

## **“A comparison of parent reported outcome with experience of services”**

### **Structured Abstract:**

**Purpose:** Routine Outcome Monitoring (ROM) is currently seen as a key driver for service improvement at individual, team and service level. This paper explores the relationships between a Patient (parent) Reported Outcome Measure (PROM); a Practitioner reported outcome measure, and a Patient (parent) Reported Experience Measure (PREM).

**Design/methodology/approach:** We recruited a cohort of 302 primary school-age children who were followed for one year from consecutively accepted referrals to three teams within two English Child and Adolescent Mental Health Services (CAMHS). Parents completed the Strengths and Difficulties Questionnaire (a PROM) and Practitioners completed the Clinician Global Assessment Scale at baseline, six and twelve months; parents completed the Experience of Services Questionnaire (a PREM) at 6 and 12 months.

**Findings:** PROM and Practitioner reported outcome measure data suggested poor clinical outcome in terms of symptoms, impact and levels of functioning but were accompanied by PREM evidence of high levels of satisfaction. There was an unexpectedly low correlation ( $<0.2$ ) between both measures of outcome and satisfaction

**Practical implications:** A range of outcome measures are required to achieve a holistic view of service performance, and single measures may be extremely misleading.

**Originality/Value:** This paper fulfils a need to explore the relationships between different outcome measures to contribute to the understanding of ROM its validity.

**Keywords:** PROMS, PREMS, Service-user satisfaction, ROM, CAMHS

### **Introduction**

Policy makers and commissioners recommend Routine Outcome Monitoring (ROM) as a key method for quality assurance (Department of Health, 2010). ROM can support good practice and innovation while highlighting areas of difficulty at an individual, practitioner, team or service level, particularly for areas where the evidence-base is sparse (Kazdin, 2008). The publication of routinely collected data on post-operative mortality in cardiac surgery may have contributed to a reduction in post-operative mortality (Bridgewater et al, 2007). However, the chronic and fluctuating nature of psychopathology, the need for multiple informants and the impact of school, home and other environments beyond the control of Child and Adolescent Mental Health Service (CAMHS) interventions make the use of ROM particularly complex in relation to children's mental health. The use of standardized outcome measures in any health specialty, however, is also not without controversy, in terms of the availability of appropriately sensitive and focused measures, as well as the potential but yet to be demonstrated influence of case complexity (Bridgewater et al 2007; Norman, 2011; Yates et al 1999). However, in the current economic and policy climate, CAMHS need to be able to demonstrate that the interventions offered have had an impact, and the use of ROM is a major part of the Children and Young People's Increased Access to Psychological Therapy programme (IAPT) (Wolpert et al 2012<sup>a</sup>).

Patient Reported Outcome Measures (PROMs) and Patient Reported Experience Measures (PREMs) have been promoted as preferable alternatives to the measures of process and activity that have previously dominated service assessment (Department of Health, 2008). PROMs assess patients' health status and/or health-related quality of life while PREMs assess patient's satisfaction with services received. Although PROMs and PREMs are referred to as measures of outcome and satisfaction respectively (Ayton et al., 2007) others argue that PREMs are an indicator of service quality as experienced by the patient (Day, et al, 2011). In CAMHS, measures may be completed by different informants, including parents, children, teachers and practitioners, who may have differing opinions about difficulties and functional level of the child or family. The primary target of the service is the child's mental health, but that said, children rarely access the service on their own behalf and may not necessarily agree with their parent's or carer's opinions about the nature and course of their difficulties (Hawley and Weisz, 2005). To make matters even more complex, most children under the age of nine cannot reliably complete most questionnaires (Schwab-Stone et al, 1996).

Little is known about the relationship between PROMs and PREMs. A cross-sectional study of the relationship between satisfaction and outcome in young people in contact with a community CAMHS reported higher levels of satisfaction expressed by parents compared to their children (Barber et al, 2006). Young people who reported higher levels of conduct problems and difficulties with a greater level of impact also reported lower levels of satisfaction with intervention, while there was no link between parental satisfaction and parental reporting's of outcome. Using data collected from more than 7000 families who attended 41 CAMHS via the Child Outcome Research Consortium ([www.corc.uk.net](http://www.corc.uk.net)), Brown and colleagues (2012) demonstrated that responses on the Experiences of Services questionnaire suggested two related areas of satisfaction; the experience of the intervention(s) offered and the environment in which the service was provided. Both factors were subject to strong halo effects, which suggest that parents report a strong overall positive or negative experience. The care-related concept, in particular, was sensitive to differences between less and more satisfied respondents, with strong service level influences.

The current study was a secondary analysis of data from a natural cohort study of primary school aged children attending CAMHS. We aimed to compare PROMS to PREMS collected from the parents of primary school children who were recruited from consecutive referrals to three CAMHS teams and the practitioners working with them. We specifically wanted to examine whether there was a correlation between parent and practitioner reported outcome and with outcome reported by both informants with parental satisfaction at six and twelve months after assessment. We expected that parental satisfaction would correlate highly with outcome, particularly parent-reported outcome.

