

8 Health

Introduction

8.1 This chapter explains the current UK Health output measure, and recommendations for the future, in five sections.

- Introduction – including scope and objectives of health care
- Current methods of output measurement, and a critique against the criteria in Recommendation 6.1 (paragraph 6.5)
- Future methods of output measurement
- Inputs and deflators
- Triangulation and productivity measurement

8.2 Health is the largest government service, measured by spending: 31 per cent of government final consumption in 2003. In the United Kingdom, central government and the Devolved Administrations provide, through the National Health Service (NHS), a comprehensive health care service funded from general taxation. The provision of health care services in the United Kingdom is a devolved responsibility and there are important differences in the organisation of services in England, Wales, Scotland and Northern Ireland.

8.3 The Department of Health (DH) in England provides funds to Primary Care Trusts (PCTs), which in turn fund NHS Trusts for providing hospital and some community health services. PCTs arrange for primary care for their local populations, and may provide some community health services directly. PCTs have an important role in planning and providing public health services, including health promotion. Hospitals are operated by NHS Trusts, which were classified as public corporations until June 2004 but, following reconsideration of their status, have now been reclassified as part of the general government sector. PCTs may also purchase services from independent sector hospitals; this is part of NHS output as included in GDP (Expenditure). Services are free of charge at the point of delivery except where patients are liable to pay, for example, prescription and dental charges.

8.4 The Welsh Assembly Government provides funds to 22 Local Health Boards, which assess local needs and pay primary and secondary health care providers.

- 8.5** The Scottish Executive Health Department funds 15 Health Boards responsible for planning health services for people in their area. The NHS in Scotland was restructured on 1 April 2004, with NHS Trusts in Scotland abolished to create single, unified local health systems. Hospital and community health services are delivered by Boards through their operating divisions, and new community health partnerships (CHPs) have been established to work closely with local authorities in planning services and promoting health improvement.
- 8.6** In Northern Ireland, health and personal social services are provided as an integrated service and an institutionally separate NHS does not exist. For convenience, this report uses the term NHS in the Northern Ireland context as if there were an institutional separation. The Department of Health, Social Services and Public Safety (DHSSPS) the four health and social service boards plan, commission and purchase services. Nineteen Health and Social Service Trusts provide health and social services.
- 8.7** The aims of DH, under the 2004 Public Service Agreement, are to improve the health and well-being of the population, to improve patients' experience of care, to reduce inequalities in both, and to continue to deliver value for the taxpayer. Specific targets include:
- Health of the population: improving health outcomes by tackling key risk factors such as smoking, child obesity and teenage pregnancy and reductions in mortality from key diseases and health inequalities;
 - Chronic care management: improving health outcomes by providing a personalised chronic care management plan for those most at risk;
 - Access to services: introducing a maximum waiting time of 18 weeks from GP referral to hospital treatment;
 - Increasing participation in drug treatment programmes: increasing the proportion of users of illegal drugs successfully sustaining or completing treatment programmes;
 - Improving the patient, user or carer's experience: securing sustained improvement in patient experience of the NHS.
- 8.8** The Welsh Assembly Government's aim is to promote the health and well-being of everyone living in Wales and provide effective and efficient health services. The aim is to shift the balance from the acute sector towards preventing ill health in the first place and then to addressing problems at an early stage. The Welsh Assembly Government has made commitments to phase out prescription charges and scrap home care charges for disabled people. The NHS and local government work closely together to deliver integrated health and social care services and to promote well-being and an effective anti-poverty agenda.

8.9 The Scottish Executive Health Department's key aims are to improve health and quality of life, and deliver integrated health and community care services, making sure there is support and protection for those members of society who are in greatest need. The Department has set out 12 National Priorities, including targets for:

- Health improvement: to improve the health of everyone in Scotland and to reduce the gap between the health status of people living in affluent and more deprived communities by tackling teenage pregnancy, dental health, alcohol and drugs misuse;
- 48-hour access: to ensure that anyone contacting their GP surgery has guaranteed access to a GP, nurse or other health care professional within 48 hours;
- Patient focus/public involvement: to actively involve the people of Scotland, including communities, patients and carers, in planning and delivering NHS services; and
- Three clinical priorities, cancer, coronary heart disease/stroke and mental illness: particular priority to tackle Scotland's two biggest killer diseases and improve services for people with mental health problems.

