OPEN ACCESS

From 'reckless' to 'mindful' in the use of outcome data to inform service-level performance management: perspectives from child mental health

Miranda Wolpert,[1] Jessica Deighton,[2] Davide De Francesco,[2] Peter Martin,[2] Peter Fonagy,[3] Tamsin Ford[4]

For numbered affiliations see end of article.

Correspondence to
Dr Miranda Wolpert, Child and Adolescent Mental Health Services (CAMHS) Outcomes Research Consortium, Evidence Based Practice Unit, UCL and Anna Freud Centre, 21 Maresfield Gardens, London, NW3 5SU, UK;
ebpu@annafreud.org

Miranda Wolpert and colleagues draw on their learning from national outcome data collection across child mental health services in the UK to propose a framework for lead clinicians and funders to collaboratively adopt to ensure effective and safe use of outcome data for performance management and quality improvement across the range of healthcare settings.

Received 3 October 2013
Revised 20 December 2013
Accepted 4 January 2014
Published Online First
23 January 2014

To cite: Wolpert M, Deighton J, De Francesco D, *et al*. *BMJ Qual Saf* 2014;**23**:272–276.

INTRODUCTION

In the recent report on patient safety in the National Health Service (NHS) in England, Don Berwick calls on the NHS to align the necessity for increased 'accountability' with the necessity to 'abandon blame as a tool' in order to develop a 'transparent learning culture'.[1] Sir Bruce Keogh, Medical Director NHS, and colleagues' recent analysis of outlier hospitals based on mortality data marks a key step on this journey, but has led to high-profile debate about the risk of possible 'reckless' (Sir Bruce Keogh's term) use of data if appropriate parameters are not established.[2 3] If these and other equivalent proxies for outcomes are to be used safely and effectively to support performance management and quality improvement in the ways envisioned by both Keogh and Berwick, it is crucial to establish clearly agreed operational procedures. Drawing on our experience of collecting and interpreting outcome data in the challenging context of child mental health across the UK, we suggest adoption of a MINDFUL framework involving consideration of *m*ultiple perspectives, *i*nterpreting differences in the light of current evidence base, focus on *n*egative differences when triangulated with other data, *d*irected discussions based on 'what if this were a true difference' (employing the 75–25% rule), use of *f*unnel plots as a starting point to consider outliers, appreciation of *u*ncertainty as a key contextual reality and the use of *l*earning collaborations to support appropriate implementation and action strategies.

COMPLEXITIES

Any attempt to measure 'impact' of a service using a given 'outcome' is complex. The Keogh report acknowledges: "two different measures of mortality, HSMR [Hospital Standardised Mortality Ratio] and SHMI [Summary Hospital Level Mortality Indicator] generated two completely different lists of outlier trusts." This was 'solved' by using both lists, but with a suggestion to move to one measure of morbidity in the future. Yet challenges remain: other measures of outcome may be relevant to consider (eg, years of high quality life) and any measure of risk adjustment (even one as well accepted as the European System for Cardiac Operative Risk Evaluation (EuroSCORE) for heart surgery[4]) may not control for all factors that impact on outcomes.[5]

If consideration of performance of hospitals in terms of morbidity data is complicated, then the challenges of applying outcome measures in mental health may appear insurmountable. There is no equivalent 'hard' indicator with the status of mortality. Relevant mental health outcomes include symptom change, adaptive functioning, subjective well-being and experience of recovery. There is no one commonly accepted risk adjustment model equivalent to EuroSCORE, although there is evidence that case severity at the outset may be among the most powerful predictors of outcome.[6] The evaluation of children's outcomes is additionally complicated by the need to elicit and interpret the views of children at different developmental stages, and the

need to consider these views alongside those of carers and other stakeholders.

CHILD AND ADOLESCENT MENTAL HEALTH SERVICES OUTCOMES RESEARCH CONSORTIUM

Despite, or perhaps because of, these challenges, child mental health has been pioneering among mental health specialties in efforts at collecting information relevant to assessing impact of services, with particular emphasis on outcomes from the service user perspective. Child, parent and clinician reports of change in symptomatology, impact on daily living, achievement of long-term goals and experience of service offered have been voluntarily collected and centrally collated by a practice-research network of child and adolescent mental health service (CAMHS) teams across the UK; the CAMHS Outcomes Research Consortium (CORC).[7]The collaboration has grown from 4 subscribing organisations in 2004 to over 70 in 2013, and now holds data on one or more indicators of outcome of treatment for over 59 826 episodes of care. An independent audit found that implementation of CORC protocols across a service (2011–2013) was associated with a doubling in use of repeated outcome measurement (from 30% to 60% of cases who had at least one outcome measure assessed at two time points).[8]

Below we draw out a framework for collaborative use of outcome data between clinical leads and funders drawing on learning from CORC's experience in the challenging environment of child mental health.