## **Method**

### **Clinical Setting**

This study took place within two CAMHS in adjacent greater London Boroughs. CAMHS A was the only mental health service for children in Borough A. The practitioners who took part in this study were from the Children's Team, which worked with children who had established psychiatric disorder(s) up to the age of 16 (Tier 3 in the of the Health Advisory Service (1995) model of CAMHS), and an Early Interventions Team, which aimed to work with children who had less complex problems and / or presented with less established difficulties (Tier 2) up to the age of 18. A separate team worked with young people who had emerging severe mental illness who were aged 14-18 years but were not

included in the current study. CAMHS B comprised a single multidisciplinary team for children with established psychiatric disorder, and included specialist sub-teams within it for ADHD, adolescents, and children who were looked-after. In this borough, Tier 2 services were provided elsewhere by a separate team. Both CAMHS served areas that were broadly representative of the British population, with the exception of a higher proportion of children from Black and Ethnic Minority Groups in the area served by CAMHS A and a higher proportion of managerial / professional people with lower proportion of semi-or-un-skilled workers in the area served by CAMHS B (Norman et al, 2014). Neither offered highly specialised Tier 4 services or had an extensive experience of research involvement.

### Participants

Consecutive referrals of children aged between 5 and 10 years 9 months were invited to participate in a study of clinical outcomes; 302 children were recruited (252 from CAMHS A, 50 from CAMHS B). Figure 1 demonstrates the flow of participants through the study. The study was focused on primary school - aged children to limit the range of presentations and the number of informants for a study of outcomes due to resource constraints. Children who were looked after by local authorities were excluded because of anticipated repeated changes in parental responsibility and placement during the course of the study and the resulting difficulty in finding informants who know the child well enough to complete measures. Children referred as an emergency were also excluded because of the obvious ethical and practical difficulty in gaining consent and completing the base line assessment between referral and first assessment. The opt-in rate among eligible families was 62% and 55% actually participated as some children's situation changed so that they required an emergency assessment (n=22 or 48%) or via communication failures when new members of the clinical team booked in multiple new cases without the knowledge of the research team (n=24) so that initial assessment proceeded prior to recruitment. Inability to establish contact by the research team was the commonest reason for non-participation of eligible families.

Insert Figure 1 about here

The characteristics of the sample are described in detail elsewhere (Norman, 2011), but are briefly summarised below (Table 1). The children's mean age was 8 years (S.D = 1.78), and 74% were boys. Most children presented with externalising problems (conduct disorder 55%, attention deficit hyperactivity disorder 35%, autism spectrum condition 14%, emotional disorders 31%, tics, eating and other disorders 17%); the proportions add to greater than 100% as many children had more than one

type of disorder. Sadly we were not permitted to access anonymised data that illustrate the proportions of children with emergency presentations, or who were children looked after or over the age of 11 years in terms of all referrals to the clinic to see how it compared with our data. The differences between participants and those who did not participate in Table 1 appear to be closely related to the planned exclusion criteria and thus to relate to eligibility rather than differences in our achieved sample among those who were eligible. Children who did not participate were older and were more likely to be referred by social services. Over the three data points there was a loss of practitioner reported data due to drop out or discharge of children from CAMHS; practitioners could only report on the function of children who had attended a least one appointment since the last data point (see Figure 1 and Table 2).

Place Table 1 here

### Measures

Parents completed two measures, the Strengths and Difficulties questionnaire (SDQ; a PROM, Goodman, 2001) at baseline and at six and twelve month follow up and the Experience of Service Questionnaire, (ESQ, a PREM, Attridge-Stirling, 2002; Brown et al 2012) at both follow ups. The Practitioners completed the Children's Global Assessment Scale, (CGAS, Shaffer et al, 1983) at baseline and six-monthly while the child attended CAMHS. These measures were selected either because they were recommended by the external working group for the National Service Framework for children and maternity services (Department of Health, 2003b: Outcomes subgroup of the child and adolescent mental health working group, 2003), because they were part of the Child Outcome Research Consortium protocol ([www.corc.uk.net](http://www.corc.uk.net)) and / or because they have been proved to be feasible for use in routine clinical practice (see [www.corc.uk.net](http://www.corc.uk.net); Slade et al., 1999).

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

The SDQ is a well-validated, widely used 25-item questionnaire composed of five scales that assess conduct problems, inattention-hyperactivity, emotional symptoms, peer problems and pro-social behaviour (Goodman, 2001). Scores in the abnormal range (>90<sup>th</sup> centile) are associated with a nearly 16 times increase in the likelihood that the child has a psychiatric disorder. The measure is reliable as judged by internal consistency (Cronbach's  $\alpha$  0.73), cross informant correlation (mean 0.34) and retest stability (mean correlation 0.62) (Goodman, 2001). The SDQ total difficulties score is generated by

summing all the scales except the prosocial scale. The SDQ also has an impact supplement, of which the items on overall distress and social impairment can be summed to generate an SDQ impact score.