8.10 The aim of the DHSSPS, under the 2004 Public Service Agreement, is to improve health and well-being through a reduction in preventable disease and ill health and by providing effective, high quality, equitable and efficient health, social and public safety services to the people of Northern Ireland. There are specific targets on:

- Improving outcomes from life threatening diseases (circulatory, cancer and renal disease) and incidents;
- Reducing maximum waiting time for all patients requiring inpatient or day case treatment (other than in exceptional circumstances) to six months by March 2010;
- Improving health outcomes for people with long-term conditions by offering a personalised care plan for people most at risk;
- All patients to be able to see an appropriate primary care professional within two working days; and
- Improving the quality of life and independence of people in need so that 40 per cent of all people who receive care managed community services, and at least 88 per cent of all people aged 75 or over, are supported as necessary in their own homes.

8.11 These objectives indicate the aspects of quality which are intrinsically part of health output, but are not included in current measures. This is discussed further in paragraphs 8.46-8.66.

8.12 Specific international guidance on measurement of health care is found in *Eurostat Handbook on Price and Volume Measures in National Accounts* (section 4.13):

‘For hospital services, output (= treatments) can be measured on the basis of so-called Diagnosis Related Groups (DRGs) type classifications. DRG systems are used to classify hospital stays into groups that are medically meaningful and as homogeneous as possible with regard to resource use ... In recent years DRG systems have been introduced in many countries to assist hospital management and funding decisions. DRG systems vary across countries, but they are sufficiently similar. They are always very detailed, consisting of several hundreds of diagnosis related groups.’

8.13 The Handbook suggests a number of possible methods together with their A/B/C ranking.

- For services to inpatients by general and specialised hospitals: ‘the use of fully quality-adjusted DRGs is an A method. While DRGs capture changes in the treatment mix well, changes in the quality of individual treatments are difficult to measure. They may be due to better performing equipment, better performing doctors and nurses or changes in the “hospital environment” such as the occurrence of infectious diseases in the hospital, medical errors, additional facilities for patients etc. Further research on appropriate indicators is needed.

DRGs that cover only changes in the treatment mix will fulfil the requirements for a B method ... Use of crude output indicators like the simple number of discharges is classified as a C method.’

- For medical practice services (general practitioners): ‘the services of GPs are such that each visit can be considered as constituting one treatment. Consequently, the recommended A method is the number of consultations by type of treatment, adjusted for changes in quality. It might be difficult, however, to obtain the corresponding cost weights. In the case of proxy weights or only partial quality adjustment the number of consultations by treatment is a B method. The simple number of consultations can also be accepted as a B method if the different types of treatment are sufficiently homogeneous with regard to the resource requirements (similar cost weights).’

Current and Previous Methods of Output Measurement, and Critique

8.14 The UK Health output measure used before June 2004 reflected movements in 16 different activity series measuring health care in England. A single series counting total inpatient and day cases accounted for about half the expenditure covered by the index; outpatient and community health treatments, GP prescribing and dental treatments were measured separately. All the series were taken from NHS operational data systems, except the measure of general practitioner consultations, which was taken from the General Household Survey (GHS) (see paragraphs 8.20-8.22). An aggregate index was formed by weighting the separate series together to reflect the amount spent on each. This index was therefore heavily influenced by movements in the count of inpatients and day cases. Only annual figures were available, after a substantial delay: the *Blue Book 2003* used figures from 2001/02.

