

Liberating the NHS: Transparency in outcomes. A framework for the NHS

The Nuffield Trust welcomes the proposed moves towards better measurement of health outcomes, and we support the increased use of such information in the planning and management of health services. By measuring outcomes rather than processes, it may be possible to assess better the underlying benefits derived from investing in healthcare. Moreover, the use of outcome measures may allow organisations to be held to account for the health benefits they achieve rather than for the processes they undertake.

The principle of using outcome measures is undoubtedly appealing. However, the reality is highly complex so we would stress that the proposed framework should to be developed carefully and cautiously. In particular, we believe that outcome measures should be seen as an accompaniment to a range of other metrics not a replacement. The relative merits of outcome measures and process measures have been debated for many years. A consensus has emerged suggesting that both types of measure are important and that they should therefore be viewed as being complementary to each other (Donabedian, 1988; Mant and Hicks, 1995; Lilford et al, 2004).

Any short list of proposed indicators will have significant limitations given the complexity of what is being measured. However, we feel that such limitations are inevitable with outcome measures, and that they should be seen as the next step in a journey rather than end in themselves. For this reason, we would advocate building a suite of measures that capture a much wider range of different perspectives. These will be both qualitative and quantitative, process measures and outcome measures including patient-reported outcomes. Such an approach will encourage open scrutiny and

debate particularly where a particular service performs differently according to different measures.

We suggest the greatest challenges will be to:

- Identify outcome measures that are sufficiently broad to capture the experiences of many patients yet specific enough to be attributable to the effects of health care rather than other confounding factors.
- Find outcome measures that are amenable to change in reasonable time scales. Where such outcome measures do not exist, it will be important to acknowledge that certain process measures are legitimate and valid. Indeed, process measures are often more valuable than outcome measures which take a long time to be changed.
- Selected outcome measures are useful at a range of at different organisational levels – not simply at national level but also at the local and provider levels. A clear understanding is required about how the measures will be used for accountability purposes.
- Develop ways of exploiting existing routine information. Many operational information systems capture information about a person's health status, albeit imperfectly – and by linking information over time this can form the basis of many outcome measures.

Although many of the indicators fall short of an ideal, we are confident that they can be developed and refined over time and that their scope can be extended. The level of data quality and completeness is an important consideration when deciding which indicators to use. The initial sets of indicators need to be designed in ways that recognise the current state of NHS data and information systems.

Scope, purpose and principles of an NHS Outcomes Framework

The measurement of healthcare outcomes has a long and distinguished history, with its pioneers including Florence Nightingale and Ernest Codman. Over the past twenty years there has been much debate about how outcome measures can be incorporated into the running of health services (Brook 1977; Fries, 1983; Charlton et al, 1984; Bardsley and Coles, 1992) Although there have been some notable successes overall, it is fair to say that progress in this field has been rather slow.

Part of the reason for the apparent lack of progress is that there are several important challenges in developing outcome measures. In the following sections we outline the challenges and propose some solutions.

There can be multiple causes of any given outcome: In a complex system such as healthcare, it is not always clear what actions did or did not lead to a given outcome (Donabedian, 1988; Lilford et al, 2004; Bardsley and Coles 1992). This generates particular problems if rewards are given for achievements that might in fact be attributable to some external factor, or indeed when punitive actions result from adverse outcomes that were actually beyond the control of the organisation being judged (Spiegelhalter, 2005). This concern is especially pertinent for very broad indicators, such as premature mortality, which may change as a result of factors unrelated to health services. For example, the

current reduction being seen in the rates of lung cancer deaths is largely the results of changes in smoking behaviours 10-20 years ago (Westlake and Cooper, 2008).

Solution: Include in selection criteria the need to demonstrate that there is evidence to support how that specific indicator can change as a result of differences in health care delivery (Lakhani et al, 2005). This in itself is no easy task and is dynamic and so what is need is a function that continuously appraises the validity and utility of indicators, adapts and amends them over time as new evidence arises. The mindset is not to produce a pre-ordained list but be alive to changes. These types of indicators often raise as many questions as answers – they have been described as ‘tin openers’ rather than ‘dials’ and so might be best used to trigger some more targeted analysis - rather than making a final judgement.

