King's Fund response to:
Assessment of the Performance of Healthcare Organisations
A Healthcare Commission consultation

This paper is a formal response by the King's Fund to the Healthcare Commission consultation on the *Assessment of the Performance of Healthcare Organisations*. The King's Fund is an independent charitable foundation working for better health, especially in London. We carry out research, policy analysis and development activities, working on our own, in partnerships, and through grants. We are a major resource to people working in health, offering leadership and education courses; seminars and workshops; publications; information and library services; and conference and meeting facilities.

Overview

The goals which the Healthcare Commission has set for the assessment process are very ambitious. They involve meeting the needs of different constituencies – patients, clinicians, managers and the public at large – while at the same time keeping the burden of inspection and data collection to a minimum.

As the Consultation notes in respect of the benefits and costs of discharging its role (p16), there remain a number of technical issues to be resolved before the principles as a whole can be implemented with confidence. Until they are resolved, the process will run the risk of appearing and being arbitrary and of failing to meet the goals that the Healthcare Commission has set itself.

For these reasons we believe that it would be advisable for the Healthcare Commission to set realistic targets for the process of developing the assessment process. It should aim to give sufficient time to allow for the substantive technical work which remains to be done and for the type of learning which can only be acquired through practical experience of how any new system works in practice.

Unless the Healthcare Commission adopts such an approach, there is a risk that it will appear to promise more than it can deliver, with consequent risks for its own reputation and, more broadly, for the process of external assessment itself.

Introduction

The response below considers briefly some of the technical issues which are inherent in any assessment process of organisational performance in the health field. These issues are not new and the fact that they have been long recognised and remain imperfectly addressed reflects the difficulties and challenges of *any* inspection and assessment process in a field such as health with its multiple goals and its mixture of the measurable, the non-measurable, the commensurate and the non-commensurate.

We recognise that the general framework within which the Healthcare Commission has to work, in particular, the two categories of standards published by the Department of Health and the requirement placed upon it to publish annual performance ratings for each health care organisation within its remit. This framework, however, does pose a number of difficulties since any overall performance measures for large health care organisations risk hiding as much as they reveal.

Aggregation

The Healthcare Commission is committed by its terms of reference to awarding annual performance ratings. Given the diversity of activity and objectives in NHS organisations, it is unlikely that any will perform well across the board. Annex 5 of the Consultation Paper sets out two options, both of which have some plausibility. However, by their nature they lose information through whatever arithmetical process is used to aggregate the elements.

The same is true of the process of coming to a judgement on particular standards/domains and their constituent elements. The consultation adopts the domains set out in the Department of Health paper *Standards for Better Health* within which standards are to be met. However, each domain comprises a range of elements. This means that a range of measures must be combined into a single one to form a judgement as to whether a hospital is 'safe'. But hospital activity gives rise to a range of different risks to its patients; similarly a wide range of policies are relevant to the promotion of a safe environment. The consultation does not make it clear how this aggregation is to be done and hence how in practice it will be determined whether a trust has reached an 'adequate (i.e. not provoking intervention) level of performance.

We consider that neither good nor bad performance in an element should be 'buried' in a single overall score where it is possible to avoid doing so. This point is acknowledged (p 39) with respect to the overall performance rating, but it holds for the components as well, particularly if individuals/GPs are intended to make use of the Healthcare Commission's work when deciding which trust to use. (This appears to be recognised in the box on page 40, but not the example on page 38).

A particular aspect of this issue is the question of how the process/system measures are intended to relate to the 'hard' data measures: what if they tell different 'stories'? In our view the hard data measures rather than the process measures should be given priority. In some cases, however, the views of patients may clash with the hard data i.e. a hospital which appears to be safe may not be perceived as being so. Such conflicts must be revealed and, if possible, explained, not hidden in a summary figure.

In our view the Healthcare Commission should modify the scale proposed for the annual performance rating (and for judgements on which is based) so as to allow for the fact that good and bad performance may co-exist. It would be better to use a shorter scale with each level described more discursively so as to reflect the mix of performance and to allow for good 'marks' to be achieved by organisations with some areas of poor performance. The summation process should explicitly be designed to ensure that trusts do not move up or down on whatever scale is used because of a change in a single domain. For these reasons it may be better to use a shorter scale: serious underperformance in a number of areas: mixed i.e. some slight underperformance but most OK and OK all round.

Consistency

Performance in some elements, for example, infection or mortality risks may be measurable provided that all trusts can employ the same system of measurement in the same way. Consistency of judgement between trusts should then be attainable. However, experience with the measurement of waiting times indicates that consistent measurement systems are hard to achieve particularly when there is a strong incentive to game the system.

Where the elements are not measurable there is the further problem of how to define whether or not the situation examined is 'acceptable' or not, without recourse to a scale or metric. If no scale exists then subjective judgements are required. By their nature these run the risk of appearing arbitrary and will not carry weight, particularly among those judged to be performing poorly. Unless consistency is perceived as being achieved, the process will not be seen to be fair and its impact will be reduced.