- 8.15** As a result of work carried out by the review team and ONS, in close cooperation with Department of Health officials, an improved version of this methodology was introduced into the National Accounts published on 30 June 2004, and is described in articles on the ONS website dated 30 June 2004 (ONS 2004a) and 18 October 2004 (ONS 2004b). The method uses information about volume and cost weights for the 1,200 Healthcare Resource Groups (similar to Diagnosis Related Groups used internationally) and 400 other activity groupings from DH's National Schedule of Reference Costs, of 200 categories of general practice prescribing and includes other activity such as NHS Direct. They range from a GP prescribed drug valued at less than £10 to a bone marrow transplant costing £45,000. The main data source, the National Schedule of Reference Costs, is updated annually and is audited, enhancing the reliability of the estimates.
- 8.16** As noted in paragraph 6.6, the improvement in the June 2004 measure over the previous one comes from wider coverage, an increased level of detail, better cost weights and improved timeliness. This became possible because the NHS has developed robust costs for a standard list of treatments, known as reference costs, which are being introduced progressively as a basis for 'payment by results'. These give detailed cost weights with matching counts of activities, within what were previously broad categories like 'inpatient and day case treatment'. This avoids the weakness that any inpatient treatment was regarded as giving equal output, despite very different costs and clinical benefits. The categories used in the current measure are each more homogenous than previously. As a result, changes in the mix of activities carried out will be reflected in total output proportionately to relative cost, as with most other indices in the National Accounts. Since a higher priced treatment adds more to the overall volume than a lower priced one, the more differentiated index captures some of the increase in the quality received by patients, as set out in paragraph 6.28.
- 8.17** In principle, the more differentiated method might have led to either an increase or decrease in the growth rate of measured output. The impact depends on the changes in the mix of treatments that has actually taken place. Comparison between the May 2004 and June 2004 lines in Table 8.1 shows the impact on NHS output growth from the change to the new method. The third line, for October 2004, shows information published by ONS later (ONS 2004b), using further improvements to the methodology, i.e. an improved link between the old and new series, and replacing gross expenditure weights with net (the weights should refer only to government expenditure and exclude private payments such as prescription charges and dental fees). Comparison shows that changes in methodology can have a substantial effect on the annual growth rates, affecting both the increase in output over the period as a whole and the attribution to individual years.

Table 8.1 General government final consumption expenditure on Health, chained volume measure: annual growth, UK, 1996-2003, per cent

Dataset	1996	1997	1998	1999	2000	2001	2002	2003	Cumulative change 1996-2003
May 2004	2.6	2.3	2.6	2.1	0.9	1.9	2.6	2.6	16.0
June 2004	3.9	1.3	1.8	3.1	3.0	4.2	4.1	4.1	23.7
October 2004	2.9	1.5	1.8	3.2	3.0	4.2	4.1	4.1	23.9

Source: Office for National Statistics

- 8.18** The June 2004 method covers a wider range of NHS activities, but not all. It has added measures for the new NHS Direct and NHS walk-in centres (less than 1 per cent of spending). DH estimates that the activity volume measures used in the new output measure cover about three-quarters of all expenditure on NHS health care activity in England, measured by expenditure in 2002/03. Work underway, as a result of the extension of reference costs, will increase coverage to 81 per cent in 2003/04 and even further in subsequent years. This is consistent with Principle D (see paragraph 4.43) on coverage of output measures. Areas not yet covered include a wide range of activities, each with relatively low expenditure. They include continuing and intermediate care, services for people with learning and physical disabilities, national screening programmes and certain other public health interventions.
- 8.19** Timeliness is much improved, as is the availability of information on within year changes. Estimates of GDP and its components are published by ONS on a quarterly basis. A first estimate is published about a month after the end of the quarter being measured. Two further estimates then follow at monthly intervals. These three separate estimates appear because there is a demand for regular updates and new information becomes available that users want to see included in GDP estimates as soon as possible. The introduction of the latest methodological change has advanced the date at which health care activity estimates are available. The ONS First Release *UK Output, Income and Expenditure, 3rd quarter 2004*, published on 26 November 2004, contained an estimate of Health output based on data on just over half of the activities actually carried out in that quarter, measured by value. This represents a substantial increase in timeliness as compared with the previous method. This achievement goes some way towards removing the criticism that early estimates of GDP are based almost wholly on data from business while data from government activities take much longer to become available.
- 8.20** The Health output measure data source for general practice consultations is the GHS, a sample survey of 20,000 people living in Great Britain. One question asks respondents if they have made contact with their GP practice in the previous two weeks. This includes contacts such as a visit or a telephone consultation with a GP or practice nurse; respondents are asked on how many occasions each of these occurred.

