Martin, P., Martin, T., Payne, C., Ritchie, B., Senior, R., York, A., & Whale, A. (2015b) *Child and Adolescent Mental Health Services Payment System Project: Final Report*. London: CAMHS Press. Retrieved from <http://pbrcamhs.org/final-report/>

- Wolpert, M. (2016). Failure is an option. *The Lancet Psychiatry*. Retrieved from [http://www.thelancet.com/pdfs/journals/lanpsy/PIIS2215-0366\(16\)30075-X.pdf](http://www.thelancet.com/pdfs/journals/lanpsy/PIIS2215-0366(16)30075-X.pdf)
- Yeh, M., & Weisz, J. R. (2001). Why Are We Here at the Clinic? Parent-Child (Dis) Agreement on Referral Problems at Out-Patient Treatment Entry. *Journal of Consulting and Clinical Psychology, 69*, 1018–1025. doi: 10.1037/0022-006X.69.6.1018
- YoungMinds. (2014). Stop cutting CAMHS services. Retrieved from http://www.youngminds.org.uk/about/our_campaigns/cuts_to_camhs_services



Chapter 10:
Appendices

Appendix A

CYP IAPT: Service transformation programme

The Children and Young People’s Improving Access to Psychological Therapies (CYP IAPT) service transformation programme launched in 2011 involved geographical partnerships between NHS Local Authorities and voluntary sector providers (termed partnerships). Groups of partnerships were linked with specific higher education institutions (HEIs) across five areas (termed collaboratives):

London and the South East, the North West, Oxford/Reading, the Yorkshire, Humber and the North East, and the South West. The programme was rolled out over 4 years in four waves of implementation: wave 1 launched in 2011, wave 2 launched in 2012, wave 3 launched in 2013 and wave 4 launched in 2014 (see table below).

Table A1: List of HEIs and partnerships

Higher Education Institution	Wave	Partnership
Greater Manchester West Cognitive Therapy Training Centre	1	Derby
	1	Manchester and Salford
	1	Pennine North
	1	Pennine South
	1	Barnsley
	2	Bolton
	2	Central Lancashire
	2	North Lancashire
	3	Ashton, Wigan and Leigh
	3	Cheshire and Wirral
	3	Liverpool
	3	North Staffordshire
	3	Trafford
	4	South Staffordshire
	4	Cumbria
	4	Sefton
4	Bradford	
4	St. Helens, Knowsley, Warrington and Halton	
University College London & King’s College London	1	Hertfordshire
	1	West Sussex
	1	Westminster
	1	Haringey
	1	Cambridge
	1	Wandsworth
	1	Greenwich
	1	Lambeth and Southwark
	2	Waltham Forest

Table A1: List of HEIs and partnerships (continued)

Higher Education Institution	Wave	Partnership
University College London & King's College London	2	Hackney
	2	Croydon
	2	Camden
	2	Richmond
	2	Islington
	2	Tower Hamlets
	2	Bromley
	3	Newham
	3	Barking and Dagenham
	3	Redbridge
	3	Lewisham
	3	Ealing
	3	Barnet
	3	East Sussex
	3	Merton
	4	North Essex
	4	Brent and Harrow
	4	Hounslow, Hammersmith and Fulham
4	Bexley	
University of Reading	1	Oxford and Buckinghamshire
	1	Wiltshire, Bath and North East Somerset
	1	Gloucestershire
	1	Bournemouth, Poole and Dorset
	1	Swindon
	2	Berkshire
	2	Bedfordshire
	2	Luton
	2	Kensington and Chelsea
	3	Hampshire
	3	Nottinghamshire
	3	Kingston
	3	Birmingham
	4	Worcestershire
Northumbria University	2	Stockton and Redcar and Cleveland
	2	North and South Durham
	2	Hambleton and Richmondshire
	2	Darlington
	2	Rotherham
	2	Doncaster
	3	Hartlepool, Middlesbrough and Easington
	3	Scarborough and Harrogate
	3	Sheffield
	3	Newcastle

Table A1: List of HEIs and partnerships (continued)

Higher Education Institution	Wave	Partnership
Northumbria University	3	Leeds
	4	Sunderland
	4	Gateshead
	4	Humber
University of Exeter	2	Devon
	2	Plymouth
	2	Torbay
	3	Somerset
	3	Cornwall
	3	Bristol and South Gloucestershire
	4	Herefordshire
	4	North Somerset

The programme drew on a model of implementation based on one devised by Fixsen and colleagues (2005; 2009; 2011). The premise was to train a selection of practitioners, supervisors and managers, alongside providing additional resources for infrastructure and building regional and national collaborations to support best practice. In this way, the aim was to maximise limited resources and facilitate the embedding of sustainability.

