

What kind of evidence do we need for evidence-based mental health policy?

The case of the Better Access initiative.

Nicholas B. Allen^{1,2}

Henry J. Jackson¹

1. Psychological Sciences, University of Melbourne, Australia
2. Orygen Youth Health Research Centre, Parkville, Victoria

Nicholas B. Allen (Correspondence), Professor, Psychological Sciences and Orygen Youth Health Research Centre, University of Melbourne.

Email: nba@unimelb.edu.au

In this issue of the journal, Jane Pirkis and colleagues report some of the core findings from the evaluation of the Better Access scheme that was funded by the Commonwealth Department of Health and Aging. [1]. This report, along with others [2-5], provides some of the first evaluations of this innovative government program. As is noted in the companion editorial by Tony Jorm [6], the program is significantly more expensive than was projected; the Department of Health and Ageing initially estimated forward expenditure on the Program of \$538m over 4 years. The cost over this period is now understood to be about triple that figure. In a context where every dollar of public mental health spending is hard won by the sector, the effectiveness of this program becomes a critical matter. If it is effective, then it represents one of the most significant, and praiseworthy, expansions of mental health services ever instigated in public policy. Indeed, if the program is effective, the fact that it has become much more costly than expected reflects not only the popularity of the program with consumers and referrers, but will presumably result in an even more beneficial impact on the nation's mental health. However, if it is not effective, then the program represents misspent public funds *and* an enormous opportunity cost in the sense that the money could be spent on other programs that might be much more influential in improving the mental health of Australians.

Perhaps because of these considerations, along with the fact that the scheme has the potential to aggravate professional rivalries, both between and within the professions, the program has attracted some controversy - often from

commentators who lack any data to back up either their criticism or defence of the scheme. For this reason the recent evaluations of the scheme, including the paper in the current issue, are to be welcomed as they have the potential to serve as a vital resource in this important public debate. Indeed, in the companion editorial by Professor Jorm [6], he clearly lays out a series of criticisms that have been levelled at the scheme, and gives a generally positive scorecard for Better Access, largely based on these recent evaluations.

Our interest here is not so much to comment on whether a positive or negative evaluation of the scheme can be made, but rather to comment on the limitations to what conclusions can be drawn given the very considerable methodological shortcomings of the evaluation. Indeed, we believe there are many key questions that are simply not answered by the current evaluation, in large part due to inadequate methodology. We fully appreciate the reasons why it would not have been possible to conduct a “methodologically pristine” evaluation of the scheme, such as a traditional randomised controlled trial, in the current environment. Clearly significant limitations apply to the evaluation that can be conducted of a government scheme that has already been rolled out on a nationwide basis (as noted by Tony Jorm [6]), and where limited financial resources are made available for the evaluation (as alluded to by Pirkis and colleagues in their discussion of limitations [1]). It is also true to say, as Pirkis and colleagues do, that this evaluation is more rigorous than that applied to most government health spending - a clearly

lamentable state of affairs. However, none of these limitations addresses the fundamental issue of whether the relevant questions have actually been answered. Just because an evaluation has been conducted and may be the best given the circumstances, does not mean that it is adequate to answer the questions. Indeed, there comes a point at which one can reasonably ask how much methodological compromise can be tolerated before an evaluation causes more harm than good. Even a flawed evaluation, simply because it carries the word “evaluation”, especially one with the imprimatur of government funding, is likely to enter into the public discourse as a series of “facts” that interested parties can use to advance their interests.

What questions must an evaluation of the Better Access Scheme address? Given the COAG National Action Plan on Mental Health (2006–11) [7], and the stated aims of the Better Access Scheme, a series of fundamental questions have been proposed by Rosenberg and Hickie [8], including:

1. To what extent does the program improve access to evidence-based mental health care for people who present to a GP with a common mental disorder?
2. What are the demographic and illness characteristics of consumers with mental disorders attending GPs who do and do not receive the new service enhancements?
3. To what extent do the various service enhancements result in better mental health outcomes for people attending GPs with a common mental disorder?

4. To what extent do the various service enhancements meet consumer needs and expectations?

We agree that these are an appropriate set of questions; however, to answer them any evaluation needs to wrestle with certain fundamental design issues that include the representativeness of the sample, unbiased (reliable) and appropriate (valid) measurement of outcomes, and appropriate control or comparison conditions. On each of these criteria, which are fundamental to any scientific endeavour, we have serious concerns regarding the current evaluation.