- 8.21** As paragraph 2.20 explained, output indicators are used to estimate growth in output volume, year on year. It is therefore important to consider the likely error in the estimated growth rate of GP consultations, given the use of a sample survey and the particular calculations used. Calculations based on the GHS indicate that, between 2001/02 and 2002/03, the number of consultations with a general practitioner went up from 217m to 241m, an apparent increase of 11 per cent. However, the characteristics of the sampling scheme and the nature of the target population are such that the 95 per cent confidence interval around this growth rate is estimated at +/- 7 percentage points. Knowing that the annual growth rate of consultations with GPs is between 4 per cent and 18 per cent is not precise enough, given the national accounts focus on growth rather than levels. The GHS is widely acknowledged as a useful source for data such as consultation rates for different age groups, and as a snapshot for any one year. However, it is not sufficiently robust as a way of measuring growth in consultations with GP practices from one year to the next.
- 8.22** The output measure for GPs also depends on the weight given to each type of consultation. The weight used is the average cost of each type of consultation. It is calculated for DH by the Personal Social Services Research Unit at the University of Kent (see <http://www.pssru.ac.uk/uc/uc2004.htm>) as a function of the average length of each type of consultation – including any travel time – and the cost of the GP's time and immediate overheads. Clearly, a home visit is likely to cost more than the average consultation in the surgery and this is reflected in the weights currently used. PSSRU calculations make use of up to date information on GP earnings and other cost elements, but the information on the length of GP consultations of various types dates back to 1992, using a survey commissioned by the Doctors and Dentists Review Body (DDRB) to inform pay determination. There have been many developments in general practice since 1992 and consultation lengths may have changed. There may also be an issue on the adequacy of the DDRB survey sample size for calculating length of the rarer events like telephone consultation. There are thus weaknesses in both the volume and cost weights used for GP consultations in the current Health output measure.
- 8.23** The output estimates currently used in the National Accounts take data from England and gross up by expenditure weights to the United Kingdom. Health care activity in Scotland, Wales and Northern Ireland might exhibit different trends from those in England, and – in line with Principle E (see paragraph 4.48) – steps should be taken to include relevant data from those countries.
- 8.24** The current methods capture *activities* carried out. Under Principle B (see paragraph 4.24), an output measure should be adjusted for the attributable incremental contribution of the activity to individual or collective welfare. This should include capturing any change in outcomes which is attributable to the use of the inputs. A basic count of activities does not measure the quality of the output such as change in quality of patient experience or clinical effectiveness. This is a continued weakness of the current method and is discussed further in paragraphs 8.46-8.66.

Future Methods of Output Measurement

8.25 In the light of the critique above, improvements are needed in a number of areas, including primary care, UK coverage and quality measurement. This section presents proposals for change under five headings:

- Primary care.
- Extending the coverage of output volume indicators.
- Improving UK coverage.
- Whole courses of treatment, technical change and substitution.
- Measuring quality change.

8.26 Improvements are already planned in respect of the first three areas, which are relatively straightforward. The last two areas, measurement of output in terms of whole courses of treatment and measurement of quality change (including attribution of outcomes to NHS activities), are more difficult and will not be resolved in the short term. In line with Recommendation 6.7 (see paragraph 6.33), a relatively high threshold should be set for inclusion of quality adjustments in the National Accounts. A preliminary period where possible measures are published for discussion by relevant experts, including user views, would be helpful. The development work on quality measurement proposed below may lead to such publications.

Primary care

8.27 Primary care is the first point of contact with the NHS for most patients. It includes general dental services, pharmacies and ophthalmic services, but this section is limited to general medical services provided by general practitioners, practice nurses and support staff. Services include primary diagnosis, advice and reassurance, prescribing medicines, health promotion, screening for disease and referral to specialists. Some minor surgical procedures are carried out in primary care.