Multiple voices

A decision was made early on in the CORC collaboration not to seek to combine perspectives in any one overall performance score but rather to consider the views of at least three judges of how good or bad the outcome of a treatment is: child, parents or carers, and clinician. In practice these views often differ, with parents and children, for example, sharing no more than 10% of the variance in their perception of difficulties.[9] Each may be important in terms of understanding different aspects relevant to performance management. Children's own views may be crucial to ensure the voice of the child influences review of services, and there is evidence that children as young as 8 years old can reliably comment on their experiences and outcomes.[10] However, parents can also offer rich insights on particular areas, such as reporting changes in behavioural difficulties exhibited by children.[9] Clinicians are important reporters particularly in relation to complex symptomatology and functioning.[11] For other areas of healthcare, deciding whose perspectives need to be included is a crucial first decision to be agreed by funders and clinical leads.

Interpreting differences

In child mental health where emphasis is on integrated interagency working and care pathways, the most meaningful focus for interpretation of difference for clinical leads and funders to consider is likely to be at the level of the integrated clinical unit rather than individual practitioners. In CORC, the central team aggregate data from participating services and produce a report that compares outcomes with those of other services in the collaboration. A range of methods is employed to try to ensure fair comparisons. One metric derived from the parent report is based on comparison with outcomes for a non-treated group to control for fluctuations in psychopathology and regression to the mean.[12] CORC members can additionally request bespoke comparisons with services who are 'statistical neighbours' in relevant domains. Any statistically significant differences between outcomes are highlighted. For other areas of healthcare, determining the level of granularity most relevant for interpreting differences and what metric is to be used to control for case complexity are key decisions to be collaboratively made at the outset.

Negative outcomes

CORC recommends member services use these reports as a starting point for exploration in team meetings and in conversations with commissioners and others. A sequenced approach is recommended whereby the focus is on areas where there is statistical difference in a negative direction between the service and the comparators. This sequence consists of consideration of the following hypotheses through exploration of the available data:

1. The measure used is not right for this population.
2. There were data entry or analysis errors.
3. The presenting problems are different from comparators.
4. There were cultural differences between groups.
5. There were differences in the service type or treatments offered.

It also involves triangulation with other indicators of performance (eg, did not attends, adverse incidents and user satisfaction).

This systematic consideration of 'negative' outcomes data may be helpful for other healthcare contexts where the inherent uncertainty of statistical estimates needs to be taken into account when evaluating whether an ostensibly 'poor outcome' reflects an underlying problem with service performance or not.

Directed discussion

CORC has developed guidance to help counteract the human tendency to explain any potential alerts to poor performance as due to flaws in the data or lack of appropriate case complexity control.[13 14] The guidance is to spend 25% of any allocated discussion time on consideration of the hypotheses set out above, and 75% of time on a thought experiment: "if these data

were showing up problems in our practice what might they be and how might we investigate this and rectify these issues."

Across healthcare systems we need to train and support a generation of clinical leaders who can make use of such data and ally it with a detailed understanding of what works for whom derived from the literature and from rigorous research trials. In the surgery data, mortality outliers were noted to be characterised by the "sub-optimal way in which emergency patients are dealt with, particularly at the weekend and at night".[3] In child mental health, poorer outcomes have been associated with lower levels of clinician morale and involvement in service-level decision making,[15] less use of evidence-based practice[16] and less reflection on outcome data.[17] Clinical leaders are encouraged to discuss with commissioners strategies to test out different hypotheses as to the link between outcomes and processes of care.