The first and very major methodological weakness of the current evaluation is the absence of a control or comparison condition of any type. Any study of the effectiveness of a new program or initiative must be able to say whether it provides significant incremental benefits when compared to the absence of such spending. In other words, simply showing that participants in this program show improvements over time cannot answer this question. Indeed, any epidemiological study that tracks symptoms over time in the community will show evidence of reduction amongst those who initially score highly on symptom scales. This is due to regression to the mean effects as well as the natural history of chronic but episodic high prevalence disorders like clinical depression and some anxiety disorders. Given that the current evaluation has essentially followed a group of help seeking individuals across an initial phase of treatment, it is not at all surprising that the pre- post-test results are highly significant. But the big question is how do these

changes compare to another comparison condition; a question that cannot be answered by the current evaluation. As such, the study cannot conclude that the treatment provided by the Better Access initiative is effective, let alone more effective than the treatment provided prior to the scheme. The most it can probably say is that the scheme does no harm.

Pirkis and colleagues [1] do identify this limitation, correctly stating, “without a control group it is not possible to say definitively that Better Access contributed to these improvements” (page XX). However they go on to assert that two other factors can be taken as support for the conclusion that the Better Access interventions were responsible for the improvements. First was that the consumers predominantly received Cognitive Behaviour Therapy (CBT), which is considered an evidence-based treatment [9]. However, this begs the question as to whether these consumers were actually receiving CBT. This was not evaluated aside from a report from the practitioner –a practitioner whose receipt of Better Access funding relies on them asserting that they are providing an evidence-based treatment such as CBT. Given that recent studies have found that many practitioners claiming to provide CBT either do not provide a treatment that conforms to the basic tenets of CBT, or do not deliver the treatment with adequate fidelity [10], the assumption that these consumers have been receiving an evidence-based treatment akin to that evaluated in randomised controlled efficacy trials is unwarranted as we have no real data on what the clinicians are actually doing.

The other factor in support of the effectiveness of the Better Access initiative that is cited by Pirkis and colleagues is that participants attribute their improvement to the treatment they have received. Although this is undoubtedly important consumer satisfaction information, it has little evidential weight when determining whether a treatment or program is effective. If it did we might be surprised at some of the interventions we would have to declare effective! We are sure that practitioners of homeopathy (just to name one example) would welcome the opportunity to assert to government that consumer satisfaction constituted support for their effectiveness, but we doubt that most mental health professionals would consider this adequate evidence.

The second major methodological limitation of the current evaluation is the fact that practitioners selected participants for inclusion in the study, and collected and entered the primary outcome data. Moreover, the participants knew the practitioners would be doing this. This has implications for both the representativeness of the sample, and the reliability of outcome measurement, and leads to a number of highly plausible potential biases in the data collection on both the practitioner and consumer's part. Practitioners are likely to be motivated to overestimate the consumer's benefit from their work and may have over selected consumers they considered likely to improve from their intervention. The fact that the sample is demographically representative does not allay these concerns. On the other hand consumers may feel a compulsion to provide socially desirable answers,

given the practitioner's expectation that they improve. Moreover, approximately 30% of the consumers enrolled in the evaluation did not provide outcome data and were not included in analyses, precluding the use of "intention to treat" analyses that are often considered a gold standard for treatment outcome research.

Adding to these concerns is the fact that the diagnoses were based on practitioner-provided diagnoses. Systematic evaluations of the agreement between practitioner diagnoses and those obtained with research standard structured diagnostic interviews have found poor agreement, especially for depressive and anxious disorders, and have furthermore found that practitioners are more likely to rate a diagnosis as present than are those administering research quality interviews [11]. This latter fact may reflect a bias on behalf of practitioners to conclude that because a consumer is seeking and receiving professional services, they must have a case level disorder - a bias that is understandable given the nature of their work, but that speaks to exactly why rigorous research studies do not utilise practitioner-provided diagnoses without at least some independent assessment of their reliability and validity. Further adding to these concerns is that the practitioners in the survey would be likely to vary widely in their training and experience in diagnosis of mental health disorders, and many general practitioners and general psychologists in particular may have received very little systematic supervised training in diagnostic assessment. Basing any inference about the severity and type of the mental health problems that consumers are experiencing on unstructured practitioner diagnoses alone is therefore extremely problematic.

We are also concerned that the breadth and type of outcome assessment utilised in the evaluation was too limited. The evaluation only reports data from two self-report scales that assess symptoms of distress, depression and anxiety. This leaves us with no information about critical aspects of consumer outcome, including comorbid symptoms (especially those relating to complexity), functioning and disability, and longer-term outcomes that provide insight into whether treatments promote resilience and reduce relapse. (This latter point is a particularly critical matter for relapsing conditions such as high prevalence mental health problems, and one on which the psychological therapies provided by the Better Access Scheme might be expected to provide benefits [12]).