8.28 The output of primary care would ideally be measured in a way that took account of the range of different activities and resulting benefits for patients, and any change in the mix of services and their quality. Many patient health episodes are treated entirely within primary care. Where other parts of the NHS are also involved, it would be best if the general practice contribution could be measured as part of a package (a care pathway) that also involves diagnosis and treatment provided by hospitals or community health staff.

- 8.29** Keeping those objectives in view, the immediate priority is to consider whether a better data source could be found for estimating the number of GP consultations, and their cost weights. The review team, with DH officials, have been discussing the use of more accurate and more timely data on GP consultations – annually and quarterly – from anonymised research databases created from GPs’ computer records. Arrangements are being made by DH for this work to be done by QResearch, based at the University of Nottingham. It is likely that first results will be available, following quality assurance work, in March 2005. Thereafter, results will normally be available within four weeks of the end of each quarter. The database contains records of approximately 6 per cent of all patients registered with general practice. These come from nearly 500 practices widely spread through the United Kingdom. The sample data can be re-weighted to make it more representative of the UK population. This source will provide data on a much larger number of consultations than currently obtained from the GHS. The likely error ranges of the growth rate have not yet been determined but they are expected to be lower than those from the GHS.
- 8.30** The Scottish Executive are considering use of data on GP consultations from the Practice Team Information (PTI) Database, which has been expanded in 2003/04 to collect information on GPs, practice nurses, district nurses and health visitor activity. The information is collected from 45 PTI practices, covering around 6 per cent of the Scottish practice population. The PTI Database allows for analysis of the number of consultations with GPs (and other staff), including information on patients’ condition, treatment, age, gender and deprivation characteristics. It is available quarterly. In Wales, data on GP consultations could be taken from the GP Morbidity database, which contains information from 43 practices covering over 10 per cent of the population. Other similar databases exist or are in development, covering different groups of general practices (including one for Northern Ireland), and it might also be useful to determine whether they should be substitutes or complements, and take account of any apparent differences in trends from different sources.
- 8.31** Further work is also needed to update cost weights used to combine different types of general practice consultation. If it continues to be necessary to estimate the number of minutes spent on each type of activity, a means should be found of verifying or updating the 1992 survey estimate.
- 8.32** **Recommendation 8.1:** we recommend that ONS should continue working with the four health departments to make use of information from computerised GP research databases to improve measurement of GP output, and should update cost weights.

Extending the coverage of output volume indicators for each function

- 8.33** As the coverage of the National Schedule of Reference Costs expands, we expect ONS to continue to work with DH to extend the coverage of the current activity measure, to 81 per cent (see paragraph 8.18) and beyond. Failure to measure every type of activity does not automatically lead to an understatement of the volume of output. The ONS convention is to assume that the excluded activities, as a group, are growing at the same rate as those that are measured. This may either understate or overstate actual change. The potential error from this assumption reduces as the proportion of all activity covered by specific measures increases. While further extension of the output measure to capture change in all distinct activities is desirable in principle, there may be diminishing returns from effort to extend coverage once it reaches about 90 per cent, and other areas of change may be more important. The timeliness of Health output data is already much improved (see paragraph 8.19), but it may be possible to increase further the proportion of actual data used in the first estimate of each quarter's GDP.
- 8.34** ONS and DH are continuing to examine areas of expenditure for which the corresponding activity is not yet included in the overall volume of output measure. Even if statistics are not available at national level, health service managers and clinicians may be able to advise on changes in case mix which might lead to a particular area of excluded activity becoming more (or less) important and so affecting the overall measure. Regular review is needed to ensure that new areas of activity, or new ways of providing existing activity, are identified and considered for inclusion in the overall measure.
- 8.35** The output estimates currently used in the National Accounts presume that all expenditure on Health is on individual services. However, some expenditure provides a collective service, eg public health campaigns and, under Recommendation 6.4 (see paragraph 6.16), should be measured by a volume index of activity or by the volume of inputs, taking account of quality change of inputs. Where the expenditure provides neither individual nor collective services, the costs should be seen as overheads and only included as inputs; they play a role in producing output, but the output will be measured against the expenditure incurred in delivering either the individual or collective services. ONS and DH are discussing the borderlines between individual and collective services, and between service provision and overheads and this should lead to further improvements in the methodology.
- 8.36** **Recommendation 8.2:** we recommend that ONS and the health departments should work together to incorporate the widening scope of Reference Costs and equivalent sources into the Health output measure as this becomes possible, with further improvements in timeliness; should keep under review the NHS services for which there are no direct output measures, taking expert advice on the potential impact on overall NHS output and productivity estimates; and should distinguish appropriately between individual services, collective services and overheads.