The programme sought to embed seven principles in services:

1. **To support whole service transformation through leadership**
2. **To improve access through self-referral**
3. **To work in partnership with the young person and their parent/carer in service delivery and design**
4. **To deliver evidence-based psychological treatments**
5. **To deliver outcomes-focused psychological treatments**
6. **To work in partnership with the young person and their parent/carer throughout treatment**
7. **To provide supervision to support the delivery of evidence-based, service user-informed, and outcomes-informed practice.**

To embed these principles, the programme was described as:

- Working in partnership with children and young people to shape their local services and supporting local services participating in the programme to do likewise.
- Supporting services to develop a culture of reflective practice and accountability.
- Improving the workforce through training in best evidence-based practice.
- Developing mechanisms to deliver frequent/session-by-session outcome monitoring to help the therapist and service user work together in their session, and to help supervisors support therapists in improving outcomes.
- Supporting local areas in improving the infrastructure they use to collect and analyse data to see if children and young people are getting better.

Specific training programmes were developed for both practice and supervision in cognitive behavioural therapy (CBT) for anxiety; parent training for behavioural difficulties in children under the age of 9; systemic family therapy for eating disorders, conduct disorders and depression; interpersonal psychotherapy for adolescent depression; leadership, service development, supervision skills and service transformation skills; and enhanced evidence-based practice (see tables A2 and A3). In the first instance YoungMinds and then the GIFT Consortium worked with the programme to ensure that young people’s participation was embedded within the sites taking part. Subsequently YoungMinds supported parents to participate in the programme. This was with

the intention of ensuring that sites were committed to improving services based on feedback provided by children, young people and their parents/carers. A best practice curriculum was developed, refined and delivered in a CYP IAPT Curriculum (NHS England, 2013). Model service specifications for targeted and specialist children and young people’s mental health services (Tier 2/3) were developed for commissioners wishing to commission using CYP IAPT principles and values. In particular, best practice standards for service delivery were codified in *Delivering With, Delivering Well* (CYP IAPT, CORC & CAMHS providers, 2014). A major emphasis was placed on ensuring service

user feedback was at the heart of delivery, and best practice guidance was provided on how best to use child- and parent-report outcome and feedback tools, such as in the Guide to Using Outcomes and Feedback Tools with Children and Young People (Law & Wolpert, 2014). Service Development and Accreditation subgroups were developed that included quality and accreditation networks, young people and service representatives, which now underpin Care Quality Commission work and are incorporated into service standards.

Tables A2 and A3 below give a breakdown of staff trained by year.

Table A2: Training in evidence-based practice by year

Year	CBT	Parenting	SFP (ED)	SFP (DEP & CD)	IPT-A	EEBP	Total
2011/12	64/19	33/11	-	-	-	-	97/30
2012/13	99/33	43/12	-	-	-	-	142/45
2013/14	99/29	56/23	23/12	45/17	29/7	-	252/88
2014/15	100/24	61/19	18/9	58/22	21/12	114	372/86
Total	362/105	193/65	41/21	103/39	50/19	114	

Notes: Counts include either trainee practitioners or supervisors in the specified training modality.

Table A3: Leadership and transformation training programme by year

Year	Leadership trainees
2011/12	35
2012/13	48
2013/14	73
2014/15	51
Total	207

Appendix B

Outcomes and evaluation group

The outcomes and evaluation task group oversaw measure selection and approach to data collection; chose measures based on review of psychometric properties, feasibility, utility, compatibility and cost; advised on how to

implement routine outcome measures and how to report findings; consulted with wider networks and held regular public consultations on measures to include in the dataset.

Membership 2011

1. Miranda Wolpert (Chair)
2. David Clark
3. Margaret Oates
4. David Wells
5. Bill Badham
6. Duncan Law
7. Margaret Murphy
8. Jessica Deighton
9. Ann York
10. Amandeep Hothi
11. Kathryn Pugh
12. Paul Wilkinson
13. Claire Maguire
14. Damian Hart
15. Paul Stallard
16. Stephen Scott
17. Andy Fugard
18. Kevin Mullin
19. Raphael Kelvin

Membership 2015

1. Miranda Wolpert (Chair)
2. Duncan Law
3. Cathy Troupp
4. David Trickey
5. Margaret Murphy
6. Cathy Street
7. Barbara Rayment
8. Margaret Oates
9. Ro Rossiter
10. Anne O’Herlihy
11. David Clark
12. Paul Stallard
13. Stephen Scott
14. Philippe Mandin
15. Rabia Malik
16. Anne York
17. Emma Morris
18. Peter Stratton
19. Jessica Deighton
20. Jenna Jacob

Table A4 outlines the full list of child- and parent-reported measures that had been agreed by December 2015.