Funnel plots

In all areas of healthcare, presentation of differences between services by use of funnel plots helps redress the tendency to spend time on interpreting differences that are likely to be the result of random variation. Figure 1 shows risk-adjusted change in child mental health difficulties from the three key perspectives routinely collected by CORC (child, parent/carer and clinician). For each mental health team, the difference between predicted and observed outcomes (y axes) is plotted against the number of cases seen (x axes) for (a) child-reported, (b) parent-reported and (c) clinician-reported outcomes based on change scores for the Strengths and Difficulties Questionnaire

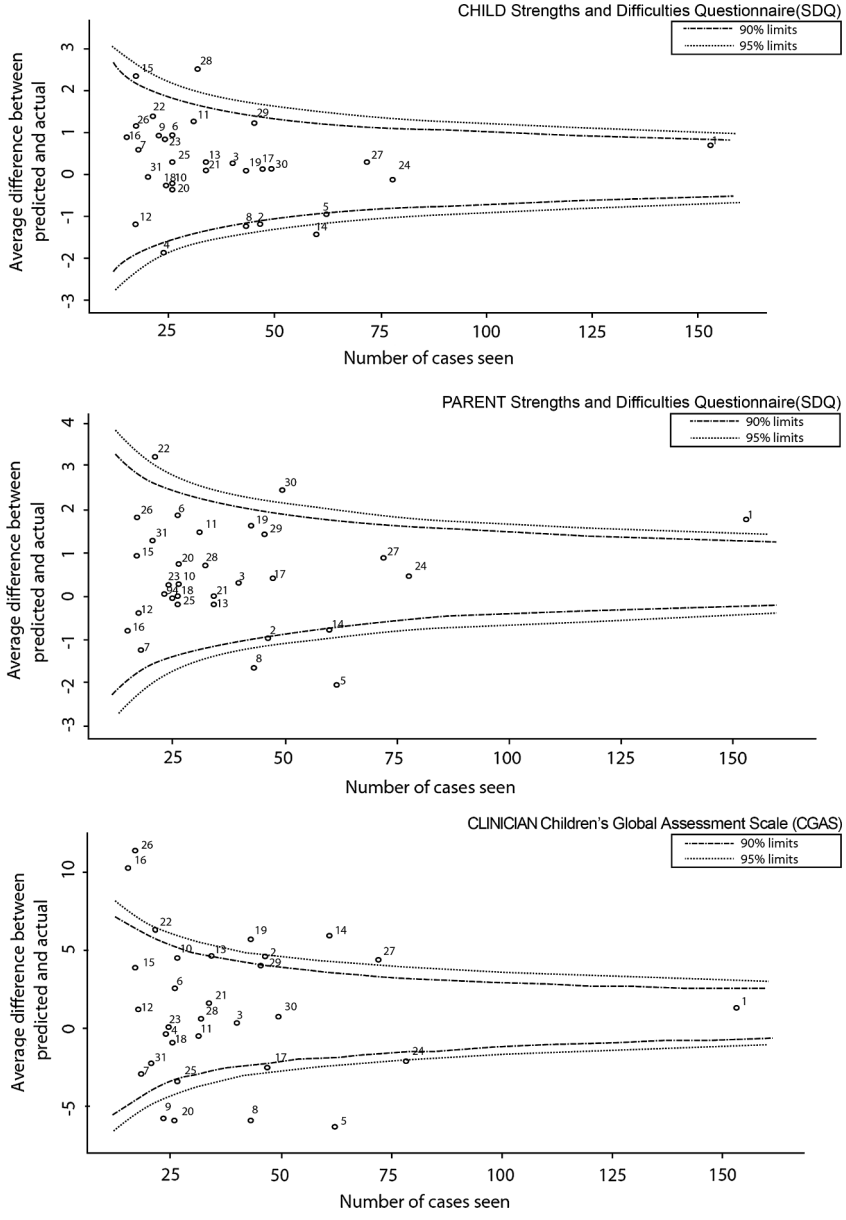

Figure 1 Funnel plots of service performance based on differences between observed and expected scores on a range of outcome measures.

(a, b)[18] and the Child Global Assessment Scale (c).[19] These funnel plots show the outcomes for 1194 children (aged 9–18, median age 14, 60% female) from 31 teams (2008–2012).

Figure 1 suggests there are more services that fall outside the 95% CI than would be expected by chance alone for both parent-rated and clinician-rated outcomes. The services that fall outside of the confidence bounds (shown for both 90% and 95% CI) are different for children-rated, parent-rated and clinician-rated outcomes, although two services are identified by two perspectives (clinicians and parents). All outliers identified from any perspective in figure 1 should be taken as the starting point for directed discussion, using the approach outlined above, to consider potential reasons for variation in service outcomes.[20]

Uncertainty

Consideration of multiperspective paediatric mental health data may be helpful in counteracting the tendency to paint overly simplistic reductionist portraits that serve only to gratify journalistic and public appetite for vilification. In all areas of healthcare, no one measure of outcome is perfect. No one perspective trumps all others. No one way of controlling for case complexity is fool proof. No services are all good or all bad.