Improving UK coverage

- 8.37** ONS, with DH assistance, is working with the health departments in Wales, Scotland and Northern Ireland to see how far relevant data from their health systems can be incorporated into the National Accounts, using the techniques adopted for England in 2004 or adapting them as appropriate. This would meet Principle E (see paragraph 4.49) on UK coverage. Good progress is being made.
- 8.38 Recommendation 8.3:** we recommend that ONS and the health departments in Scotland, Wales and Northern Ireland should introduce measures of Health output in those countries into the National Accounts once sources and methods have been verified.

Whole courses of treatment, technical change and substitution

- 8.39** Ideally, we should look at the whole course of treatment for an illness rather than at its components. This might include several linked outpatient attendances, investigations, inpatient stays where the patient may be transferred between consultants, and follow-up care including GP consultations and prescriptions. At present, each of these counts as an independent unit of activity in different parts of the current Health output index. A change in medical practice could change the total count of activities without a corresponding change in outcome. For example, new ‘one stop shop’ hospital procedures for the timing of diagnostic tests and medical consultations can lead to a judgement on the diagnosis and treatment plan during a single outpatient appointment rather than a number of separate appointments. The health outcome for patients may be the same or better, but the count of activities is lower, and the cost weighted index may also be lower. If it were achievable, a unit of output based on a course of treatment would be less prone to artificial distortion as a result of changes in procedure.
- 8.40** Such an output measure could take better account of the quality of care given. Readmission soon after treatment may be a sign that the treatment has been unsuccessful and this should ideally be flagged in the way the data are collected (and perhaps treatment for problems created by clinical failure should not be counted as a new output). NHS data systems have not in the past attempted to record complete courses of treatment, though more may be possible when current major IT developments are complete. NHS policies for chronic disease management also suggest a need for supporting information which links the activities of hospitals, community staff and general practitioners together, to measure the costs and effectiveness of an overall package of care.
- 8.41** Early progress in measures based on aggregating NHS episodes into whole courses of treatment is unlikely, but we would encourage further exploration of the concepts, to consider what might be done to give the clearest picture of value added by the NHS. We also consider it very important that developments in NHS information systems, while intended primarily to support patient care, should also support better analysis of health care outputs.

8.42 Nicola Mai, an ONS economist working on the review, explored a method for measuring health output using a diagnosis-based approach, presented in a paper published in *Economic Trends* (Mai, 2004). This work is based on similar work carried out in the USA by David Cutler et al (Cutler, 2001). The method has three key characteristics:

- the unit of measurement is the whole course of treatment of a patient rather than activities performed;
- units of measurement are grouped by diagnosis; and
- treatments are adjusted for quality factors.

8.43 Mai illustrates this approach with a case study on a single diagnosis – coronary heart disease. He uses a common cost weight for two substitutable treatments (coronary artery bypass graft (CABG) and percutaneous transluminal coronary angioplasty (PTCA)). The clinical pattern has been a fall in the more expensive treatment, CABG, and a rise in the cheaper one, PTCA. His proposed methodology suggests higher output growth compared with the results of the 2004 ONS method. If clinical opinion and research evidence is that patients have done as well or better with PTCAs as with CABG, then it is misleading to say that health output is lower simply because the cheaper treatment has been used. Where treatments are substitutes, and especially where the cheaper one gives better results for the patient, Mai's technique has a distinct advantage. However, unless clinical information about the outcome of treatment can be used consistently, the current cost weights are a better way of reflecting the overall pattern of change in output.