Table A4: List of measures used by services

	Full title of measure (listed alphabetically)	Acronym	Subscales	Child version	Parent version
1	Brief Parental Self Efficacy Scale	BPSES	–		Y
2	Child/Outcome Rating Scale	C/ORS	–	Y	
3	Goal Based Outcomes	GBO	–	Y	Y
4	Clinical Outcomes in Routine Evaluation	CORE-10	–	Y	
5	Eating Disorder Examination – Adolescent	EDE-A	–	Y	
6	Eating Disorder Examination – Questionnaire	EDE-Q	–	Y	
7	Experience of Service Questionnaire	ESQ	–	Y	Y
8	Generalised Anxiety Disorder	GAD-7	–	Y	Y
9	Impact of Events Scale	IES	–	Y	
10	Me and My School (Behavioural subscale only)	M&MS	–	Y	
11	Oppositional Defiance Disorder – parent rated	ODDp	–		Y
12	Patient Health Questionnaire	PHQ-9	–	Y	
13	Revised Child Anxiety and depression Scale (Total anxiety and depression)	RCADS	Generalised anxiety Separation anxiety Depression Social anxiety Panic Obsessive compulsive disorder (OCD) Total anxiety	Y	Y

Table A4: List of measures used by services (continued)

	Full title of measure (listed alphabetically)	Acronym	Subscales	Child version	Parent version
14	Sheffield Learning Disabilities Outcome Measure	SLDOM	–		Y
15	Short/Warwick Edinburgh Mental Wellbeing Scale	S/WEMWBS	–	Y	
16	SCORE-15	SCORE-15	–	Y	Y
17	Session Rating Scale (Group and Child versions)	SRS (GRS/CRS)	–	Y	Y
18	Strengths and Difficulties Questionnaire (Total Difficulties)	SDQ	Emotional problems Hyperactivity Conduct problems Peer problems Prosocial behaviour Impact on life	Y	Y
19	Session Feedback Questionnaire	SFQ	–	Y	Y
20	Session by Session (previously known as RMQ; Regular Monitoring Questionnaire)	SxS (RMQ)	–	Y	Y
21	Young Person's Clinical Outcomes in Routine Evaluation	YP CORE	–	Y	

Table A5: Timeline of measures added to the dataset

Measure	Date added to dataset	Included in data completeness target	
		Yes (✓)/No(X)	From:
SDQ (child and parent)	April 2011	✓	April 2011
RCADS (child and parent)	April 2011	✓	April 2011
Session by session subscales of RCADS	April 2011	✓	April 2011
Impact of Events Scale (child)	April 2011	✓	April 2011
Me and My School (child)	April 2011	✓	April 2011
Oppositional Defiance Disorder (parent)	April 2011	✓	April 2011
GAD-7 (child)	December 2011	✓	December 2011
PHQ-9 (child)	December 2011	✓	December 2011
Goals (child and parent)	December 2011	✓	December 2011
CORS (child and parent)	December 2011	✓	June 2013
SRF (child and parent)	December 2011	X	Not included
SFQ (child and parent)	December 2011	X	Not included
ESQ (child and parent)	December 2011	X	Not included
CGAS (practitioner)	December 2011	X	Not included
HoNOSCA (clinician)	December 2011	X	Not included
SxS (child and parent)	February 2012	✓	February 2012
Current view (clinician)	December 2012	✓	December 2012
SCORE 15 (family)	June 2013	✓	June 2013
BPSES (parent)	June 2013	X	Not included
SLDOM (parent)	June 2013	X	Not included
SWEMWBS (child)	June 2013	✓	June 2013
EDE-Q (child/clinician)	June 2013	✓	June 2013
ORS (child)	June 2013	✓	June 2013
YCORS (child)	June 2013	X	Not included
Kessler-10 (parent)	June 2013	X	Not included
GSRS (child and parent)	June 2013	X	Not included
YP CORE (child)	June 2013	✓	June 2013