#### *The SDQ Added Value Algorithm*

The SDQ Added Value algorithm is an equation generated (see [www.sdqinfo.com](http://www.sdqinfo.com); Ford et al, 2009) from data from the British Child and Adolescent Mental Health Survey 2004 (Green et al, 2004) on children with a psychiatric disorder in the baseline survey and / or whose parents had contacted primary health care or a teacher with concerns about their child's mental health. It was derived empirically from the parental SDQ scores and designed to estimate the difference between the expected and the observed scores on a parental SDQ at follow up, based on baseline scores, using the epidemiological sample as a proxy control group. A positive SDQ Added Values score (SDQ AVS) indicates that the children have better outcomes than would have been predicted by the algorithm, while a negative score suggests that the child is doing worse than expected. A score of 0 indicates no difference.

The SDQ AVS is calculated from the following formula (see [www.sdqinfo.org](http://www.sdqinfo.org)):

$$\text{SDQ Added Value} = 2.3 + 0.8 * T1\text{Total} + 0.2 * T1\text{Impact} - 0.3 * T1\text{Emotion} - T2\text{Total}.$$

T1Total= Time 1 SDQ Total Difficulties score

T1Impact= Time 1 Impact score

T1Emotion= Time 1 Emotional SDQ subscale

T2Total = Time 2 SDQ Total Difficulties score

Two studies have tested the SDQ AVS algorithm against data from the control and intervention arms of randomised controlled trials separately. They have demonstrated that the score does approximate to expected effect sizes; that is an effect size of 0 in the control group who received no intervention and for whom no change would be expected while the intervention group demonstrated an SDQ Added Value Score that approximated to the effect size detected for the intervention in the original trial analysis (Ford et al, 2009; Rotheray et al, 2014). In both these studies, simple difference / change scores provided large effect sizes that greatly exceed that suggested by the SDQ AVS as the influence of attenuation, random fluctuation, and regression to the mean were not controlled. The use of the SDQ AVS provided a proxy control group because of ethical and practical constraints of studying a control group of children who required referral to CAMHS.

### *The Experience of Services Questionnaire (ESQ)*

The ESQ was developed from responses from parents / carers and young people attending a variety of children's services by the Commission for Health Improvement (Attride-Stirling, 2002). It includes 12 quantitative items that ask the participant to rate their agreement on a 4 point Likert scale ranging from "Certainly True" (rated 3), "Partly True" (rated 2), "Not True (rated 1)" to "Don't know (unrated)" in relation to statements about their service experience. As stated above, an analysis of data from 41 CAMHS suggested that the ESQ was able to demonstrate both service level and family level differences in satisfaction, particularly in relation to dissatisfaction (Brown et al, 2012).

### *The Children's Global Assessment Scale (CGAS)*

The CGAS provides an estimate of the level of function for children aged 4-16 along a continuous scale of 1-100. It has high levels of inter-rater reliability (intraclass correlation coefficient 0.84), test-retest stability (intraclass correlation coefficients 0.69-0.95), and discriminant validity ( $p=0.001$  for difference between mean scores for outpatients and inpatients), and can detect clinically significant change (Shaffer et al 1983). The denominators decreased in the CGAS at each time point as the clinicians were only asked to complete it if the child was seen by CAMHS within the previous six month data collection period, and some children were discharged or dropped out of treatment during the study.

### Statistical analysis

The statistical analysis was carried out using SPSS 15.0. The sample size calculation was based on pilot data from a smaller and shorter survey conducted by Byrne et al(1999) that suggested that a sample size of 148 children would be required in order to detect the predicted 2-3 point drop in mean SDQ total difficulties score (equivalent to an effect size of 0.5) with 85% power and a 5% significance level. The larger sample was chosen to allow subgroup analysis among the more common disorders and a nested randomised controlled trial of exposure to a standardised assessment (Ford et al, 2013). The SDQ AVS score was calculated using the equation described above. Pearson's correlations were run to explore the relationship between the parent's satisfaction with the CAMHS service (measured by the ESQ), and reports of clinical outcome according to the parents (total difficulties score and the SDQ impact score) and practitioner (CGAS). Individual items of the ESQ were examined descriptively to explore aspects of service use that parents were particularly satisfied or dissatisfied with, with low levels of satisfaction defined as fewer than 60% of parents reporting "certainly true".

## **Results**

As illustrated in Figure 1 and 2, data were provided by 75% of parents and 82% of practitioners over the three time points.

### **Practitioner reported outcome**

Practitioners reported similar levels of functioning in the children at each time point (Table 2), which suggests no improvement occurred and which may relate to the inevitable loss of data about children who were functioning well enough to be discharged or who dropped out of treatment.