Perhaps the most provocative aspect of the report, given the self-interest of the professional groups involved, is the separate analyses of the clinical psychologist, generalist psychologist, and general practitioner groups. Although Pirkis and colleagues state that it was not appropriate to either pool these results, nor to perform statistical comparisons between the groups, the results are nevertheless given some discussion in the paper, and will undoubtedly be like catnip to professional groups who are sometimes more interested in protecting their members' access to public funding of their work than they are in making an unbiased evaluation of what might be the best type of mental health system for the Australian community. Indeed, in Tony Jorm's companion editorial, he already

attempts to draw some conclusions from these patterns of data (although we hasten to point out that he does not belong to one of the professions being evaluated and is therefore in that sense a relatively unbiased observer).¹

The overall problem here is that conclusions are more compelling than caveats, so even with an appropriate set of limitations enumerated in the manuscript and associated reports [1-5], the compulsion for stakeholders to consider the questions outlined above as being “answered” will be hard to resist. The professions wish to justify their members’ access to public funding of their work, the government wishes to justify their significant investment in the program as an efficient use of taxpayer funds, and consumers clearly appreciate the government subsidies supporting their out of pocket expenses for their mental health services. Each of these interests are perfectly understandable and to be expected. However, scientific research has developed a series of procedures that are precisely designed to provide answers to questions that are relatively unbiased by these types of interests. These particularly have to do with independent (i.e., disinterested and unbiased) sampling, observation and inference, as well as provision of appropriate control conditions. The current evaluation could have done much better in all of these regards. For example, although some have asserted that it is not possible to provide a control condition for a program that is already available on a population wide basis, this is not true. There are a number of methodologies, including the use of historical (pre-scheme) control samples, and multiple within-participant baseline measurements prior to treatment, that would be very feasible in the current

circumstances. Indeed such designs have been proposed for the evaluation of the current scheme [8] but have not been utilised.

It is quite ironic that the evaluation of a multi-billion dollar government scheme has methodological shortcomings that in another context would be likely to preclude it from receiving an NHMRC grant at a tiny fraction of this cost. Most of the limitations of the current evaluation are not inherent to the scheme or the current circumstances, but rather could have been overcome with more adequate funding for the evaluation. We look forward to a more compelling set of answers to these pressing matters of public interest, not just in the evaluation of the Better Access scheme, but in all public health spending. And that is mainly a matter of the political will to subject extensive public funding to appropriately funded evaluations – evaluations that really can answer questions.

References

1. Pirkis J, Ftanou M, Williamson M, Machlin A, Spittal M, Bassilios B, Harris M. An evaluation of Australia's Better Access initiative. *Aust N Z J Psychiatry* in press.
2. Bassilios B, Pirkis J, Fletcher J, Burgess P, Gurrin L, King K, Kohn F, Blashki G. The complementarity of two major Australian primary mental health care initiatives. *Aust N Z J Psychiatry* 2010; 44:997-1004.
3. Byles JE, Dolja-Gore X, Loxton DJ, Parkinson L, Stewart Williams JA. Women's uptake of Medicare Benefits Schedule mental health items for general practitioners, psychologists and other allied mental health professionals. *Med J Aust* 2011; 194:175-179.
4. Harris MG, Burgess PM, Pirkis JE, Slade TN, Whiteford HA. Policy initiative to improve access to psychological services for people with affective and anxiety disorders: population-level analysis. *Br J Psychiatry* 2011; 198:99-108.
5. Pirkis J, Harris M, Hall W, Ftanou M. *Evaluation of the better access to psychiatrists, psychologists and general practitioners through the medicare benefits schedule initiative: summative evaluation*. Melbourne: Centre for Health Policy, Programs and Economics, University of Melbourne, 2011.
6. Jorm, A.F. (2011) Australia's Better Access Initiative: Do the Evaluation Data Support the Critics? *Aust N Z J Psychiatry* in press.
7. Council of Australian Governments. *National Action Plan on Mental Health 2006-2011*. Canberra: Department of Health and Ageing, 2006.

8. Rosenberg S, Hickie I. How to tackle a giant: creating a genuine evaluation of the Better Access program. *Australasian Psychiatry* 2010; 18:496-502.
9. Jorm, A.F., Allen, N.B., Purcell, R., Hetrick, S., & Morgan, A. *A guide to what works for depression*. Melbourne, Australia: Beyondblue, 2009.
10. Shafran, R., Clark, D. M., Fairburn, C. G., Arntz, A., Barlow, D. H., Ehlers, A., Freeston, M., et al. Mind the gap: Improving the dissemination of CBT. *Behaviour Research and Therapy*, 2009; 47:902–909.
11. Jensen, A. L., & Weisz, J. R. Assessing match and mismatch between practitioner-generated and standardized interview-generated diagnoses for clinic-referred children and adolescents. *Journal of Consulting and Clinical Psychology* 2002; 70:158–168.
12. Lynch, D., Laws, K. R., & McKenna, P. J. Cognitive behavioural therapy for major psychiatric disorder: does it really work? A meta-analytical review of well-controlled trials. *Psychological Medicine* 2010; 40:9–24.

Footnote.

1. Further to this point, and in the interest of full disclosure, it may be relevant to note that we are both clinical psychologists, although neither of us receives income from the Better Access Scheme, and we have for some time had an open mind about the ultimate value of the scheme, seeing both potential benefits and serious risks.