**Why measuring clinical change at the individual level is challenging but crucial: Commentary on Jensen and Corralejo (2017)**

**Abstract**

Jensen and Corralejo highlight that most mental health intervention research focuses on the group level rather than individual level. They looked at parent-reported outcomes for up to 24 children and found that individual-level findings appear to tell a less positive story than the more traditional group level finding. They call for more reporting of outcomes at individual level.

Professor Wolpert supports this proposal and highlights some further areas for consideration. In particular, she raises the question of what constitutes a good outcome and the lack of consensus around this. She looks at the possible metrics that can be used to measure outcomes, and measurement issues that might arise from using different measures. In conclusion, she supports Jensen and Corralejo’s call to routinely report individual-level outcomes, and suggests using three key metrics – becoming symptom free, substantially improving/deteriorating and moving towards reaching goals – as a starting point for reporting individual-level outcomes. Wolpert stresses the need for more consideration of what constitutes a good outcome and an appropriate range of metrics, particularly taking into account the views of children, young people and families.

**Small data, big implications**

Jensen and Corralejo use a small dataset to illustrate a big issue. Their paper can be seen as part of a groundswell of emerging work that challenges the way therapy outcomes for children and adults are reported, and may even challenge how confident we are in the effectiveness of current interventions (Dragioti, Karathanos, Gerdle, & Evangelou, 2017; Wolpert et al., 2016; Wolpert et al., 2017).

Jensen and Corralejo point out that it has become the norm in most mental health intervention evaluation research to “rely heavily on group statistical analysis to determine if the treatment group and one or more other groups are different in such a way that the difference was unlikely to have occurred by chance”. They refer to the fact that this approach does not indicate whether any individuals in either group have returned to “normal functioning” or “made clinically meaningful change”. They comment that even when these latter approaches are employed they are still used more often at the group level rather than at the individual level. They point out that considering and reporting change only at the group level, including “statistical significance, large effect sizes, and even group clinical significance level”, makes it impossible to distinguish between very different scenarios in terms of outcomes.

The authors consider parent-reported outcomes at the individual level for up to 24 children where parents have completed one or more of the following parent report measures: Parenting Scale, Eyberg Child Behavior Inventory and the Parenting Stress Index. For each measure they calculated improvement for each individual using the reliable change index and a normative comparison of whether the child was now considered to be below the clinical threshold according to the measure developers. They note that these individual-level findings appear to tell a less positive story than the more traditional group level findings:

“while the group statistical analysis and effect size measures would lead one to conclude a strong positive effect from treatment, analysis of clinical significance suggests that more caution should be used in drawing conclusions. Treated individuals made meaningful improvements in decreasing parental stress, but improvements in child behaviour and parenting skills were more moderate and did not meet the cut-off for meaningful change. None of the measures suggested treated individuals would have returned to being within the normative range on these measures following treatment.” (pxx)

They conclude with a call for more reporting of outcomes at the individual level and that this might even be required by journal editors. They note: “We believe that intervention treatment outcome research would benefit from consistent individual analysis”.

I support this proposal. Indeed, this is why we undertook individual-level analysis for our recent report on outcomes across child mental health services involved in the Children and Young People’s Improving Access to Psychological Therapies (CYP IAPT) programme (Wolpert, et al., 2016). Drawing on our learning from that report, earlier work and subsequent consultation, I think Jensen and Corralejo’s suggestions could be extended to include more consideration of some of the complexities and challenges in the approach they advocate, in relation to both conceptual and practical issues. I outline some of the issues below.

**What constitutes a good outcome?**

Jensen and Corralejo take as entirely uncontentious that reliable improvement and return to “normal functioning” are straightforward and uncontested concepts. However, what counts as a good outcome in child mental health is an under-developed area.

There is a lack of clear consensus on the most important outcome to prioritise, why an outcome should be prioritised, and whose perspective should be taken into account. With regard to the most important outcome, most focus to date has been on the measurement of symptoms, with some focus on measuring function (Hoagwood et al., 2012). Various outcome priorities have been identified by different stakeholders, including policy makers, researchers and service users (Childs, Deighton, & Wolpert, 2013) and different frameworks have been suggested, including Hoagwood (2012), and Childs, Deighton and Wolpert (2013). With regard to different perspectives, there is known disparity between child and parent views on goals of therapy (Edbrooke-Childs et al., 2016; Jacob, Edbrooke-Childs, Law, & Wolpert, 2015; Yeh & Weisz, 2001).

Karolin Krause, a PhD student in the Evidence Based Practice Unit, has highlighted that outcomes are not fixed and have changed over time:

“What constitutes a legitimate and desirable outcome of therapy has also varied historically and across socio-cultural contexts. Conceptualisations of outcomes in child mental health are influenced by the political economy and policy priorities, ideological considerations, systems of religious belief, and the normative and scientific paradigms dominating psychological and psychiatric practice and research.”