Appendix C

Table A6: KPI summary table

No.	KPI	Comments	Sample used as denominator
1a and 1b	Movement towards recovery (Recovery Units and standardised effect size of treatment)	1a) As previously agreed, analysis not run as there is a danger of misinterpretation 1b) All scales showed positive effect sizes for both child and parent-reported scales (range was 0.48 to 1.38)	Paired outcomes for above threshold closed treatment sample 6
2	Number of children/young people who had a problem with EET attendance at the start of treatment and the extent of the problem decreased by the end of treatment in reporting period	Of those who indicated they had difficulties in the areas at the start of treatment, 55% (565/1,025) of children had improved in terms of attendance and 53% (604/1,138) had improved in terms of attainment	Derived from closed treatment cases sample 3, but only considering cases with difficulties rated mild or above in the relevant EET item at a first time point
3a	Number of periods of contact showing recovery in reporting period	36% (2,114/5,918) of children and 28% of parents (1,046/3,699) with the relevant criteria reported recovery at the follow-up point	Paired outcomes for above threshold closed treatment sample 6
4	Numbers of children/young people seen by treatment type in reporting period	Just over half of the total sample (65%; 62,865/96,325) did not have type of therapy recorded. The most frequent type of defined therapies were Cognitive Behaviour Therapy (17%) followed by Multimodal Integrated therapy (11%), Family Systemic Therapy (8%) and Parent Training (7%). The most commonly-recorded type overall was "other therapy" (19%)	Full sample 1
5	Number of self-referral as a proportion of total referrals in reporting period	Out of all the cases with referral source data in the dataset (66,295), 4% (2,709) were recorded as being self-referrals	Full sample 1
6a and 6b	Proportion of children or young people who finished treatment who said they were satisfied and proportion of parents who said they were satisfied with the treatment their child had received	6a) 14% (3,196/23,373) of closed cases had child reported satisfaction data. Of those, 82% (2,629) recorded "very satisfied" for their overall care. 6b) 11% (2,698/23,373) of closed cases had parent reported satisfaction. Of those, 88% (2,382) recorded "very satisfied" for their overall care.	Closed treatment cases sample 3

Table A6: KPI summary table (continued)

No.	KPI	Comments	Sample used as denominator
7	Number of therapists, supervisors and managers in training	Between waves 1 to 4, a total of 863 practitioners, 248 supervisors and 207 managers received training	Full sample 1
8a	Data completeness of paired outcome scores. Based on closed cases seen for treatment and at least three recorded events in the reporting period [symptom measure/impact measure] ²	50% (7,808/15,536) had paired outcome information from symptom or impact tracking normed scales from either child or parent perspective	Above threshold closed treatment sample 5
8b	Data completeness of paired outcome scores. Based on closed cases seen for treatment and at least 3 recorded events in the reporting period [symptom measure only] ³	35% (5,437/15,536) of closed cases had paired outcome information from symptom tracking scales from either child or parent perspective	Above threshold closed treatment sample 5
9a	Proportion of children or young people where clinical tools used to monitor progress against goals	12% (2,784/23,373) of cases had paired goal information from the child perspective and 3% from the parent perspective	Closed treatment cases sample 3
9b and 9c	Proportion of children or young people where clinical tools used to monitor progress against general wellbeing ⁴ and proportion of children or young people where some evidence of service user feedback collected ⁵	6% (923/15,536) of cases had parent wellbeing information 79% (18,411/23,373) of cases had service feedback information at case closure	Above threshold closed treatment sample 5
10	Proportion of young people who have information on ethnicity	71% (68,663/96,325) of cases had ethnicity information	Full sample 1

² Measures included are: SDQ Hyperactivity, SDQ Conduct, SDQ Emotional, SDQ Impact sub-scale, SDQ Impact sub-scale followed by RMQ (SxS), RMQ (SxS), RCADS sub-scales, PHQ-9, GAD-7, IES, M&MS, YP CORE, CORE-10, CORS, OR.

³ Measures included are: RCADS sub-scales, PHQ-9, GAD-7, IES, M&MS, YP CORE, CORE-10, CORS

⁴ Measures included are: RCADS sub-scales, PHQ-9, GAD-7, IES, M&MS, YP CORE, CORE-10, CORS, ORS.

