Can mental health outcome measures be developed which meet the following three criteria: (1) standardised, (2) acceptable to clinicians, and (3) feasible for ongoing routine use? We shall argue that the answers at present are “yes”, “perhaps”, and “not known”, respectively.

Standardised ratings of outcomes of interest to patients and clinicians which go beyond symptom severity have been available for almost two decades. Some are composite global ratings, which include disability, while most measure specific domains of patient or carer function such as quality of life, needs, or satisfaction. When assessed against an array of psychometric characteristics, many of these scales have been shown to be well constructed in terms of their validity, reliability, and sensitivity. There is strong evidence therefore that the first criterion has been satisfied.

Are at least some of these scales acceptable to clinicians rather than to researchers? Here the evidence is less compelling. In this issue of Quality in Health Care McClelland and colleagues present results to show that one particular scale, developed in the context of the Health of the Nation initiative of the last government, is not only reasonably sensitive and valid when used by staff in ordinary clinical settings, but also that in eight of 10 study sites in Northern Ireland it was found to be relevant to their routine clinical tasks. Nevertheless, at the beginning of each episode of care they reported that the Health of the Nation Outcome Scale (HoNOS) was less relevant than at the end of treatment, as the information necessary for its completion was often unavailable until a full assessment had been carried out.

Previously, a study in South London showed that, with one brief training session, multidisciplinary staff in a community mental health team could use the Global Assessment of Functioning Scale in a reliable way that had predictive validity. Even so, the use in routine settings of assessment scales of known psychometric adequacy has taken place more often for particular diagnostic conditions (using specific scales to rate the severity of, for example, depression or anxiety) than it has for whole treated patient populations. The second criterion has therefore been met only in part.

Thirdly, is it common for staff in routine clinical psychiatric practice to use standardised assessments? The evidence here is far less convincing. Despite the professional drive to practise in accordance with evidence based medicine, and the requirements of governmental policy to meet new standards with a National Service Framework for Mental Health, the literature about the use of standardised scales on an ongoing routine basis, rather than for a time limited study period, is largely absent. Only recently have operational definitions of “feasibility”, for example, been proposed for outcome scales. Slade and colleagues propose that a feasible scale should be brief, simple, relevant, acceptable, available, and valuable. For a scale to be valuable, certain characteristics of a healthcare system are required, including the infrastructure to support the collection, analysis and feedback of outcome data, and professional and managerial expectations that routinely collected data will shape clinical care and subsequent service plans. It follows that implementing the routine use of outcome measures is a complex task involving the characteristics of the scales, the motivation and training of staff, and the wider clinical and organisational environment.

By comparison, Andrews et al have identified three dimensions of feasibility—namely, applicability, acceptability, and practicality. The “acceptability” of a measure is the degree to which it addresses dimensions of importance to the rater. Acceptability of a measure describes the ease with which a rater can use a particular measure. Practicality refers to the training requirements and the complexity of scoring, reporting and interpreting the data. When assessed using these criteria it is clear that our current knowledge tells us more about barriers to implementing routine outcome measures than about the necessary and sufficient ingredients for their successful translation into clinically meaningful everyday use.