8.44 Mai suggests further work (as discussed in paragraphs 8.39-8.41) to combine outpatient, hospital, drug, GP consultation and GP prescribing data with hospital inpatient data to give a clearer analysis of the cost of care pathways. His methodology could be used to measure quality gain where one form of substitutable treatment has a clear advantage over another. An approach which identifies benefits from technical innovation and substitution, especially if it could be based on the costs and benefits of whole courses of treatment, might be used to analyse changes in health outputs in certain areas. If a wider range of examples become available, results could be explored as part of the triangulation approach recommended in Principle H (see paragraph 4.64).

8.45 Recommendation 8.4: we recommend that ONS should explore, with DH and the wider health information and research community, ways of taking forward work on whole courses of treatment, technical change and substitution, and should make use of the results in Health productivity articles.

Measuring quality change

8.46 The summary in paragraphs 8.7-8.11 of the priorities for health set by governments across the United Kingdom suggested that the main dimensions for understanding quality of health care are:

- a) saving lives and extending life span;
- b) preventing illness and mitigating its impact on the quality of life;
- c) speed of access to treatment; and
- d) quality of patient experience.

8.47 All these dimensions of quality are relevant for the National Accounts, and work is needed to find ways of measuring them, in addition to the disaggregated count of activities which (as explained in paragraph 6.28) already incorporates some aspects of quality change. The National Accounts would need to measure the year on year change in quality of health care received by individual patients, attributable to the NHS. Change may be easier to measure than the absolute level of quality. Quality has several dimensions, and so vectors for change in each of these, if measurable, would need to be combined. That requires some common basis for valuing (say) a 10 per cent improvement in one aspect of quality against (say) a 5 per cent decrease in a different aspect. This ideally requires the marginal social valuation of different aspects of quality so that changes in different dimensions are commensurate, eg by expressing them in money terms. Further work would be needed to find a basis for these relative values which would be accepted as robust.

8.48 Quality measurement is important in health care and medical research more widely, and expertise from public health medicine, epidemiology, health service management, health informatics and health economics could all be valuable. There are various sources of quality indicators, such as Healthcare Commission performance indicators and National Centre for Health Outcomes Data (NCHOD) clinical outcome indicators, but these are often more suited for use in cross sectional analysis rather than as time series. Consistency of definitions is important for measurement of quality change over time, and Healthcare Commission (or their predecessor) performance indicator definitions change fairly frequently, since their prime purpose is to measure and incentivise improvements on matters of current concern. Nevertheless, these bodies and others in the health area have expertise in understanding and defining quality measures and it would be helpful to engage them in development work

8.49 DH is funding the Centre for Health Economics at University of York and the National Institute of Economic and Social Research to develop new approaches to measuring and understanding NHS outputs and productivity. Their work began in March 2004. A first interim report (Dawson et al. 2004), focusing on data sources in secondary care and setting out ideas for possible methods, was published in July 2004. A second interim report is planned for publication early in 2005 and a final report is due in August 2005. This initiative is very welcome and the final results will be helpful in taking forward practical steps to measure Health output, within the framework proposed here.