⁵ Measures included are: ESQ, SFQ, C/G/SRS

Appendix D

Table A7: Thresholds and reliable change indices for child- and parent-reported outcome measures

Clinical Thresholds and Reliable Change Indices for Child-reported Measures			
Child-reported Measure	Clinical Threshold	Source	Reliable Change Index in the CYP IAPT data**
CORE-10	11	Barkham et al., 2013	6.27
CORS*	32	Duncan et al., 2006	10
GAD-7	8	Spitzer et al., 2006; also recommended in the IAPT Data Handbook	4.22
IES	17	Threshold reported in this document: http://em-drresearchfoundation.org/toolkit/cries-13.pdf	11.92
M&MS	6	Approx. equivalent to top 10% of population; Deighton et al., 2013	2.82
ORS*	28	Duncan et al., 2006	6.55
PHQ-9	10	Kroenke et al., 2001; also recommended in the IAPT Data Handbook	5.99
RCADS Depression (T-score)	66	Approx. equivalent to top 6% of population; T-scores based on Chorpita et al., 2000	18.01
RCADS GAD (T-score)	66	Approx. equivalent to top 6% of population; T-scores based on Chorpita et al., 2000	14.18
RCADS OCD (T-score)	66	Approx. equivalent to top 6% of population; T-scores based on Chorpita et al., 2000	16.96
RCADS Panic (T-score)	66	Approx. equivalent to top 6% of population; T-scores based on Chorpita et al., 2000	19.34
RCADS Separation Anxiety (T-score)	66	Approx. equivalent to top 6% of population; T-scores based on Chorpita et al., 2000	24.03
RCADS Social Phobia (T-score)	66	Approx. equivalent to top 6% of population; T-scores based on Chorpita et al., 2000	13.96
SxS (RMQ)	2	Approx. equivalent to top 10% of population; Categories 'High' and above reported in: http://www.sdqinfo.com/py/sdqinfo/c0.py	3.27
SDQ Conduct Problems	5	Approx. equivalent to top 10% of population; Categories 'High' and above reported in: http://www.sdqinfo.com/py/sdqinfo/c0.py	3.74
SDQ Emotional Problems	6	Approx. equivalent to top 10% of population; Categories 'High' and above reported in: http://www.sdqinfo.com/py/sdqinfo/c0.py	4.26
SDQ Hyperactivity	7	Approx. equivalent to top 10% of population; Categories 'High' and above reported in: http://www.sdqinfo.com/py/sdqinfo/c0.py	3.87
SDQ Total Impact	2	Approx. equivalent to top 10% of population; Categories 'High' and above reported in: http://www.sdqinfo.com/py/sdqinfo/c0.py	3.24
YP-CORE	14	'All' category reported in Twigg et al., 2015	10.77

* higher score = less severity; in all other instances, the reverse applies. ** This is the amount scores have to change between a first and a last time point for the change to be considered reliable, based on the CYP IAPT data.

Clinical Thresholds and Reliable Change Indices for Parent-reported Measures

Child-reported Measure	Clinical Threshold	Source	Reliable Change Index in the CYP IAPT data**
ORS*	28	Duncan et al., 2006	6.42
RCADS Depression (T-score)	66	Approx. equivalent to top 6% of population; T-scores based on Chorpita et al., 2000	23.07
RCADS GAD (T-score)	66	Approx. equivalent to top 6% of population; T-scores based on Chorpita et al., 2000	17.63
RCADS OCD (T-score)	66	Approx. equivalent to top 6% of population; T-scores based on Chorpita et al., 2000	22.48
RCADS Panic (T-score)	66	Approx. equivalent to top 6% of population; T-scores based on Chorpita et al., 2000	40.25
RCADS Separation Anxiety (T-score)	66	Approx. equivalent to top 6% of population; T-scores based on Chorpita et al., 2000	27.62
RCADS Social Phobia (T-score)	66	Approx. equivalent to top 6% of population; T-scores based on Chorpita et al., 2000	16.61
SxS (RMQ)	2	Approx. equivalent to top 10% of population; Categories 'High' and above reported in: http://www.sdqinfo.com/py/sdqinfo/c0.py	3.16
SDQ Conduct Problems	5	Approx. equivalent to top 10% of population; Categories 'High' and above reported in: http://www.sdqinfo.com/py/sdqinfo/c0.py	4.33
SDQ Emotional Problems	6	Approx. equivalent to top 10% of population; Categories 'High' and above reported in: http://www.sdqinfo.com/py/sdqinfo/c0.py	4.39
SDQ Hyperactivity	7	Approx. equivalent to top 10% of population; Categories 'High' and above reported in: http://www.sdqinfo.com/py/sdqinfo/c0.py	3.82
SDQ Total Impact	2	Approx. equivalent to top 10% of population; Categories 'High' and above reported in: http://www.sdqinfo.com/py/sdqinfo/c0.py	3.17

* higher score = less severity; in all other instances, the reverse applies. ** This is the amount scores have to change between a first and a last time point for the change to be considered reliable, based on the CYP IAPT data.