### **Parent reported outcome and satisfaction**

Parents also reported little change in the difficulties or their impact, and the SDQ AVS suggested that children in this sample had worse outcomes at six months than would be predicted by epidemiological data as calculated using the SDQ AVS (CAMHS A mean = -1.43 [95% confidence interval -2.31 to -0.55], and CAMHS B mean = -2.82 [-4.44 to -1.22], see [www.sdqinfo.com](http://www.sdqinfo.com); Ford et al, 2009).

Despite reporting apparently poor outcomes on the SDQ, parents reported high levels of satisfaction on the ESQ (Tables 2 and 3). As Table 2 indicates, correlations between ESQ scores and parent and practitioner reported outcome measures were surprisingly weak.

Place Table 2 here.

Table 3 examines the proportion of parents endorsing each possible response for each item of the ESQ to examine satisfaction in more detail. It is striking that at least 10% responded “don’t know” at time 2 with nearly 20% endorsing this answer at Time 3. Between 51% and 80% reported “certainly true” in response to the 12 positive statements about their child’s care at the six month follow up, while between 41% and 67% expressed this level of satisfaction after 12 months contact with the clinic. Five items were found to have strikingly lower levels of satisfaction than the others; parents were less satisfied that everything had been explained to them, that the practitioner knew how to help them, that staff collaborated together, as well as the comfort of the clinic and the convenience of appointment times. In addition, these items also tended to have a higher proportion who answered “don’t know”, which might indicate reluctance to criticise.



Place Table 3 here.

## **Discussion**

Parents expressed considerable satisfaction with the service that they had received; although five items stand out as having lower levels of endorsement. Parents were less satisfied that everything had been explained to them, or that practitioners knew how to help them or worked together and that they had received an appointment at a convenient time in a comfortable place. The last two of these are more to do with administration than the performance of the clinic and are potentially easy to remedy. The first three may be related to poor communication as this may leave parents unsure as to the practitioners' skills, and may also link to the low levels of agreement between practitioner and parent reports on the outcome measures. The CORC data (Wolpert et al, 2012b) also demonstrated a similar pattern of generally high levels of satisfaction, with less positive responses for receiving adequate explanations, practitioner knowledge and convenient appointments. Although the researchers were independent of the clinics, the data was collected in relation to named individuals, which may explain the high proportion of "don't know" response in our sample. Previous work suggests that satisfaction surveys that are not anonymous produce more positive findings than those where the respondents are not identifiable to those running the service (Stallard and Potter, 1999; [www.corc.uk.net](http://www.corc.uk.net)).

There are several possible explanations for the lack of change in symptoms, impact and functioning. These include a lack of statistical power or methodological issues in the way we attempted to detect change, sample related factors or that our findings reflect poor clinical outcomes. Our preparatory power calculation suggested that the sample recruited should have been amply sufficient to detect an effect of intervention, so lack of power seems unlikely. General measures, such as the SDQ, allow comparison across children of different ages and with different disorders, but are likely to be less sensitive to change than condition specific measures that have more items that relate directly particular difficulties reported by the child and their family (Lee et al, 2005). Practitioners who work with specific groups could consider the use of condition specific measures to allow more detailed and sensitive analysis of the change in specific difficulties alongside more general measures to allow benchmarking. It seems unlikely, however, that our choice of measures completely explains our failure to detect improvement as others using these measures have been able to do so. The mean SDQ AVS for clinics sending data to CORC have been positive (mean=0.19 see [www.CORC.uk.net](http://www.CORC.uk.net), Wolpert et al, 2012b; Fugard et al., 2014), and reanalysis of trial data suggests that the SDQ AVS functions as expected (Ford

et al, 2009; Rotheray et al 2014). Similarly, the smaller study based on one of the services that participated in the current work (Byrne et al. 1999), Australian clinics working secondary school-age children (Matthai et al, 2003) and routine data from an Inner London clinic on parenting (Hurst et al, 2014) detected clinically and statistically significant change over a similar time period using the SDQ.

Statistically significant change may not necessarily indicate clinically meaningful change, and there are various methods for assessing whether the change noted is clinically meaningful or not. Wolpert et al (2014)<sup>a</sup> compared four such methods, which included movement that Crossed a Clinical Threshold (CCT), the Reliable Change Index (RCI), Difference / change Scores (DS) and the SDQ Added Value Score (AVS). They found that the level of agreement between these approaches was only moderate, and recommend the CCT and RCI when considering clinically meaningful change in individual cases and the DS and AVS as more appropriate for reporting on groups of cases. Two studies, as described above (Ford et al, 2009; Rotheray et al, 2014) have demonstrated that DS greatly inflate the estimates of change while the SDQ AVS does adjust for attenuation, random fluctuation and regression to the mean as expected. The SDQ AVS, therefore, does seem to be the most rigorous way to assess change in our study.

