

FEATURE ARTICLE

# Standardized routine outcome measurement: Pot holes in the road to recovery

**Richard Lakeman**

*Townsville Integrated Mental Health Services, Queensland, Australia*

**ABSTRACT:** *Routine ‘outcome measurement’ is currently being introduced across Australian mental health services. This paper asserts that routine standardized outcome measurement in its current form can only provide a crude and narrow lens through which to witness recovery. It has only a limited capacity to capture the richness of people’s recovery journeys or provide information that can usefully inform care. Indeed, in its implementation nurses may be required to collude in practices or account for practice in ways which run counter to the personal recovery paradigm. Nurses should view a focus on outcomes as an opportunity for critical reflection as well as to seek ways to account for recovery stories in meaningful ways.*

**KEY WORDS:** *Health of the Nation Outcomes Scale, mental health nursing, nursing practice, outcome measurement, recovery.*

## INTRODUCTION

The term ‘outcome’ has only relatively recently infiltrated the everyday discourse of the mental health community. Traditionally, administrators of mental health services have measured outputs (e.g. hospital length of stay or contact time with health professionals) but increasingly nurses and clinicians are being asked to measure ‘outcomes’, that is things which are purported to be the ‘outcome’ of care (e.g. measures of pathology, functioning and aspects of quality of life). Despite much rhetoric ‘talking up’ the benefits of routine standardized outcome measurement (ROM) the field is relatively uncharted and the benefits remain largely speculative.

A recent editorial in the *International Journal of Mental Health Nursing* (Coombs & Meehan 2003) focused on the introduction of ROM in Australia. This process has already commenced in Queensland where the focus of care, ICD-10 diagnosis, age-appropriate Health of the Nation Outcomes Scale (HoNOS), and the Life Skills Profile (LSP-16) is required to be completed for all mental health service users when admitted or discharged from services

and at least every 3 months. The Mental Health Inventory (MHI) may also optionally be given to service users. This paper offers some critical comments about ROM in mental health services and urges that nurses maintain a cautious and critical stance towards this new fashion.

## ROUTINE OUTCOME MEASUREMENT AND INDIVIDUAL DIFFERENCE

The term outcome permeates the fabric of everyday public discourse. Watson (2003; p. 96) states that the term outcome is

‘... a modern portmanteau to hold such words as result, consequence, upshot, product, effect, return and happenstance, which are all subtly different and don’t suggest as outcome does that everything in the world is, or can be, governed by strategies’.

‘Improving outcomes’ in mental health services has become a catch cry. It means subtly and sometimes profoundly different things to different people. All mental health professionals ought to be mindful of the consequences of their involvement with others. However, what the ‘outcomes’ ought to be depends, to a large extent, on what they are needed for and employed to do. For example, a psychiatrist might anticipate that medical treatment will lead to a reduction in ‘psychiatric symptoms’, a crisis nurse might anticipate that a person remains safe until the

---

**Correspondence:** R. Lakeman, C/- Townsville Integrated Mental Health Services, 100 Angus Smith Drive, Townsville, Qld 4814. Email: rlakeman@iqconnect.net.au

R. Lakeman, DipCpNsg, BN, BA Hons, FANZCHMN  
Accepted August 2004.

person is able to resolve their problems, or a case manager may hope to secure stable accommodation for someone who is homeless. For the service user, feeling safe, respected, listened to, and helped will be essential elements and hopefully outcomes of care. There may be some specialized services in which employing standardized measures makes good sense (e.g. an alcohol and drug treatment centre ought to demonstrate a reduction in alcohol or drug-related problems in its targeted population). However, for the most part the shared desired outcome will vary from individual to individual. In the parlance of strategy, outcomes ought to be negotiated in care planning and may well change as new needs emerge and the person's circumstances change.