Appendix E

Current View: development and algorithm

Key information below is reproduced with permission from Wolpert et al., 2015b.

Current View development

'The Current view tool' (see page 82) is a data collection tool initially developed by the Children and Young People's Improving Access to Psychological Therapies (CYP IAPT) programme. It was subsequently expanded by the CAMHS Payment System Project Group to cover a wider range of presenting problems, complexity factors and contextual problems affecting children and young people (CAMHS EBPU, 2011). Detailed guidance on completing the form was published in 2013 (Jones et al., 2013).

Development of the complexity factors section of the tool was informed by consultation events organised by the Project Group in May 2012. Events were held in Leeds and London and were attended by 34 and 57 people respectively. Attendees consisted of a mix of clinicians and non-clinicians. At each event a group exercise was conducted in which participants were asked to rank 12 factors in terms of how much they might increase the amount of direct and indirect work needed to help a child or young person and their family manage a mental health problem. The 12 factors were: Looked after Child; Current protection plan; Learning disability; Pervasive developmental disorder (including Autism and Asperger's); Youth offending; Experience of war, torture or trafficking; Serious physical health issues (including Chronic Fatigue); Deemed "child in need" of social service input; Refugee or asylum seeker; Young carer status; Neurological issues (such as tics or Tourette's); Access issues (difficulties travelling to sessions). 'Access issues' was ranked lowest at both events and was not taken forward for inclusion in the complexity factors section of the Current View form. However, it is captured in a different way as part of the 'Service Engagement' contextual problem on the form (CAMHS EBPU, 2011).

An opportunity was provided in the group exercises to suggest additional factors that were not already included in the exercise. Parental health was the most frequently mentioned overall (by 6 delegates in Leeds and 14 delegates in London) and was subsequently added to the tool. Deprivation was suggested by three delegates in London. As social deprivation is known to be associated with higher rates of mental health difficulties and thought to link with potential difficulties in accessing services, a

complexity factor named 'Living in financial difficulty' was added to the tool as a proxy for social deprivation.

The 30 provisional problem descriptions were developed from consultation with expert clinicians, including widespread consultation with both professionals as part of the CYP IAPT Outcomes and Evaluation Group (OEG) and young people via the YoungMinds Very Important Kids (VIK) group.

Preliminary analysis of the Current View tool's psychometric properties has been undertaken by Dr Andy Whale, Dr Amy Macdougall and Dr Peter Martin, with mixed results and further work is needed. See the full report (Wolpert et al., 2015b) for further detail.

The needs-based grouping algorithm

The algorithm is based on 31 ratings from the Current View Form, namely the thirty presenting problems and the complexity factor "Pervasive developmental disorders". For some groupings, the age of the child or young person (CYP) is also taken into account. All groupings are mutually exclusive, with the exception of grouping NEU (Neurodevelopmental assessment), which may be combined with any of the other groupings. The algorithm never suggests PBP (Presentation suggestive of potential BPD) because we think that this grouping may not be identifiable from presenting information captured on the Current View. The following presents a common-sense explanation of the algorithm. Some details are left out in order to facilitate a conceptual understanding. For a full specification, please see Appendix A of Wolpert et al., 2015b).

The "NICE groupings" (ADH, AUT, BEH, BIP, DEP, GAP, OCD, PTS, SHA, SOC, EAT, PSY): For the algorithm to suggest one of these groupings, the CYP must have the associated "index problem" rated at least moderate. The index problem is the main symptom associated with a particular NICE clinical guideline [e.g. "Low mood" for DEP ("Depression"), "Eating issues" for EAT ("Eating disorders")]. In addition, the child must NOT have high ratings on a selection of other presenting problems. These "exclusion criteria" are different for each grouping and are based on clinical judgement regarding which kinds of symptoms may mean that the group indicated by the

“index problem” may not be appropriate.

NEU (Neurodevelopmental assessment): If “Unexplained Developmental Difficulties” is rated moderate or severe, the algorithm suggests NEU. Note that a child can be in NEU at the same time as being in any of the other groupings.