Taken together, this supporting literature suggests that for some reason outcomes were poorer than would be expected, which might possibly be related to unusual characteristics in our sample. We restricted recruitment to primary school aged children who were predominantly boys, and who presented with predominantly externalising problems, which are more likely to persist than emotional disorders (MacDermid et al, in submission; Goodman et al, 2002). These were not unexpected findings. Ethical and practical problems meant that we could not have a control group and that we had to exclude children presenting as emergencies or children who were looked after, as we needed to avoid impacting on children's care. Given the sensitivity of attendance at a mental health service we could not approach families that did not make contact with us, so our response rate was moderately low given the multiple stresses that many such families face. The differences between all referrals and participants seem closely related to our exclusion criteria. It seems unlikely that children who are looked after, who are known to have particularly complex psychological problems as well as increased risk factors (Ford et al, 2007), or children in such distress that they required assessment within 24-48 hours would have had better outcomes than those with less acute difficulties who were included, but it remains a possibility. We recruited from three teams within two CAMHS in Greater London. The populations served are fairly representative of the rest of England according to census data, as described above, but it is possible that

three CAMHS, which were selected for pragmatic reasons, may not be representative of other CAMHS. A larger sample of CAMHS would provide greater confidence in the generalisability of findings, and allow a funnel plot analysis to ascertain the extent to which the results from each clinical fall outside the statistically predicted range. Fugard et al (2014) completed such an analysis with administrative data from 51 UK CAMHS, and reported a positive point estimate of mean SDQ AVS from all services combined. They also identified a minority of services that produced SDQ AVS adjusted for the number of cases returned that were above or below either the 95% or 99.5% confidence intervals around this mean value, suggesting that some services are reporting outcomes that are significantly better or worse than expected.

Our findings that symptoms, impact and functioning did not change echo a number of longitudinal observational studies that have failed to detect an association between contact with CAMHS and improved outcomes, (Bachmann et al., 2010; Jörg et al., 2010; Zwaanswijk et al., 2006). All of these results may reflect ineffective services or methodological limitations of studying outcomes using observational methods. Some children who meet psychiatric diagnostic criteria experience chronic and severe difficulties, whilst others have more transient problems (Chaiton et al., 2013; Zwaanswijk et al., 2006), and a small proportion of children with disorders even appear to deteriorate over the course of treatment (Warren et al, 2010). Angold et al (2000) observed a positive effect of service use on symptoms only after controlling for pre-treatment symptom trajectory in the multi-wave Great Smoky Mountains Study. Children who later accessed CAMHS demonstrated higher levels of psychopathology and a deteriorating trajectory compared to children who did not contact services. Children seen in clinical practice often have more complex and severe disorders, greater levels of comorbidity as well as more adverse psycho- social backgrounds compared to participants in intervention trials who may be selected to avoid comorbidity and are often recruited from non-clinical environments (Weisz et al., 2013). This leads to heated debates about the extent to which evidence-based interventions can be applied to routine service provision (Kazdin, 2008; Novins et al, 2013; Weisz et al., 2013). Deighton et al (2015), using naturalistic service collected data reported no significant difference in outcome between evidence based practice and non-evidence based practice in children with conduct disorders, but a significantly greater improvement according to child self-report among children with emotional disorders. A meta-analysis of psychotherapies found that advantages of evidence-based psychotherapies were statistically and clinically non-significant in samples of clinically referred and diagnosed young people compared to those receiving usual care (Weisz et al 2013). The children and

young people's IAPT programme aims to embed both evidence-based and ROM in CAMHS in England and it will be interesting to observe evaluations of this programme as they emerge.

Our study adds to the growing literature on ROM in CAMHS and in comparison to others that mostly use administrative data, it provides a known response rate, more complete data with less doubt about variations in coding. However, potential limitations with our study are a relatively small sample, despite recruiting from two clinics over 18 months, and potential response bias. Respondents in clinical research studies tend to engage in treatment for longer, have better outcomes and are more satisfied (Garland et al, 2000; Stallard and Potter, 1999). In addition, we necessarily lost practitioners as children were discharged from the clinics. Practitioners can only report the functioning level of children with whom they have had contact, and even if contact after discharge were arranged, it would almost certainly be qualitatively different than normal clinical contact. Future research studies could consider the use of independent clinical researcher evaluations of children's level of functioning, which would require additional resource and represent an additional burden to families, but would also provide a more complete level of practitioner data over time. Similarly, it would be interesting to repeat this analysis on data from an older cohort where the response of young people could also be studied.