In the light of this, it would be 'reckless' to attempt to identify any single service as 'underperforming' or 'overachieving' on the basis of a funnel plot alone,[3] and these should be always seen as starting points for investigation not end points. For 'MINDFUL' use of data to occur, it is crucial that clinical leads and funders acknowledge from the outset the need to operate collaboratively under conditions of uncertainty, which means considering hypotheses for differences and aligning their findings with the current evidence on best practice.

Learning collaboration

A key learning from CORC has been the power of a learning collaboration that includes clinicians, managers, service users and commissioners. Thus, members have been able to develop their own data standards, determine measures that are relevant for their communities and agree approaches between funders and providers that allow support for quality improvement without introducing perverse incentives.[21] Of course, major challenges remain in terms of both data quality and quantity. Sir Bruce Keogh has commented: "when you start these things the data is never as good as you would like and that's where the courage comes in....data only becomes good when you use it."[2] In order for this courage to be possible, we would call for an explicit commitment to 'MINDFUL' use of data that allows for multiple perspectives, encourages focus on what differences might mean in consideration alongside the latest evidence, explicitly acknowledges uncertainty and is supported by local and national learning collaborations.

CONCLUSION

The 'MINDFUL' approach to the use of outcome data to inform service-level performance involves the consideration of *m*ultiple perspectives, *i*nterpretation focused on *n*egative differences and use of directed discussions based on 'what if this were a true difference' (employing the 75–25% rule). Funnel plots should be used as a starting point to consider outliers, always keeping in mind an appreciation of *u*ncertainty with *l*earning collaborations of clinicians, commissioners and service users supporting data analyses. Adopting this 'MINDFUL approach to the variation

Key messages

Mindful use of outcome data in collaboration between service providers and commissioners involves the following

- *M*ultiple perspectives should be considered. This means practitioners and commissioners agreeing at the outset which perspectives on outcomes may be relevant in any areas of work. In child work, this is likely to include as a minimum child, parent/carer and clinician.
- *I*nterpreting differences at the right level of granularity (individual professionals or teams or services) and in the light of best current knowledge as the factors affecting case mix is crucial. Commissioners and providers should agree the appropriate unit of comparison and the best proxy for case complexity, recognising these will inevitably be flawed and partial.
- *N*egative differences (where the unit appears to be doing significantly worse than expected compared with other units adjusted for case mix) should be a starting point for consideration and hypothesis testing.
- *D*irected discussions should focus on 'what if this were a true difference' rather than only considering why the findings may be flawed (75–25% rule).
- *F*unnel plots are a key way of presenting information to allow consideration of outliers and are to be preferred to any simple ranking, but should still be considered a starting point not an end point.
- *U*ncertainty as a key contextual reality has to be appreciated and accommodated. In the light of uncertainty, it is imperative that funders and clinical leads do not make decisions on any single data source alone, but rather triangulate data from a range of sources.
- *L*earning collaborations, both local and national, can support development of appropriate measures and approaches.

identified in child mental health services (figure 1) involves outlier services, their commissioners and service users collaboratively considering, in directed discussions, possible explanations for the variation at the level of the clinical team. This creates the possibility of collaborative consideration of data to inform service review and improvement and reduces the possibility of "reckless" interpretation by any one group.

Author affiliations
[1]Child and Adolescent Mental Health Services (CAMHS) Outcomes Research Consortium, Evidence Based Practice Unit, UCL and Anna Freud Centre, London, UK
[2]Evidence Based Practice Unit, UCL and Anna Freud Centre, London, UK
[3]Research Department of Clinical, Educational and Health Psychology, UCL and Anna Freud Centre, London, UK
[4]Medical School, University of Exeter, Exeter, UK

Acknowledgements The authors would like to thank the CORC Committee at the time of writing (Ashley Wyatt, Miranda Wolpert, Julie Elliott, Mick Atkinson, Alan Ovenden, Tamsin Ford, Alison Towndrow, Ann York, Kate Martin and Duncan Law) and the CORC Central Team at the time of writing (Jenna Bradley, Ruth Plackett, Rachel Argent and Isobel Fleming).

Contributors MW conceived of and led the writing of this article and acts as a guarantor. JD contributed to the writing and to the plan for the analysis of data. DDF contributed to the data analytic plan and undertook the data analysis. PM contributed to the analytic plan and the data analysis and discussion. PF provided comments and contributions to the content and structure. TF contributed to the writing, to the structure and the data analysis. Information used to prepare the article comes from the CORC dataset.