A serious flaw of ROM is its failure to capture the subtlety of individual difference. A recent national survey in the United Kingdom (Gilbody *et al.* 2002a), found that psychiatrists did not use outcome measures routinely. Indeed, comments were overwhelmingly negative challenging the capacity of tools to capture the subtlety of multifaceted outcome or to describe the individual patient and terms such as 'simplistic' and 'pseudo-scientific gloss' were used to describe them (Gilbody *et al.* 2002b; p. 102). My team had routinely used the HoNOS and the long form of LSP for several years. When viewing the graphed scores over time for some individuals, one can perceive no discernable pattern. It is hard to see how the aggregation of such data if meaningless at the individual level could have any possible usefulness to a service. However, for the same individuals with carefully chosen specific and more sensitive measures one may discern a reduction in specific symptoms. For others even this is futile, as the key problems have revolved around attempting to secure housing, benefits, employment or negotiating acceptable treatment or care. Whilst it may be argued that commonly used outcome measures attempt to tap into these constructs, what they fail to capture is the significance of particular issues for individuals.

## ROUTINE OUTCOME MEASUREMENT AND RESEARCH

A systematic review of large-scale outcome research projects concluded that the '... large-scale studies using "humongous" databases are largely achieved at the expense of clinically meaningful outcomes ...' (Gilbody *et al.* 2002a; p. 14). There are so many possible factors that may influence ratings that changes can rarely be attributed to any specific intervention. Holloway (2002) points out that this failure to control confounding variables can lead to findings such as depressed people who receive treatment for depression having worse outcomes than those who receive no treatment. This is hardly

surprising, however, if the measures chosen have only a tenuous relationship to factors such as illness severity or intention to treat. Measures such as HoNOS have been found to have questionable validity and only a tenuous relationship to illness severity (Adams *et al.* 2000).

Such methodological issues have seldom caused grief for the pharmaceutical industry. Gilbody *et al.* (2002a) point out that there are tremendous advantages for the pharmaceutical industry in mandated routine collection of data. For example, it may claim that expensive randomised controlled trials (RCT) may be dispensed with in order to measure clinical and economic effectiveness; ethical approval and consent may not be required as treatment is received as part of usual care and outcomes are collected routinely; and they may not need to provide and dispense drugs for free to the thousands who would otherwise be enrolled in RCTs. Indeed, ROM provides a boon for researchers of all kinds. ROM databases provide an enticing dataset that may be chosen for convenience at the expense of more sophisticated, sensitive measures, more appropriate to specific research questions. The accessibility of ROM data may even influence the kinds of questions researchers ask, effectively 'dumbing down' clinical research and evading considerations of fundamental ethical questions. As is the case with the burgeoning number of illness registries, and electronic databases kept by universities and health services, debate about issues of informed consent for inclusion or access by researchers has been conspicuously subdued.

Nurses of course do not deal with aggregates they deal with people and what they provide aside from specific technical tasks are varying types of human engagement. It has long been recognized in psychotherapy research that despite very different theories and procedures, different forms of psychotherapy appear to secure comparable outcomes (Garfield 1994). There is growing evidence that it is not the intervention that makes the difference to people but the quality of the relationship. Bambling and King (2001) assert that the working alliance is responsible for up to 30% of all treatment outcomes across psychotherapeutic approaches. A more fruitful way to infer whether a process of care is good or otherwise may be to attempt to measure the quality of relationships. That is, whether or not people feel listened to, understood, respected or helped.

## ROUTINE OUTCOME MEASUREMENT AND IMPROVING OUTCOMES

There is little if any evidence that measuring outcomes actually leads to improvement in outcomes. In a detailed and lengthy report on the field testing of outcome measures in Australia, Stedman *et al.* (1997; p. 102)

concluded that ‘... the research to date suggests that the “simplistic” application of outcomes measures will not benefit service providers and consumers in the long term’. Even when the measures chosen are highly specific to particular populations or problem areas there is scant evidence that measurement helps people or leads to better decision making on the part of clinicians. A recent Cochrane Review found no reliable evidence establishing the value of the routine administration of outcome measures and needs assessment tools in schizophrenia and related disorders (Gilbody *et al.* 2003).