In our view, consistency of judgement should be treated as a principle in its own right (or as an expansion of principle 10) and explicit consideration given as to how to achieve it. This will require genuine standardisation of data reporting systems and also, possibly, the use of measures applied in other fields where marking schemes are designed to reduce the scope for differences to emerge solely on the basis of subjective judgements as well as the risk of capture of the assessors by those being assessed.

Stability

Experience with the existing star rating system has demonstrated the potential for judgements on overall performance levels to be influenced by quite small changes in the information used in the assessment process.

Volatility may also arise because the technical base of assessment changes or new sources of information become available. Volatility from either source will reduce the value of the judgements concerned and tend to undermine their impact. If it persists, the whole system of assessment will be brought into disrepute.

The Healthcare Commission should ensure that whatever methods it employs to aggregate data within individual domains, or over all domains, that the results are not affected by small changes in the underlying data. The alternatives set out in the Consultation Paper should be assessed using real world data on the extent of year on year variation and where such variation is deemed to be due to 'chance' or factors outside the control of management, means must be found for 'smoothing' or damping

its impact. Similarly it should model in advance the implications of alternative assessment methods to try to identify where the outcome of its assessment may be sensitive to changes in data or method.

Trade-offs

The concept of a standard implies a specific level to be attained, so the key decision is whether to pass or fail. But as the consultation makes clear, there are degrees of attainment. That poses the question of how acceptable 'degrees of attainment' should be defined. Broadly speaking we would expect that higher levels of attainment would cost increasing amounts to achieve i.e. in the case of patient safety some risks can be eliminated fairly cheaply, but as the degree of risk falls, the cost of reducing risk further rises. Given limits on resources and given multiple objectives/domains, any NHS organisation must try to find the right level of attainment in each domain and the right balance between performance in different domains. The Healthcare Commission needs to consider whether it is for it or the organisations being inspected to determine what the right levels and balance are.

A further point follows from this. Because the core standards are not fixed points, but continua, there are likely to be acceptable trade-offs between performance relating to the core standards and those relating to the developmental standards. In other words, the relationship between the two categories is not hierarchical, as the language of the consultation implies.

In our view the Healthcare Commission must acknowledge the existence of trade-offs between different domains and between basic and developmental standards. It must ensure that its processes allow for these trade-offs in forming a view on organisational performance and that it does not define precise standards without ascertaining first the cost of achieving them. It should aim, in the longer term, to identify the benefits at the margin from devoting extra resources to a particular domain so to aid it and the organisations concerned define what a 'satisfactory' level of performance is in each domain.

Serving different audiences

The consultation makes it clear that the assessment process is intended to serve the needs of a number of audiences, principally patients (and their GP or other advisers), managers and clinicians in their day to day work as well as the public at large. The Consultation does not, however, discuss how its processes should be adapted to these various audiences. Their requirements are in fact diverse. In our view it is quite unrealistic for the Healthcare Commission to serve all of them equally well from existing sources of information and within the same framework.

While the Healthcare Commission is to be commended for adopting a 'synthesising' role, by aiming to use the findings of other regulators and existing data sources, it must take a view on where the processes and data sources are not adequate and where they need to be supplemented, and by whom. This implies a need for a systematic assessment of all existing data sources with a view to establishing their suitability for the tasks the Healthcare Commission has set itself and where there are gaps that must be filled.

Technical development

The Consultation Paper notes the need for continuing technical development particularly in respect of developmental standards. However, as the comments above have brought out, the assessment process requires technical development in many areas, including the relationship between cost and benefit across all domains, the interconnections between domains i.e. the extent to which pursuit of improvement in one domain reduces/increases the costs of achieving gains in another.

The Healthcare Commission needs to extend its work into efficiency measurement and case mix issues so that it can be sure that 'like is being compared with like'. It also needs to explore, with real world data, the stability/sensitivity of its proposals for deriving an overall performance rating.

The Consultation Paper states (page 4) that the intention is to make use of existing information as much as possible. To be able to do this well, it must aim to use the data and other information it collects to improve its understanding of the system it is inspecting. In the case of safety, for example, it needs to be in a position to determine what levels of safety should be achievable by what set of policies/resources i.e. the basic input/output relationship in ways which go beyond the approach, typical in the past in the work of the Audit Commission, of listing variations between trusts on a variable by variable basis even though it is well known that the actual relationships are much more complex than this.

The Consultation refers (page 13) to 'intelligent information' without defining the phrase. We take it to mean that the information presented is 'fit for purpose' and presented and interpreted in a responsible way. In the case of the examples given, quality of care and value for money, this would mean, at minimum, that the measures used and the data underlying them had been validated. To do this, however, for the whole of the health and health care sector is currently not feasible as the data required are not generally available. Furthermore, a range of methodologies are in use or in the development phase, so difficult choices lie ahead as to which methods to use or whether new ones are required.

We believe that as one output of the consultation process, a programme of analysis work should be defined to support the proposed working and /survey methods and to meet the exacting goals it has set for itself. The Healthcare Commission should also support trusts in analysing their own data. Our own fieldwork has demonstrated how limited such capacity is within NHS organisations.

King's Fund
February 2005