Competing interest MW is a paid director (1 day a week) for CORC; a not-for-profit learning collaboration. TF is an unpaid director of CORC.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 National Advisory Group on the Safety of Patients in England. *A promise to learn—a commitment to act. Improving the Safety of Patients in England*. London: Department of Health, 2013. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/226703/Berwick_Report.pdf
2 BMJ. Surgical Outcome Data. [Podcast] 2013 [6th July 2013]. http://www.bmj.com/podcast/2013/07/05/surgical-outcome-data?utm_source=feedburner-bmj.com+download&utm_medium=feed&utm_campaign=Feed%3A+bmj%2Fpodcasts+%28The+BMJ+podcast%29
3 Keogh B. Review into the quality of care and treatment provided by 14 hospital trusts in England: Overview report 2013. http://www.nhs.uk/NHSEngland/bruce-keogh-review/Documents/outcomes/keogh-review-final-report.pdf
4 Nashef SA, Roques F, Michel P, *et al*. European system for cardiac operative risk evaluation (EuroSCORE). *Eur J Cardiothorac Surg* 1999;16:9–13.
5 Lilford R, Mohammed MA, Spiegelhalter D, *et al*. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. *Lancet* 2004;363:1147–54.
6 Norman SM, Henley W, Ford T. Case complexity and outcomes: what do we know and what do we need to know? In: Essau C. ed *Inequalities*. London: Association for Child and Adolescent Mental Health, In press.
7 Wolpert M, Ford T, Trustam E, *et al*. Patient-reported outcomes in child and adolescent mental health services (CAMHS): use of idiographic and standardized measures. *J Ment Health* 2012;21:165–73.
8 Hall CL, Moldavsky M, Baldwin L, *et al*. The use of routine outcome measures in two child and adolescent mental health services: a completed audit cycle. *BMC Psychiatry* 2013;13:270.
9 Verhulst FC, Van der Ende J. Using rating scales in a clinical context. In: Rutter M, Bishop D, Pine DS, Scott S, Stevenson J, Taylor E, *et al*. eds. *Rutter's Child and Adolescent Psychiatry*. 5th edn. Malden, Oxford and Carlton: Blackwell, 2008:289–98.
10 Department of Health. *Children and Young People's Health Outcomes Strategy*. London: Department of Health, 2012. http://www.ncvys.org.uk/UserFiles/DH_CYP_Health_Outcomes_Strategy_Briefing.pdf
11 Garralda ME, Yates P, Higginson I. Child and adolescent mental health service use. HoNOSCA as an outcome measure. *Br J Psychiatry* 2000;177:52–8.
12 Ford T, Hutchings J, Bywater T, *et al*. Strengths and difficulties questionnaire added value scores: evaluating effectiveness in child mental health interventions. *Br J Psychiatry* 2009;194:552–8.
13 Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. *Science* 1974;185:1124–31.
14 Gray JAM, Gray JAME-bh. *Evidence-based healthcare and public health: how to make decisions about health services and public health*. 3rd edn. Edinburgh: Churchill Livingstone, 2009.
15 Glisson C, Hemmelgarn A. The effects of organizational climate and interorganizational coordination on the quality and outcomes of children's service systems. *Child Abuse Negl* 1998;22:401–21.
16 Weisz JR, Kazdin AE. The present and future of evidence-based psychotherapies for children and adolescents. In: Weisz JR, Kazdin AE. eds *Evidence-based psychotherapies for children and adolescents*. London: The Guilford Press, 2010:439–51.
17 Bickman L, de Andrade AR, Athay MM, *et al*. The relationship between change in therapeutic alliance ratings and improvement in youth symptom severity: whose ratings matter the most? *Adm Policy Ment Health* 2012;39:78–89.
18 Goodman R, Scott S. Comparing the strengths and difficulties questionnaire and the child behavior checklist: is small beautiful? *J Abnorm Child Psychol* 1999;27:17–24.
19 Shaffer D, Gould MS, Brasic J, *et al*. A children's global assessment scale (CGAS). *Arch Gen Psychiatry* 1983;40:1228–31.
20 Spiegelhalter DJ. Funnel plots for comparing institutional performance. *Stat Med* 2005;24:1185–202.
21 CORC. *CORC's Position on Commissioning for Quality and Innovation (CQUIN) targets concerning outcome measures*. London: CORC, 2013. http://www.corc.uk.net/wp-content/uploads/2012/03/CORCs-Position-on-CQUIN-targets-03042013.pdf