Similar reviews to examine the effect of routinely administered psychiatric questionnaires on the recognition, management, and outcome of psychiatric disorders in non-psychiatric settings have concluded that the exercise is costly and there is little evidence showing any improved psychosocial outcomes (Gilbody *et al.* 2001). This is consistent with earlier reviews (Schade *et al.* 1998) that acknowledge the validity of screening instruments for depression but conclude that there is little evidence of any benefits to those that are screened. For example, Dowrick and Buchan (1995) found that disclosure of cases of unrecognized depression (as indicated by the Becks Depression Inventory) to general practitioners had no effect on patient outcomes. Clearly, screening might be useful if it led to better treatment, and those who have undertaken detailed cost analysis (Valenstein *et al.* 2001) have recommended that resources should first be invested in the area of improving treatment over and before screening.

### ROUTINE OUTCOME MEASUREMENT AND PERSONAL RECOVERY

The term ‘recovery’ is belatedly beginning to take hold in Australia. Professional groups such as social work (Carpenter 2002), occupational therapy (McGruder 2001), nursing (Barker 2003) and psychiatry (Andresen *et al.* 2003) are beginning to challenge the assumptions underpinning positivistic bio-medical approaches to mental health care and promote personal recovery. The facilitation of the right conditions for recovery to take place is the proper focus of nursing and indeed, mental health services (Barker 2003). Nurses have an essential role in supporting people on their recovery journeys (Whitehill 2003). Unlike ‘clinical recovery’, personal recovery refers to a process rather than an end point or outcome. As Helm (2003, p. 50) state, it is ‘... the journeying task of making sense of life itself’.

Coleman (1999) argues that outcomes are a way of sanitizing a non-recovery process rather than measuring a recovery process. This need he argues, is rooted in the psychiatrist’s need to use a medical frame of reference as a means to understand mental distress. Within the

medical model, there is little or no notion of recovery for conditions such as schizophrenia (Andresen *et al.* 2003). Coleman (1999, p. 23) also states, ‘The use of an outcome, biased approach to mental health carries with it the risk of alienating the client from the recovery process through the limited measures used to determine outcomes’. Recovery is a transformative process (Deegan 1996), one that calls us to find meaning in and give value to our experience. The application of clinician-rated standardized tools such as HoNOS or LSP have little, if anything, do with or offer to the recovery process. Indeed, they strip the person’s experience of all meaning and reduce it to predetermined categories. The type of knowledge gleaned from such an exercise is ‘... at best narrow and exclusive, at worst artificial, sterile and inapplicable to everyday realities’ (Keen & Keen 2003).

Coombs and Meehan (2003) suggest that mandatory outcome measurement, as introduced in Queensland, provides an opportunity to identify those that have not progressed as planned and nursing practice can then be motivated by the objective evidence of the measures rather than subjective feelings and perceptions. However, quantification does not make an observation more objective, ‘evidenced based’ or meaningful. Furthermore, evolving recovery-focused approaches to caring with people such as the Tidal Model (Barker 2003), do not presume to know the proper course of people’s lives or exactly how they ought to progress. The focus is instead on the kind of support that people believe they need now to take the next step in their recovery journey.

Concepts such as coping, hope, connectedness, a sense of self-efficacy, empowerment, and self-esteem are likely to be far better indicators of recovery processes than psychiatric symptomatology or indeed any kind of rating scale. Unfortunately, whilst more meaningful, these concepts don’t readily lend themselves to quantification. They are subjectively experienced and defy ‘objective’ rating. The outcomes initiative is part of the positivist dominant discourse stressing the need to standardize, quantify, collate and reduce. Nurses should understand the limitations of this way of viewing the world and not privilege it over other ways of knowing. Nursing should align itself with a recovery agenda that stresses the need to develop wisdom, strive towards understanding, and expand our creativity in understanding and responding to people in distress.