What then, should practitioners take from our findings? Wolpert et al (2014<sup>b</sup>) have developed the 'MINDFUL' approach to outcome data, which recommends the combination of multiple sources of data from multiple informants (parents, children and practitioners), using a combination of types of data (process, outcome, activity and experience), and ensuring interpretation using the right methods at the right level of analysis (practitioner, team, clinic, regional or national level). In the current study, both the parent and practitioner reported outcome data suggested that there was no improvement in symptoms, their impact or levels of functioning at six and 12 months follow up in this sample, yet parents reported high levels of satisfaction with most aspects of their experience of the service. The correlation between outcome and satisfaction was unexpectedly low. This suggests that PROMs and PREMs assess different but important aspects of service quality and the use of either alone may provide only a partial indication of how a service functions (Hansen et al, 2010). Our findings emphasise the importance of the use of multiple measures and/or informants as suggested by the MINDFUL approach.

Wolpert and colleagues (2014<sup>b</sup>) also recommend that data is examined for trends over time rather than single data points. We were able to report on the same children over three data points, but only on one

cohort of children. Shortly before data collection commenced, CAMHS A was integrated with social care with changes in staff roles, management and orientation. It seems possible that ongoing upheaval interfered with the effectiveness of the service, as the earlier smaller study that involved CAMHS A reported positive clinical outcomes (Byrne et al, 1999). Sadly, we lacked data from our study to report whether reported outcomes would have improved once the changes had had time to mature in successive cohorts of children, but the impact of repeated service reconfiguration is not without its impact on staff morale, and whether this influences clinical outcomes should be studied empirically.

The MIINDFUL approach also argues that ROM should be applied to enable learning and support innovation at both a local and national level but be divorced from funding decisions due to the inherent uncertainties about what results may mean (Wolpert et al, 2014<sup>b</sup>). We wholeheartedly agree. Funnel plots and scatterplots of intervention effects based on a sample size and precision are a useful method to explore grouped outcome data in CAMHS (Wolpert et al., 2014<sup>b</sup>) and have been recommended for use in regional and national analyses of CAMHS outcomes (Fugard et al, 2014) and physical healthcare (van Dishoeck et al 2011).

The primary goal of CAMHS should surely be the delivery of effective interventions and reduction in psychopathology with improvements in functioning. However, it is encouraging to see that the parents report high levels of satisfaction with clinics even when outcomes may not have been good as this is an important part of engagement, which itself may influence outcome. Our findings suggest that a greater focus on the convenience of the appointment time and setting, and careful attention to explanations about interventions may improve experience, although this would need replicating and testing empirically. Parents' relative lack of confidence in the skills of the practitioners that they met is worrying, and should be explored in other datasets. It may be that the forthcoming CAMHS transformation and the Children and Young People's IAPT programme (Wolpert et al, 2012<sup>a</sup>) both provide a solution (increased practitioner skills and a focus on evidence-based cognitive behavioural approaches that include a significant amount of psychoeducation) and ROM as the means to assess it.

In summary, we found a surprising lack of correlation between parent and practitioner measures of clinical outcome and parent reported experience of services. The measurement of clinical outcomes in CAMHS is complex and requires careful thought. The use of practitioner, parent /carer measures of both

psychopathology and satisfaction is an important method to explore the impact of interventions but should be triangulated with other data about the context and activity of the service.

## **References**

Angold, A., Costello, E. J., Burns, B. J., Erkanli, A. & Farmer, E. M. (2000), Effectiveness of nonresidential specialty mental health services for children and adolescents in the "real world", *Journal of the American Academy of Child and Adolescent Psychiatry*, 39(2), 154-160.

Attride-Stirling, J. (2002), *Development of methods to capture user views of Child and Adolescent Mental Health Services for clinical governance reviews*. Commission for Health Improvement: Project Evaluation Report

Ayton, A., Mooney. M., Sillifant, K., Powls, J., & Rasool, H. (2007), The development of the child and adolescent versions of the Verona Service Satisfaction Scale (CAMHSSS). *Social Psychiatry and Psychiatric Epidemiology*, 42, 892-901. DOI: 10.1007/s00127-007- 0241-9

Bachmann M, Bachmann CJ, John K, Heinzl-Gutenbrunner M, Remshmidt H & Mattejat F. (2010). The effectiveness of child and adolescent psychiatric treatments in a naturalistic outpatient setting. *World Psychiatry*, 9, 111-117.

Barber, A. J., Tischler, V. A. and Healy, E. (2006), "Consumer satisfaction and child behaviour problems in child and adolescent mental health services", *Journal of child health care: for professionals working with children in the hospital and community*, 10 (1), 9-21.

Bridgewater, A., Grayson, A., Brooks, N., Grotte, G., Fabri, B., Au, J., Hooper, T., Jones, M., & Keogh B. (2007), Has the publication of cardiac surgery outcome data been associated with changes in practice in Northwest England? An analysis of 25,730 patients undergoing CABG surgery under 30 surgeons over 8 years. *Heart*, 93, 744-748.