**Possible metrics of outcome**

In our recent report on outcomes in children seen in child mental health services across England, we derived metrics for good outcomes based on categories adapted from Adult IAPT (Gyani, Shafran, Layard, & Clark, 2013) with an additional metric of movement towards goals."**Recovered**" referred to children and young people who started treatment with self-reported levels of symptoms indicative of significant problems and who ended treatment with no self-reported symptoms indicative of significant problems. "**Reliably improved**" referred to children whose improvement in scores on at least one questionnaire was greater than could be likely due to measurement error alone and there was no reliable deterioration in scores on any measure greater than could be expected from measurement error.[[1]](#footnote-1) "**Moved towards goals**" referred to a child who indicated they had moved towards their pre-agreed goals by at least one point on a 10-point scale. These are far from perfect categories of good outcomes, and measures of functioning, such as attendance and attainment in education, would be valuable additions where data is available. However, they do provide a starting point. Following the report, we undertook consultation on terms used to describe these outcomes and would now suggest “**symptom free**”, “**substantial improvement/deterioration**” and “**movement towards achieving goals**” as the basis for developing a common language for reporting on outcomes at the individual level (<http://www.corc.uk.net/features/talking-outcomes-with-children-and-families-miranda-wolpert-reflects-on-recent-seminars-and-calls-for-your-ideas/>).

**Measurement issues**

Jensen and Corralejo do not comment in detail on what their findings of different recovery rates for the different measures might mean. In our report on outcomes across services in England, where a wide variety of different child- and parent-reported measures was used, we 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 and PHQ9. 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). 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). The fact that different scales use different thresholds presents challenges for comparing across scales and populations.” (Wolpert, et al., 2016).

There needs to be more analysis and publication of outcomes achieved by different measures that purport to measure the same thing, so that they can be compared in more meaningful ways (Wolpert, Cheng, & Deighton, 2015).

**Conclusion**

We support the call to action set by Jensen and Corralejo of routinely reporting individual-level outcomes. In relation to this, a starting point for common metrics is outlined above and might focus in the first instance on metrics related to becoming symptom free, substantially improving, and moving towards achieving goals. However, there is an urgent need for more detailed consideration of what constitutes a good outcome in child mental health and what the appropriate range of metrics should be, with particular regard for the views of children, young people and families. The current author and colleagues hosted an interdisciplinary conference to debate this, open to all, on 3 July in Central London (<http://www.ucl.ac.uk/evidence-based-practice-unit/news>). It is hoped that many of you will join this debate to so that we, as a community, can agree a common set of individual-level outcome indicators. This will allow us to rise to the challenge posed by Jensen and Corralejo.

**References**

Childs, J., Deighton, J., & Wolpert, M. (2013). Defining and measuring mental health and wellbeing in children: A response mode report requested by the Department of Health for the Policy Research Unit in the Health of Children, Young People and Families. London: CAMHS Press.

Dragioti, E., Karathanos, V., Gerdle, B., & Evangelou, E. (2017). Does psychotherapy work? An umbrella review of meta-analyses of randomized controlled trials. *Acta Psychiatrica Scandinavica*. doi: 10.1111/acps.12713

Edbrooke-Childs, J., Jacob, J., Argent, R., Patalay, P., Deighton, J., & Wolpert, M. (2016). The relationship between child- and parent-reported shared decision making and child-, parent-, and clinician-reported treatment outcome in routinely collected child mental health services data. *Clinical Child Psychology and Psychiatry, 21*(2), 324­–338. doi: 10.1177/1359104515591226

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

Hoagwood, K., Jensen, P. S., Acri, M. C., Olin, S. S., Lewandowski, R. E., & Herman, R. J. (2012). Outcome domains in child mental health research since 1996: Have they changed and why does it matter? *Journal of the American Academy of Child & Adolescent Psychiatry, 51*(12), 1241–1260.

Jacob, J., Edbrooke-Childs, J., Law, D., & Wolpert, M. (2015). Measuring what matters to patients: using goal content to inform measure choice and development. *Clinical Child Psychology and Psychiatry, Online first*. doi: 10.1177/1359104515615642

Wolpert, M., Cheng, H., & Deighton, J. (2015). Measurement Issues: Review of four patient reported outcome measures: SDQ, RCADS, C/ORS and GBO – their strengths and limitations for clinical use and service evaluation. *Child and Adolescent Mental Health, 20*(1), 63–70. doi: 10.1111/camh.12065

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.

Wolpert, M., Vostanis, P., Martin, K., Munk, S., Norman, R., Fonagy, P., & Feltham, A. (2017). High integrity mental health services for children: focusing on the person, not the problem. *BMJ, 357*. doi: 10.1136/bmj.j1500

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

1. Reliably deteriorated referred to a child whose scores on any measure had got reliably worse, regardless of whether there had been reliable improvement on any other measure. [↑](#footnote-ref-1)