### ROUTINE OUTCOME MEASUREMENT AND OPPRESSION

The outcome initiative may be essentially benign if treated for what it is and only limited significance is ascribed to the measures by nurses and other clinicians. However, it may be positively damaging if nurses in

practice shift their focus from attendance to the person's experience to reducing and measuring it. Little is gained by reducing a person's story to a 12, four-item ordinal scale and further reducing these to behaviours, impairment, symptoms and social functioning (which is required with HoNOS). Such an exercise is akin to asking someone to describe a Picasso in 12 words or less. The meaning, richness and potential are removed, leaving a decontextualized impersonal skeleton that cannot do any justice to the art let alone be described as objective.

The myth of objectivity and certainty pervades psychiatric discourse and indeed the discourses of public life in general. It provides a justification for psychiatric power. The reality of mental illness is that there is little certainty even about the basic constructs such as diagnosis that are claimed to be the central purview of the field. Kendell and Jablensky (2003) point out that there is little evidence of the validity of most currently recognized mental disorders, that is, that they possess discrete natural boundaries that separate one disorder from another. Crowe (2000) reminds us that because a psychiatric diagnosis often fails to describe the individual's experience of mental distress it can be regarded as a categorization process that, while not necessarily intentional, serves to maintain oppressive power relations within society. Barker (2003), Crowe (2000) and a growing personal recovery movement, call for clinicians to recognize the oppressive history and present reality of psychiatric power and to give privilege to the person's narrative.

The outcomes initiative, though likely well-intentioned, draws nurses into a potential web of collusion with psychiatric oppression. Nurses, until recently, could claim to be detached from the diagnostic process and realize roles that enabled them to be a bridge between psychiatric knowledge and the person (Barker *et al.* 1999). The Queensland outcomes protocol, however, requires the entry of an ICD-10 diagnosis before any other measures can be entered. The Queensland directive (Queensland Health 2003) is that

'All mental health clinicians in public mental health services are expected to have, or develop, the skills in assessment of mental health consumers with a view to ongoing planning of care. All mental health clinicians should formulate a provisional diagnosis.'

Many nurses possess the skills to make medical diagnosis and some roles may well call on these skills. Some will not, and nurses ought to be able to choose other methods of problem formulation to plan or inform care. Nurses are being forced into making diagnostic statements in order to ensure that the protocol is followed. Nurses are being required to engage in a process that

many service users find oppressive, which is ethically questionable, and is not what nurses have been found to be needed for (Barker *et al.* 1999). The protocol itself has become the dictator of care and nurses merely instrumental (see Lakeman 2000) to the purposes of others with little regard for how nurses see themselves or their relationships with others.

The proper focus of nursing involves engagement with people in their recovery process. At all stages, nurses require humility and to 'care with' people rather than objectifying the person or their needs. Coombs and Meehan (2003) propose that the inclusion of a consumer rated measure (MHI) in the outcomes battery provides an incentive for nurses to engage the consumer in dialogue around those issues that are important to the consumer. Whilst such tools may be useful, particularly for the more 'psychologically minded', such discussion ought to be central to the nurse-person relationship regardless. Nurses must be careful not to be constrained by a questionnaire. Imposing professional constructions on issues such as quality of life is potentially disempowering and limits possibilities for exploration (Martin 2000).

## ROUTINE OUTCOME MEASUREMENT AND RESOURCE ALLOCATION

Coombs and Meehan (2003; p. 163) claim that poor outcomes at a local level could be used to argue for more resources. However, the need for more resources has been apparent and well articulated prior to the initiation of routine outcome measurement. Indeed, the National Standards for Mental Health (Commonwealth of Australia 1997) from which this project stemmed, requires, amongst other things, that mental health services ensure that supported accommodation is provided and/or supported in a manner which promotes choice, safety and maximum possible quality of life for the consumer. Yet, Federal and State Governments have made dismal commitments to ensure that this standard is met, ensuring that homelessness is a reality for a great many service users. Or that the consumer and the consumer's family/carer have access to a range of safe and effective therapies, yet psychotherapy or non-pharmacological approaches to care remain quite inaccessible to the majority of service users. The setup and compliance costs associated with universal ROM are enormous. The most significant of these is the time lost spending time with people in order to comply with the protocol and the resources diverted from actual service provision (in terms of training, data entry, analysis etc.). These costs must be weighed against the potential (but unproven) benefits to people who use mental health services.