Outcomes are multi-dimensional: There are, quite legitimately, a range of different perspectives as to what constitutes a good outcome (Fries, 1983; Fitzpatrick et al, 1992; Bowling, 2002, Black and Jenkinson, 2009). Good outcomes may be clinical metrics (e.g. degree of revascularisation of a coronary artery), a prognostic indicator (e.g. change in expected 10-year coronary mortality), the patient’s experience of care (satisfaction with cleanliness, compassion of staff etc.) and the patient’s experience in change of symptoms (e.g. improvement in angina symptoms). To be comprehensive, outcome measurement needs to recognise all of these different dimensions. This multi-dimensional approach is especially important in the assessment of interventions that have implications for both health and social care. The danger here is that health needs will dominate the analysis to the exclusion of information about a person’s basic social needs.

One of the key concerns for the NHS Commissioning Board and GP commissioners will be whether these outcome measures are capable of capturing the monitoring of the most important aspects of good quality care. Within the past decade, several major investigations have identified serious lapses in the quality of care delivered by certain NHS institutions (Healthcare Commission, 2008a). In certain cases, the use of outcome measures alone may not have identified the problems concerned.

Solutions: Though the framework does recognise these issues there is scope for widening the spectrum of indicators that are developed. The concept of a ‘balanced’ set of indicators could be expanded so that, for example, it makes more explicit recognition of the range of dimensions of health and care that the framework seeks to addresses. In terms of the types of metrics for example, the framework might usefully include more direct intelligence from patients perspective (via surveys of the population), or from selected clinical markers ; or from a societal perspective on appropriate care and include markers about serious lapses in terms of patient dignity and respect e.g. finding from complaints and investigations.

We suggest that a broader perspective may then accommodate collection and analysis of information on:

- *Unfair limitations on access to care and extreme waiting times*
- *Treatment of patients with dignity and respect*
- *High-risk provider behaviours*
- *Extreme inequalities in health and health care*
- *Investment in long term preventive health measures (e.g. immunisation, contraception)*
- *Markers of healthy lifestyles and behaviours e.g. drug misuse problems, binge drinking.*

For some measures the timescales for change can be very long: The delay between receiving a healthcare intervention and experiencing an ensuing change in outcomes can

be very protracted. This is particularly the case for interventions relating to chronic diseases. In the short run, there may be no observable improvement in outcomes, despite the fact that the intervention will improve outcomes in the long run. Much of the literature on outcome measures is based on studies of short-term curative interventions, and therefore it may be less applicable to long term conditions. Another issue is the problem of population movement or “churn” where a commissioner invests in preventive campaigns, only for the future benefits to be enjoyed in another part of the country where the patient moves home (Millet et al, 2002).

Detection lags: In addition measures need to overcome the time lag between poor performance and detection. If an outcome measurement is to be reliable, then the required information must be collected, collated and processed in sufficient quantities to produce a robust conclusion. All of this processing takes time, and so it may limit the ability to detect major adverse outcomes (Bottle and Aylin, 2008; Healthcare Commission, 2008b). The danger is that major problems may be identified too late.