Brown, A., Ford, T., Deighton, J. and Wolpert, M. (2012), Satisfaction with child and adolescent mental health services: translating users' feedback. *Administration and Policy in Mental Health and Mental Health Services Research* DOI 10.1007/s10488-012-0433-9

Byrne, P.J, Croft, C., Chisolm, D., Nikapota, A., and Taylor, E. (1999), *Development of a methodology for the comparison of treatment services and costs in child and adolescent mental health*. Report to the NHS R&D executive and the Institute of Psychiatry.

Chaiton, M., Contreras, G., Brunet, J., Sabiston, C. M., O'Loughlin, E., Low, N. C. P., Karp, I., Barnett, T. A. & O'Loughlin, J. (2013), Heterogeneity of Depressive Symptom Trajectories through Adolescence: Predicting Outcomes in Young Adulthood, *Journal of the Canadian Academy of Child and Adolescent Psychiatry = Journal de l'Académie canadienne de psychiatrie de l'enfant et de l'adolescent*, 22(2), 96-105.

Day, C., Michelson, D. & Hassan, I. (2011), Child and adolescent service experience (ChASE): Measuring service quality and therapeutic process. *British Journal of Clinical Psychology*, 50, 452–464. DOI: 10.1111/j.2044-8260.2011.02008.x

Deighton, J., Argent, R., De Francesco, D., Edbrooke-Childs, J., Jacob, J., Fleming, I., Ford, T. & Wolpert, M. (2015). [Associations between evidence-based practice and mental health outcomes in child and adolescent mental health services.](#) *Clinical Child Psychology and Psychiatry*. Online first June 2015. DOI: 10.1177/1359104515589637

Department of Health (2003b) *Getting the right start: national framework for children. Emerging findings*. TSO:London

Department of Health. (2008) *Guidance on the routine collection of Patient Reported Outcome measures (PROMs)*.

Department of Health. (2010) The NHS White paper. *Equity and Excellence: Liberating the NHS 2010*.

Ford T, Vostanis P, Meltzer H and Goodman R. (2007), Psychiatric disorder among British children looked after by local authorities: a comparison with children living in private households. *British Journal of Psychiatry*, 190, 319-325.

Ford, T., Hutchings, J., Baywater, T., Goodman, A., and Goodman, R. (2009), Strengths and difficulties questionnaire added value scores: evaluating effectiveness in child mental health interventions, *British Journal of Psychiatry*, 194, 552-558.

Ford, T.J., Last, A., Henley, W., Norman, S., Guglani, S., Kelesidi, K., Martin, A., Moran, P., Latham-Cork, H. and Goodman, R. (2013), Can diagnostic assessment be used as a useful adjunct to clinical assessment in child and adolescent mental health services? A randomised controlled trial of disclosure of the Development and Well-Being Assessment to practitioners, *Social Psychiatry and Psychiatric Epidemiology*, 48, 583-593

Fugard, A.J.B., Stapley, E., Ford, T., Law, D., Wolpert, M. and York, A. (2014), Analysing and reporting UK CAMHS outcomes; and application of funnel plots. *Child and Adolescent Mental Health*, DOI: 10.1111/camh.12086

Garland, A.F., Saltzman, M., and Aarons, G. (2000), Adolescents' satisfaction with mental health services: Development of a multi-dimensional scale. *Evaluation and Program Planning*, 23, 165–175.

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

Goodman R, Ford T & Meltzer M. (2002), Mental health problems of children in the community: 18 month follow up. *British Medical Journal*, 324, 1496-1497.

Green H, McGinnity A, Meltzer H, Ford T and Goodman R. (2005). *Mental health of children and young people in Great Britain, 2004* TSO: London.



Hansen, L.K., Vincent, S., David, E, Surafudheen, S., and Kingdon, D. (2010), Patient satisfaction rating scale for psychiatric service users. *Psychiatrist*, 34, 485-8.

Hawley, K. M., and Weisz, J. R. (2005), Youth versus parent working alliance in usual clinical care: Distinctive associations with retention, satisfaction, and treatment outcome, *Journal of Clinical Child and Adolescent Psychology*, 34(1), 117–128.

Health Advisory Service. (1995) Together We Stand: The Commissioning, Role and Functioning of Child and Adolescent Mental Health Service. London: HMSO.

Hurst A, Price A, Davies R, Doolan M, Lanham W and Ford T (2014), Routine Outcome Monitoring of Evidence Based Parenting Programmes: Indications of effectiveness in a community context. *Journal of Children's Services*.9(1), 58-74. DOI 10.1108/JCS-09-2013-0030

Jorg F, Ormel J, Reijneveld SA, Jansen DEMC, Verhulst F and Oldehinkel AJ (2010), Puzzling Findings in Studying the Outcome of “Real World” Adolescent Mental Health Services: The TRAILS Study. *PLoS ONE*, 7(9): e44704. doi:10.1371/journal.pone.0044704

Kazdin, A.E. (2008), Evidence based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base and improve patient care. *American Psychologist*,63 (3),146-159.