## CONCLUSIONS

Routine outcome measurement in its present form cannot capture individual differences with any clinically useful sensitivity. The outcome of ROM is an aggregation of data that has limited validity but which may provide an enticing pool of data for analysis by researchers. However, researchers, nurses and managers must be extremely cautious that the protocol does not deflect them from asking the important research questions or impede the development or application of more meaningful ways of evaluating or accounting for care. The ethics of consent and utilitarian questions about resource allocation must keep pace with expenditure on such projects.

Nurses must maintain a healthy skepticism regarding the validity of brief clinician-rated scales – that because they are called ‘outcome measures’ that they necessarily indicate an outcome of care. The language of ‘outcomes’ and the ‘tools’ used to measure them shape reality. They hint at a uniformity, certainty and predictability in the sphere of mental distress and care, which is seductive and simplistic. However, nurses must not be seduced into blindness to the uniqueness of people’s experience. Their vision ought to remain on attending to the process of personal recovery rather than some predetermined end point. Nurses need to continually ask service users and families what would be good for them (acknowledging that this may shift and change). A data collection protocol should not drive the process of care.

Routine outcome measurement may satisfy calls for ‘accountability’ but this should not be at the expense of imagination, courage, initiative, reflection or generosity which Watson (2003; p.96) suggests have been priceless qualities for getting at the truth (until the advent of the term and which he also suggests spurned the preoccupation with outcomes). These are precisely the qualities that nurses and others need today to assist people in their recovery journeys. Coombs and Meehan (2003; p. 164) assert that nurses should welcome the current focus on outcomes and caution that it would be a mistake ‘... for nurses to simply see the collection as a bureaucratic exercise’. There is a need to ‘talk up’ the collection of routine outcome measurements as the process has been found to be disliked and considered irrelevant by most clinicians (Gilbody *et al.* 2002b). However, critical debate and reflection should accompany its introduction rather than good marketing. Nurses need to take advantage of opportunities to use the data, which they enter and collect, to achieve their purposes. However, it would be better to view the use of standardized outcome measures as simply another bureaucratic exercise than to make the mistake of considering that the selected measures

represents a cogent or particularly useful view of reality for nurses or others.

## REFERENCES

- Adams, M., Palmer, A., O’Brien, J. T. & Crook, W. (2000). Health of the nation outcome scales for psychiatry: Are they valid? *Journal of Mental Health*, 9, 193–198.
- Andresen, R., Oades, L. & Caputi, P. (2003). The experience of recovery from schizophrenia: Towards an empirically validated stage model. *Australian and New Zealand Journal of Psychiatry*, 37, 586–594.
- Bambling, M. & King, R. (2001). Therapeutic alliance and clinical practice. *Psychotherapy in Australia*, 8, 38–43.
- Barker, P. (2003). The tidal model: Psychiatric colonization, recovery and the paradigm shift in mental health care. *International Journal of Mental Health Nursing*, 12, 96–102.
- Barker, P., Jackson, S. & Stevenson, C. (1999). What are psychiatric nurses needed for? Developing a theory of essential nursing practice. *Journal of Psychiatric and Mental Health Nursing*, 6, 273–282.
- Carpenter, J. (2002). Mental health recovery paradigm: implications for social work. *Health and Social Work*, 27, 86–94.
- Coleman, R. (1999). *Recovery An Alien Concept*. Gloucester: Handsell Publishing.
- Commonwealth of Australia (1997). *National Standards for Mental Health*. Canberra: Commonwealth of Australia.
- Coombs, T. & Meehan, T. (2003). Mental health outcomes in Australia: Issues for mental health nurses (Editorial). *International Journal of Mental Health Nursing*, 12, 163–164.
- Crowe, M. (2000). Psychiatric diagnosis: some implications for mental health nursing care. *Journal of Advanced Nursing*, 31, 583–589.
- Deegan, P. (1996). Recovery as a journey of the heart. *Psychiatric Rehabilitation Journal*, 19, 91–97.
- Dowrick, C. & Buchan, I. (1995). Twelve month outcome of depression in general practice: does detection or disclosure make a difference? *British Medical Journal*, 311, 1274–1276.
- Garfield, S. (1994). Major issues in psychotherapy research. In: A. E. Bergin & S. L. Garfield (Eds). *Handbook of Psychotherapy and Behavior Change*, 4th edn. (pp. 335–359). New York: J. Wiley.
- Gilbody, S. M., House, A. O. & Sheldon, T. A. (2001). Routinely administered questionnaires for depression and anxiety: systematic review. *British Medical Journal*, 322, 406–409.
- Gilbody, S. M., House, A. O. & Sheldon, T. A. (2002a). Outcome research in mental health: Systematic review. *British Journal of Psychiatry*, 181, 8–16.
- Gilbody, S. M., House, A. O. & Sheldon, T. A. (2002b). Psychiatrists in the UK do not use outcomes measures: National survey. *British Journal of Psychiatry*, 180, 101–103.
- Gilbody, S. M., House, A. O. & Sheldon, T. A. (2003). Outcome measures and needs assessment tools for schizophrenia and related disorders. *Cochrane database of systematic reviews Online Update Software* (1), CD003081.