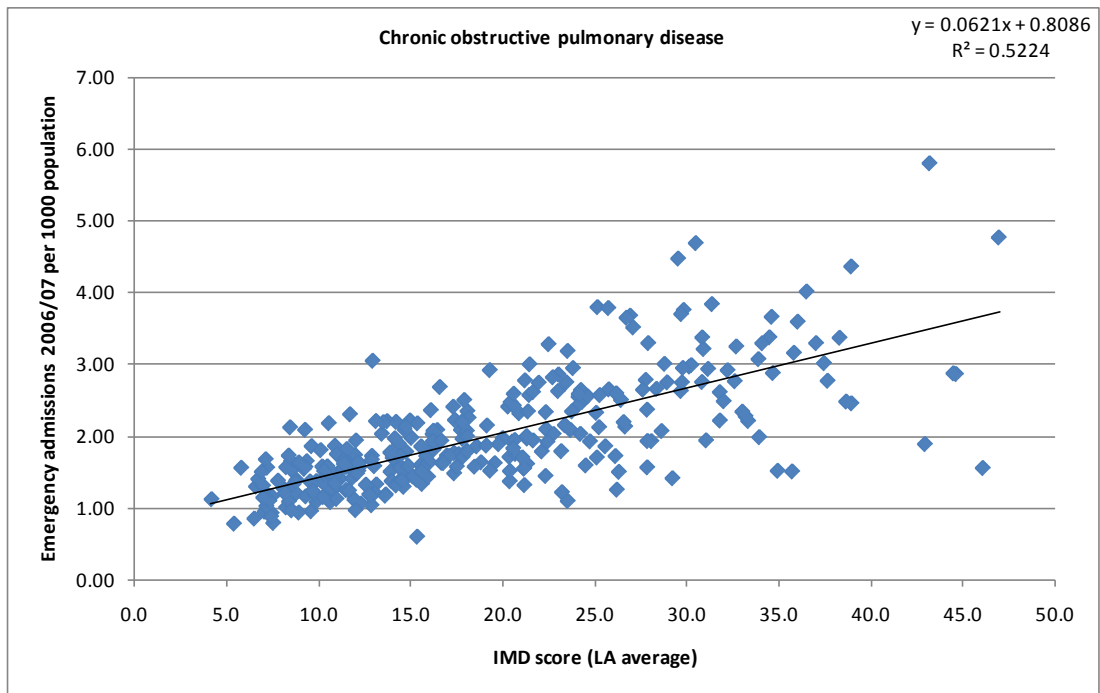
Solutions:

- ***Using information intelligently as it emerges***, for example by exploiting cumulative analytical techniques such as cumulative sum control charts (CUSUMS). A recent example of the usefulness of this approach occurred in 2008, where analysts at the national transplant centre generated an alert about the case fatality rate for heart transplants conducted at one hospital. Transplants were temporarily suspended at the hospital whilst officials from the Healthcare Commission conducted an investigation. A few weeks later, following a swift investigation, services were restarted with a clear understanding about the expected patterns of outcomes. (Healthcare Commission, 2007)
- ***Establishing mechanisms to exploit additional intelligence***. This could include qualitative information, and in particular intelligence from service users and staff. Procedures should be developed to ensure that such information can be processed in a systematic way (Bardsley et al, 2009).
- ***Include process measures***. Process measures have an important role to play (Lilford et al, 2007), particularly in the assessment of interventions that affect longer term outcomes. : Include short/medium term process measures that are indicative of long term outcomes. This would include uptakes of preventive programmes e.g. breast cancer screening, immunisation rates and some clinical quality markers e.g. inhaled steroids for asthma.

Establishing reasonable expectations: The idea of international comparability is potentially useful but getting comparable data is fraught with difficulty and in some cases is not possible - for example on PROMS measures. Evidence from international comparisons should be regarded as just one approach to the wider issue of how to determine what should be expected for a given outcome indicator.

Standardising for pre-existing health states: When considering what causes a good or bad outcome, it is important to adjust for pre-existing health status to avoid a situation whereby the best outcomes are associated with the healthiest patients prior to treatment. Often risk adjustment of some form is necessary and needs to be tailored to the individual indicators and the circumstances in which it is used. For example, the proposed use of ambulatory care sensitive conditions as an outcome measure (Domain 2) raises the question of whether these indicators should be standardised to account for known differences by socioeconomic status across areas. The figure below shows the strong correlation between emergency admission rates for chronic obstructive pulmonary

disease (COPD) by local authority and deprivation scores as measured by the Index of Multiple Deprivation (IMD). This example prompts questions about whether outcomes for this indicator should be adjusted for age, sex, deprivation and/or co-morbidity.



Source: Nuffield Trust analysis of HES data

Solutions: One approach is to use a common framework of comparing observed and expected values for each indicator in such a way that the definition of 'expected' levels can be tailored for the measure. The 'expected value' needs to be risk-adjusted but can be derived in several ways including (a) agreed norms based on expert opinion, (b) national averages, (c) averages of similar 'benchmark' geographical areas, (d) historical values, and (e) international evidence (Bardsley et al 2009).

Ensuring adequate content and coverage: Outcome measures, ideally, need to be unambiguous and should draw on information that it is available or can readily be collected. There will inevitably be a tendency towards focusing on the measurement of outcomes on specific patient sub-types. Whilst this makes the indicator more valid for those patients, the danger is that outcome measures will only be recorded on a small, sub-set of all patients receiving care. One example is that the introduction of routine PROMS is limited to a small set of elective procedures, rather than the much wider and more challenging problems of chronic diseases.