If the CYP does not have any of the index problems identifying the “NICE groupings”, or if the presence of exclusion criteria means that no “NICE grouping” could be suggested, then there are five further possibilities:

ADV Getting advice: Signposting and self-management advice: If the CYP presents with mild symptoms only – or has at most one moderate problem, but this ‘moderate problem’ is none of the “index” problems associated with the “NICE Groupings” that we have referred to above – then the algorithm suggests ADV.

EMO Co-occurring emotional difficulties: If the CYP has two or more “emotional problems” (anxieties and/or depression) rated at least moderate, the algorithm will suggest EMO. There are some additional exclusion criteria; for details see Table C1b in Appendix C of Wolpert et al., 2015b.

BEM Co-occurring behavioural and emotional difficulties: If the CYP has “Behavioural Difficulties” and one or more emotional problems (anxiety/anxieties or depression) rated at least moderate, the algorithm will suggest BEM. There are some additional exclusion criteria; for details see Table C1b in Appendix C of Wolpert et al., 2015b.

DNC Difficulties not covered by other groupings: If the CYP has either two or more problems rated as moderate, or a single problem rated as severe, but doesn’t fit into any of the groupings described above (except NEU), then the algorithm will suggest DNC. Some additional criteria apply; for details see Table C1b in Appendix C of Wolpert et al., 2015b.

DSI Replace with severe impact: The algorithm will suggest DSI if the CYP has either two or more problems rated as severe, or has either Delusional Beliefs or Eating Issues rated as moderate or severe, or Extremes of Mood rated as severe, but doesn’t fit the criteria for either grouping EAT or PSY.

Note that in cases where DSI is suggested by the algorithm and Delusional Beliefs, Eating Issues or Extremes of Mood are rated as moderate or severe, a co-occurring problem has meant that a “NICE grouping” has not been suggested.

Figure A1: The Current View tool

Current View

CYP Name

Practitioner's Name

Please indicate your reason for completing this form:

DOB:

Practitioner's ID

First Contact

NHS ID:

Service Allocated Case Id

Changed Situation

Date: / / 20

Time: h m

Changed Understanding

	PROVISIONAL PROBLEM DESCRIPTION RATING NEED NOT IMPLY A DIAGNOSIS	NONE	MILD	Moderate	SEVERE	NOT KNOWN
1	Anxious away from caregivers (Separation anxiety)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	Anxious in social situations (Social anxiety/phobia)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	Anxious generally (Generalized anxiety)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4	Compelled to do or think things (OCD)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5	Panics (Panic disorder)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6	Avoids going out (Agoraphobia)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7	Avoids specific things (Specific phobia)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8	Repetitive problematic behaviours (Habit problems)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9	Depression/low mood (Depression)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10	Self-Harm (Self injury or self-harm)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11	Extremes of mood (Bipolar disorder)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12	Delusional beliefs and hallucinations (Psychosis)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13	Drug and alcohol difficulties (Substance abuse)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14	Difficulties sitting still or concentrating (ADHD/Hyperactivity)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15	Behavioural difficulties (CD or ODD)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16	Poses risk to others	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17	Carer management of CYP behaviour (e.g., management of child)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18	Doesn't get to toilet in time (Elimination problems)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19	Disturbed by traumatic event (PTSD)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20	Eating issues (Anorexia/Bulimia)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21	Family relationship difficulties	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
22	Problems in attachment to parent/carers (Attachment problems)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23	Peer relationship difficulties	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24	Persistent difficulties managing relationships with others (includes emerging personality disorder)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25	Does not speak (Selective mutism)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26	Gender discomfort issues (Gender identity disorder)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27	Unexplained physical symptoms	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28	Unexplained developmental difficulties	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
29	Self-care Issues (includes medical care management, obesity)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
30	Adjustment to health issues	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

	SELECTED COMPLEXITY FACTORS	Yes	No	NOT KNOWN
1	Looked after child	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	Young carer status	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	Learning disability	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4	Serious physical health issues (including chronic fatigue)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5	Pervasive Developmental Disorders (Autism/Asperger's)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6	Neurological issues (e.g. Tics or Tourette's)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7	Current protection plan	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8	Deemed "child in need" of social service input	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9	Refugee or asylum seeker	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10	Experience of war, torture or trafficking	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11	Experience of abuse or neglect	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12	Parental health issues	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13	Contact with Youth Justice System	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14	Living in financial difficulty	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