- Helm, A. (2003). Recovery and Reclamation: a Pilgrimage in Understanding Who and What We Are. In: P. Barker (Ed). *Psychiatric and Mental Health Nursing: The Craft of Caring*. (pp. 50–55). London: Arnold.
- Holloway, F. (2002). Outcome measurement in mental health – welcome to the revolution. *British Journal of Psychiatry the Journal of Mental Science*, 181, 1–2.
- Keen, T. & Keen, J. (2003). Developing Collaborative Assessment. In: P. Barker (Ed). *Psychiatric and Mental Health Nursing: The Craft of Caring*. (pp. 87–96). London: Arnold.
- Kendell, R. & Jablensky, A. (2003). Distinguishing between the validity and utility of psychiatric diagnosis. *American Journal of Psychiatry*, 160, 4–12.
- Lakeman, R. (2000). *Nurses Are More Than Tools: Instrumentality and Implications for Nursing Ethics*. Proceedings of The Mental Health Nurses for a Changing World: Not Just Surviving; 3–7 September 2000, Broadbeach, Queensland.
- Martin, P. J. (2000). Hearing voices and listening to those that hear them. *Journal of Psychiatric and Mental Health Nursing*, 7, 135–141.
- McGruder, J. (2001). Life experience is not a disease or why medicalizing madness is counterproductive to recovery. *Occupational Therapy in Mental Health*, 17, 59–80.
- Queensland Health (2003). *Outcomes Initiative: Information Sheet 2 – Adult*. Brisbane: Systems and Outcomes Team.
- Schade, C. P., Jones, E. R. & Wittlin, B. J. (1998). A Ten-Year Review of the Validity and Clinical Utility of Depression Screening. *Psychiatric Services*, 49, 55–61.
- Stedman, T., Yellowlees, P., Mellsop, G., Clarke, R. & Drake, S. (1997). *Measuring Consumer Outcomes In Mental Health: Field Testing of Selected Measures of Consumer Outcome in Mental Health*. Canberra: Department of Health & Family Services.
- Valenstein, M., Vijan, S., Zeber, J. E., Boehm, K. & Buttar, A. (2001). The Cost–Utility of Screening for Depression in Primary Care. *Annals of Internal Medicine*, 134, 345–360.
- Watson, D. (2003). *Death Sentence: The Decay of Public Language*. Sydney: Random House.
- Whitehill, I. (2003). *The Concept of Recovery*. In: P. Barker (Ed). *Psychiatric and Mental Health Nursing: The Craft of Caring*. (pp. 43–55). London: Arnold.