Lee, W., Jones, L., Goodman, R and Heyman, I. (2005). Broad Outcome Measures may underestimate effectiveness: an instrument comparison study, *Child and Adolescent Mental Health*, 10,143-144.

MacDiarmid, F., Racey, D., Goodman, R. andFord, T. (In submission).The Persistence of DSM IV Disorders in the three year follow ups of the British Child and Adolescent Mental Health Surveys 1999 and 2014. *Lancet Psychiatry*.

Matthai, J., Anderson, P. and Bourne, A. (2003), Use of the Strengths and Difficulties Questionnaire as an outcome measure in a child and adolescent mental health service, *Australasian Psychiatry*, .11, 334-337.

Norman, S. M. (2011), '*Factors influencing the outcomes of children utilising CAMHS*'. PhD Thesis, Peninsula College of Medicine and Dentistry, Exeter.

Novins, D.K., Green, A.E., Legha, R.K. and Aarons, G.A. (2013), Dissemination and implementation of evidence-based practices for child and adolescent mental health: a systematic review. *Journal of the American Academy of Child and Adolescent Psychiatry*, 52, 1009-1052.

Rotheray, S., Racey, D., Rodgers, L.R., McGilloway, S., Berry, V. and Ford, T. (2014), Further Evidence on the Effectiveness of the Strengths and Difficulties Added Value Score as an Outcome Measure for Child and Adolescent Services. *Child and Adolescent Mental Health*, 19 (4) 270-273 DOI: 10.1111/camh.12059

Schwab-Stone, M. E., Shaffer, D., Dulcan, M. K., Jensen, P. S., Fisher, P., Bird, H. R., Goodman, H., Lahey, B. B., Lichtman, J. H., Canino, G., Rubio-Stipec, M. and Rae, D. S. (1996), Criterion validity of the NIMH Diagnostic Schedule for Children Version 2.3 (DISC\_2.3). *Journal of the American Academy of Child and Adolescent Psychiatry*, 35, 878-888.

Shaffer, D., Gould, M.S., Brasic, J., Ambrosini, P., Fisher, P., Bird, H., and Aluwahlia, S. (1983), A children's global assessment scale (CGAS). *Archives of General Psychiatry*. 40 (11), 1228-31.

Slade, M., Beck, A., Bindman, J., Thornicroft, G., and Wright, S. (1999), Routine clinical outcome measures for patients with severe mental illness: CANSAS and HONOS. *The British Journal of Psychiatry*, 174, 404-408.

Stallard, P., and Potter, R. (1999), Making sense of child and adolescent mental health services. *Psychiatric Bulletin*, Vol 23, pp217-221.

van Dishoeck A.M., Lingsma H.F., Mackenbach J.P and Steyerberg, E.W. (2011). Random variation and ranking of hospitals using outcome indicators. *BMJ Quality Safety*, doi:10.1136/bmjqs.2010.048058

Warren, J. S., Nelson, P. L., Mondragon, S. A., Baldwin, S. A. and 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), 144-155.

Weisz JR, Kuppens S, Eckshtain D, Ugeto AM, Hawley KM and Jensen-Doss A (2013). Performance of Evidence-Based Youth Psychotherapies Compared With Usual Clinical Care: A Multilevel Meta-analysis. *JAMA Psychiatry*, 70, 751-760.

Wolpert, M., Fugard, A.J.B., Deighton, J., and Görzig, A. (2012)<sup>a</sup>, Routine outcomes monitoring as part of children and young people's Improving Access to Psychological Therapies (CYP IAPT) – improving care or unhelpful burden? *Child and Adolescent Mental Health*, 17( 3):129-30.

Wolpert, M, Ford, T, Trustam, E, Law, D, Deighton, J., Flannery, H and Fugard, RJW. (2012)<sup>b</sup>, 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., Gorzig, A., Deighton, J., Fugard, A.J.B., Newman, R., and Ford, T. (2014)<sup>a</sup>. 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 0.1111/camh.12080

Wolpert, M., Deighton, J., De Francesco, D., Martin, P., Fonagy, P., and Ford, T. (2014)<sup>b</sup>. From 'reckless' to 'mindful' in the use of outcome data to inform service-level performance management: perspectives from child mental health. *BMJ Quality and Safety*, 23, 272-276.

Yates, P., Garraida, M. E. and Higginson, I. (1999), Paddington Complexity Scale and Health of the Nation Outcome Scales for Children and Adolescents. *British Journal of Psychiatry*, 174, 417-423.

Zwaanswijk, M., Verhaak, P. F. M., van der Ende, J., Bensing, J. M. & Verhulst, F. C. (2006) Change in children's emotional and behavioural problems over a one-year period: Associations with parental problem recognition and service use, *European Child & Adolescent Psychiatry*, 15(3), 127-131