CONTEXTUAL PROBLEMS					
	None	Mild	Moderate	Severe	Not known
HOME	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
SCHOOL, WORK OR TRAINING	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
COMMUNITY	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
SERVICE ENGAGEMENT	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
EDUCATION/EMPLOYMENT/TRAINING					
ATTENDANCE DIFFICULTIES	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
ATTAINMENT DIFFICULTIES	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

GUIDANCE ON IMPACT RATINGS		
These are broad definitions for guidance and should be considered within an age-appropriate context. Each Problem should be rated independently If Functioning and Distress levels differ, then select the higher rating		
FUNCTIONING	DISTRESS	
NONE	There may be transient difficulties and 'everyday' worries that occasionally get out of hand (e.g. mild anxiety associated with an important exam, occasional 'blow-ups' with siblings, parents or peers) but CYP is generally secure and functioning well in all areas (at home, at school, and with peers).	No distress or noticeable difficulties in relation to this problem.
MILD	Symptoms cause occasional disruption but do not undermine functioning and impact is only in a single context . All/most age appropriate activities could be completed given the opportunity. The CYP may have some meaningful interpersonal relationships.	Distress may be situational and/or occurs irregularly less than once a week. Most people who do not know the CYP well would not consider him/her to have problems but those who do know him/her well might express concern.
MODERATE	Functioning is impaired in at least one context but may be variable with sporadic difficulties or symptoms in several but not all domains.	Distress occurs on most days in a week . The problem would be apparent to those who encounter the child in a relevant setting or time but not to those who see the child in other settings.
SEVERE	CYP is completely unable to participate age-appropriately in daily activities in at least one domain and may even be unable to function in all domains (e.g. stays at home or in bed all day without taking part in social activities, needing constant supervision due to level of difficulties).	Distress is extreme and constant on a daily basis . It would be clear to anyone that there is a problem.

DEFINITIONS OF CONTEXTUAL PROBLEMS
These definitions are for general guidance purposes only and should be considered within an age-appropriate context and with reference to cultural norms where appropriate. The examples given are not exhaustive.

- HOME**
Problems in the home environment that are external to the CYP (e.g. crowded housing, homelessness, lack of social support network).
- SCHOOL, WORK OR TRAINING**
Problems in the school, work or training environment that are external to the CYP (e.g. difficulties in communications between home and school, multiple changes of teacher, breakdown in relations between teacher(s) and CYP/family).
- COMMUNITY**
Problems in the community that are external to the CYP (e.g. street violence, gang intimidation, racial discrimination and difficulties with neighbours).
- SERVICE ENGAGEMENT**
This refers to difficulties regulating the appropriate level of service engagement. This may include history of multiple or fractured contact with services, difficulties locating care records, difficulties accessing the service and problems engaging the CYP and their family appropriately.

DEFINITIONS OF ATTENDANCE AND ATTAINMENT DIFFICULTIES
These definitions are for general guidance purposes only and should be considered within an age-appropriate context and with reference to cultural norms where appropriate. They should also be considered with specific reference to the CYP you're working with (e.g. if the CYP has a learning disability, attendance and attainment should be considered in relation to peers of the same developmental rather than chronological age). The examples given are not exhaustive.

ATTENDANCE DIFFICULTIES	
NONE	No problems noted. As rough guidance, around 1-2 days absence from school per month should be considered as within normal limits.
MILD	Some definite problems. The CYP may be attending part-time or missing several lessons (includes truanting, school refusal or suspension for any cause). As a rough guidance, 1 day of absence per week might be considered here.
MODERATE	Marked problems. The CYP may be attending infrequently, or is at high risk of exclusion or dismissal. As a rough guidance, the child may be absent 2 days per week.
SEVERE	CYP is out of school the majority of the time (for reasons of truancy, exclusion or refusal) or may be in a Pupil Referral Unit, expelled or not in Education, Employment or Training.
ATTAINMENT DIFFICULTIES	
NONE	No problems noted. The CYP will be attaining at the optimum age-appropriate level moderated by that expected for their known abilities.
MILD	Some problems. For example, if the CYP is in school they may be well below the year level in at least one subject, or have problems with work rate or timekeeping if in employment or training.
MODERATE	Significant problems. If at school they may fail key exams, or be below the year group in all subjects. If in employment, they may have received formal warnings about their performance and/or behaviour.
SEVERE	CYP has dropped out of education, employment or training.

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 hold 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.

Contact:
CORC@annafreud.org
www.corc.uk.net