Solutions: One solution to this problem is to use a much broader array of indicators with an explicit aim to capture as many care groups as possible. Indeed, this ought to form a guiding principle for the development of the outcome framework. In particular, the importance of chronic disease needs to be recognised where the challenge will be in assessing how a person maintains their health status rather than on measuring short-term improvements following acute or elective care.

Fitness for purpose: The outcomes framework needs to address how the different levels at which different outcome measures may be used. Whilst the paper describes the framework from the high-level perspectives of Ministers and the NHS Commissioning Board, it should be recognised that the framework also needs to have applicability at lower organisational levels – in particular at the level of the newly emerging GP

consortia. Outcome measures will have an important role to play at the interfaces between the NHS Commissioning Board, the regulators, commissioners, providers and local government. A second consideration concerns the need to be clear about the consequences to providers and commissioners of good and bad performance on different outcome indicators. We believe that consideration of the applications is an important part of designing appropriate measurement tools, as clearly the key consideration is that any measure should be fit for its intended purpose (Lilford et al, 2007; Jenkinson et al, 1994; RSS, 2003).

Solutions: Develop the framework to indicate the types of actions that may flow from extreme values - this will usually mean poor results. Specify the organisation level that might be appropriate for considering the reasons behind poor performance and recognise that the performance characteristics of an indicator are linked to its purpose.

Limited information on health: The reason for measuring outcomes is to detect changes in health status that can be attributed to antecedent health care. Ideally, routine information systems would capture information about health status and then outcome measures would compare changes in health status over time. Such an arrangement would generate a rich flow of operational intelligence useful to clinical staff and provide the basis for monitoring results (Bardsley and Coles, 1992; Lakhani et al, 2005). In practice, however, certain types of important information are not collected consistently or reliably and analysts' abilities to extract and link records at individual level are limited.

Solutions: Linking together operational and administrative data sets has huge potential to inform our understanding of care (Roos et al, 2008; Bardsley et al, forthcoming). This means that in the short term, simply accessing information about outcomes will incur a significant cost. This investment will be critical but we believe it will be worthwhile. In particular, we would emphasise how improvements in data quality and data linkage allow the health needs, services and outcomes to be tracked over time at the individual level.

Simplicity, complexity and transparency: Often the search for transparency and accountability leads to approaches that in information terms are very simple. So for example in star ratings, performance of a multi-million pound organisation delivering thousand of different services is boiled down to a four point scale – itself the sum of a handful of indicators. Similarly, the tendency to use league tables – and the problems they cause have been well documented (Mckee and Hunter 1995). Limiting the number of indicators limits the range of outcomes that can be considered. On top of this is the vital need to ensure that measures incorporate ways to detect random statistical effects (Spiegelhalter, 2005; RSS, 2003).

Solutions: Modern information systems mean that processing and presenting complex data sets is getting easier. We suggest that rather than artificially restrict the range of information, recognise diversity and develop approaches that can summarise and extract key messages from multiple indicator systems. Similarly, the approach to structuring the indicators could be simpler and more fluid.

Comments on individual indicators

We would suggest that the right approach in developing the outcomes framework is to embed the right mechanisms for identifying, measuring and using outcome measures. The definitions of the indicators themselves are important, but should not be seen as permanent but rather a starting point for a programme. In choosing and developing indicators it may be useful to check whether there are already existing metrics that could be incorporated as part of a wider set, for example standard indicators drawn from work by the OECD (OECD, 2010).

On the specific question of the structure of the NHS Outcomes Framework there is the problem that ideally, an overarching indicator would in some sense represent the summation of more detailed indicators below. As it stands, we are not confident that the existing framework will permit this. An alternative option, that would be simpler, would be to use an array of metrics, using a common method summarising achievement across many indicators. In fact this approach could summarise indicators in different ways to address different themes e.g. summaries for older versus younger people etc. We suggest the emphasis is on ways of summarising results simply and clearly across a range of many different indicators and not to strive for a small number of indicators.

Domain 1 – Preventing people from dying prematurely

- Whilst mortality indicators are easily measured they do have the problem that they are weak in that they are too uncertain and too remote to be appropriate as markers of change in the healthcare system. This is especially so for all cause mortality rates and we suggest markers of more specific situations be used e.g. mortality within 30 days of elective surgery; years of life lost for specific chronic conditions.
- The framework currently lacks indicators that describe the prevalence/incidence of certain conditions with potentially high mortality (e.g. some cancers) yet these data are available.
- The uptake of some health screening programmes, especial breast and cervical cancer, could be added as a more direct marker of the impact that health services can have.
- There are currently no indicators of health and lifestyle factors that are known to be strongly predictive of health outcomes (e.g. smoking prevalence). These are important markers of health outcomes in the future.
- The level of mortality rates from accidents/injury is an important marker that should be considered – especially in relation to younger people and children.

Domain 2 – Enhancing the quality of life for people with long-term conditions

- There is little about outcomes for people with mental health problems yet in terms of the scale of health problems this group should be especially important. Measurement for this group is not easy but we would suggest a mix of indicators (National Social Inclusion Programme, 2009) of adverse events e.g. suicide, admissions indicating self harm together with markers of successful treatments e.g. successful discharge from treatment.
- The proposal to use a generic PROM for long term conditions is ambitious. The problem is to calibrate the right measurement tool for these conditions. We believe this is an important area for research and development work. In the

meantime we suggest using measurement based on the care history for cohorts of people with long term conditions. For example, hospital admissions for complications of diabetes for people (within aged x to y) diagnoses at least 5 years previously. Such metrics are not routinely used at present but could be developed through linkage of existing operational data. These measures would have to assume that emergency hospital admission was in general (*ceterus paribus*) undesirable for people with long term conditions.

Domain 3 – Helping people to recover from episodes of illness or following injury

- The indicator on bed day use for people with acute admissions could be extended to look more generally at the patterns of readmission relative to a risk adjusted expectation (Billings et al, 2006).
- For people of working age, we would encourage the use of indicators that seek to measure whether people are able to regain employment after serious illness, for example through the linkage of anonymised data on employment benefits – something that has not been undertaken (to our knowledge) yet is technically possible.

Domain 4 – Ensuring people have a positive experience of care

- We welcome the proposal to continue to develop information drawn from reported patient experiences. The challenge now will be to use this information to inform decision making.

Domain 5 – Treating and caring for people in a safe environment and protecting them from avoidable harm

- Indicators based on incidents reported to the NPSA are likely to be very unreliable. This is partly because of differences in reporting systems (National Social Inclusion Programme, 2009). The danger is that an area with an apparently high incidence of adverse incidents is in fact just one with better reporting practice. Instead we suggest concentrating on major events that should not happen, such as major failings, investigations, court proceedings etc.

References

- Bardsley M., Georghiou T. and Dixon, J., *Social Care Use in the final months of life*. London, Nuffield Trust, Forthcoming
- Bardsley M., Spiegelhalter D.J., Blunt I., et al (2009) Using routine intelligence to target inspection of healthcare providers in England. *Quality and Safety in Healthcare* 18, pp. 189-194
- Bardsley, M. and Coles J. (1992) Practical experiences in auditing patient outcomes. *Qual Saf Health Care* 1, pp. 124-130
- Billings J., Dixon J., Mijanovich T., et al (2006) Case finding for patients at risk of readmission to hospital: development of algorithm to identify high risk patients. *BMJ* 333(7563) pp. 327
- Black, N. and Jenkinson, C. (2009) Measuring patients' experiences and outcomes. *BMJ* 339, pp. 2495
- Bottle, A. and Aylin, P. (2008) Intelligent Information: a national system for monitoring clinical performance. *Health Services Research* 43, pp. 10-31
- Bowling, A. (2002) *Measuring Disease. A review of Disease Specific Quality of Life Measurement Scales*. 2nd ed., Open University Press, Buckingham
- Brook, R.H. (1977) Quality – Can We Measure It? *N. Engl. J. Med.* 296, pp.170-172
- Charlton J., Bauer R. and Lakhani A. (1984) Outcome measures for district and regional health care planners. *Community Med* 6, pp. 306–15.
- Donabedian, A. (1988) The quality of Care. How can it be assessed? *JAMA*;260:1743-1748
- Fitzpatrick R., Fletcher A., Gore S., et al (1992) Quality of life measures in health care. I: Applications and issues in Assessment. *BMJ* 305, pp. 1074-7
- Fries, J. (1983) Toward an Understanding of Patient Outcome Measurement. *Arthritis & Rheumatism*; 26, pp. 697-704.
- Healthcare Commission (2008a) *Learning from Investigations*.
http://www.cqc.org.uk/_db/_documents/Learning_from_investigations.pdf. Last accessed 6/10/10
- Healthcare Commission. (2008b) *Following up mortality outliers*.
http://www.cqc.org.uk/_db/_documents/Following_up_mortality_outliers_200903244704.pdf Last accessed 10/10/10
- Healthcare Commission (2007). *Intervention at Papworth Trust*. Healthcare Commission, http://www.cqc.org.uk/_db/_documents/Papworth_intervention_report.pdf last accessed 6/10/10
- Jenkinson C., Bardsley M. and Lawrence K. (1994) Do Issues of Measurement Matter? Chapter in Jenkinson C (ed). *Measuring Health: Questionnaire Design and Application*. London, UCL Press,
- Lakhani A., Coles J., Eayres D., et al. (2005) Creative use of existing clinical and health outcomes data to assess NHS performance in England: Part 2—more challenging aspects of monitoring. *BMJ* 330, pp.1486–92.
- Lilford R., Brown C.A. and Nichol J., (2007) Use of process measures to monitor the quality of clinical practice. *BMJ* 335, pp. 648 doi: 10.1136/bmj.39317.641296.AD (Published 27 September 2007)

Liberating the NHS: Transparency in outcomes. A framework for the NHS

Lilford R., Mohammed M. A., Spiegelhalter D., Thomson R. (2004) Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. *Lancet* 363, pp.1147-54

Mant, J. and Hicks, N. (1995) Detecting differences in quality of care: the sensitivity of measures of process and outcome in treating acute myocardial infarction. *BMJ* 311, p. 793-6

Mckee, M. and Hunter, D. (1995) Mortality league tables: do they inform or mislead? *Qual Health Care* 4, pp. 5-12, doi:10.1136/qshc.4.1.5

Millet C., Bardsley M. and Binysh K. (2002) Exploring the Effects of Population Mobility on Cervical Screening Coverage, *Public Health* 116, pp. 353-60.

National Social Inclusion Programme. (2009) Outcomes Framework for Mental Health Services 2009.

http://www.socialinclusion.org.uk/publications/Broadened_Social_Inclusion_Outcomes_Framework.pdf

OECD Health data 2010: Overview.

<http://www.ecosante.fr/index2.php?base=OCDE&langs=ENG&langh=ENG&valeur=&source=1>

NPSA (2007). *Safer care for the acutely ill patient: learning from serious incidents*. Patient Safety Observatory. National Patient Safety Agency, January 2007.

RSS (2003) *Performance indicators: good, bad, and ugly* [The report of a Working Party on Performance Monitoring in the Public Services chaired by Professor S. M. Bird, submitted on October 23rd, 2003] (2005) *J. R. Statist. Soc. A* 168, Part 1, pp. 1–27

Roos L.L., Brownell M., Lix L., et al (2008) From health research to social research: privacy, methods, approaches. *Soc Sci Med.* 66(1) pp. 117-29

Spiegelhalter, D. (2005) Funnel plots for institutional comparisons. *Stat Med* 24, pp. 1185–202.

Westlake S. and Cooper, N. (2008) Cancer incidence and mortality: trends in the United Kingdom and constituent countries, 1993 to 2004. *Health Statistics Quarterly* 38, pp. 33-46.

For more information about Nuffield Trust,
including details of our latest research and analysis,
please visit www.nuffieldtrust.org.uk



Download further copies of this consultation response
from www.nuffieldtrust.org.uk/publications



Subscribe to our newsletter:
www.nuffieldtrust.org.uk/newsletter



Follow us on Twitter: [Twitter.com/NuffieldTrust](https://twitter.com/NuffieldTrust)

Nuffield Trust is an authoritative
and independent source of
evidence-based research and
policy analysis for improving
health care in the UK

59 New Cavendish Street
London W1G 7LP
Telephone: 020 7631 8450
Facsimile: 020 7631 8451
Email: info@nuffieldtrust.org.uk



www.nuffieldtrust.org.uk