Saving lives and extending life spans; mitigating effects of disease

- 8.50** The first York/NIESR report discusses the concept of Quality Adjusted Life Years (QALYs) as part of a potential currency for health outcome improvement. Gaining a better understanding of how the NHS is improving the health status of patients is a goal in itself, apart from potential use in the National Accounts. Attribution is an important issue. QALYs are a useful measure where the activity of the NHS is the exclusive factor in the change in health status. This would be the case in much curative care, for example treatment of a broken arm, where all (or almost all) the health outcome is attributable to the NHS. In other situations, including long-term care, the health status of NHS patients may be explained in part by other factors. Joint production between health and social care for elderly people with disabilities and other complex medical needs is an issue (see paragraphs 6.22-6.25). Other social factors are also relevant. For example, it is not clear how far improved survival rates from cancer may be due to earlier diagnosis and treatment, more effective treatment, healthier life styles or beneficial effects of affluence. The first two factors involve health care expenditure, and to a degree health promotion activities (eg smoking cessation clinics) influence life styles. But health care activity is clearly not the only factor. The York/NIESR research team are considering further the issues on attributing changes in health outcome to the NHS, as required in Principle B (see paragraph 4.25).
- 8.51** BUPA reported on its collection of information on health outcomes (BUPA, 2004). BUPA collected such information from its patients for the last six years, covering more than 100,000 cases of elective surgery. They have used two instruments, SF36 (a generic health status measurement instrument; see <http://www.sf-36.org/>) and VF-14 (an instrument for assessing visual impairment, AO 1994). The main purpose is cross-sectional analyses to support identification of poor performance and good practice. The first three years of collection saw relatively low response rates, but they have been improved by for example issuing postal reminders to patients. This is an interesting precedent but evidence of outcomes for patients receiving private health care would not in itself be relevant to measurement of government final consumption. Further work is needed to explore whether similar information for NHS patients, if available, would be useful in measuring quality change over time, eg whether any changes in the mean of the distribution might be due to factors such as changing response rate or case mix (eg more patients who have co-morbidity which affects their responses to the survey) rather than a genuine difference in attributable benefits to individuals from health care.
- 8.52** DH is commissioning research into the collection of information on health outcomes for patients visiting independent treatment centres, focusing particularly on a limited number of treatments (cataract surgery, hip replacement, knee replacement, varicose vein procedures and hernia repair). In a first stage, which began in December 2004 and is due to report in June 2005, the London School of Hygiene and Tropical Medicine is reviewing the international evidence on routine use of patient reported health status measures from which can be derived assessments of the outcomes of treatment, including both measures specific to the procedures undertaken, and generic measures. In a second stage, to begin in April 2005, piloting of routine use of the data will begin.

8.53 Collection of information on health outcomes needs to reflect the typical natural history and care pathway for different diagnoses. The choice of timing of the collection of information on health status depends on the length of time it takes for health gains from NHS activity to be achieved. This might be days, months or (for cancer survival rates) years. It is also important to understand the counterfactual: a patient may be worse after treatment than before, but nevertheless be better than they would have been if no treatment had been given at all.

8.54 Further consideration of data to assess changes in health outcome as a result of NHS treatment should consider carefully how health status information can be used to track changes over time. Particular attention will need to be paid to:

- the choice between sample surveys (with consideration of appropriate sample size and the issues in comparing results of successive unlinked surveys – see paragraph 8.21) and a complete census;
- issues involved in using information designed for cross-sectional use to answer questions about change over time, and possibility of linked longitudinal information in some areas; and
- how to separate out the particular contribution of the NHS to the outcome.

8.55 It would be very helpful to be able to base quality adjustments for NHS output on a data set which measures the health outcome achieved as a result of treatment, collected annually by all or part of the NHS for most aspects of health care. This is a major undertaking. An alternative approach – we are grateful to Professor David Cutler of Harvard University for suggesting this to us – would be to focus initially on one or two major disease groups (eg coronary heart disease and cancer) for which there is agreement, based on research, on the clinical practices which are most likely to lead to effective outcomes. The National Service Frameworks published by DH would be a starting point. Evidence is then needed on the extent to which the desired practices are being followed – this could be monitored by the rate of prescribing of recommended drugs, for example. The overall effectiveness of treatment patterns for the disease group could be compared with previous years, taking account of known demographic factors which affect the cases presenting for treatment. Research evidence could be used to estimate the health gains (eg future years of healthy life) which will result from the use of best practice treatment patterns, rather than previous ones.

8.56 This approach might require less new data collection, and rely more on analysis of existing research evidence and expert clinical opinion. Work on these lines would be very interesting in assessing the extent to which health care spending is producing worthwhile results, beyond the count of volume activity which is already available. It may not be possible to develop a precise numeric quality adjustment which would be acceptable for the National Accounts, but publication in the context of productivity analysis could be very helpful.

- 8.57** The new GP contract introduced in April 2004 includes financial payments based on evidence of quality of patient care, in a number of specified areas. This will mean that computerised data sets will become available on the numbers of general practitioners meeting standards which have been set, on the basis of research evidence, as likely to improve patients' health – eg to prevent a first or subsequent heart attack or stroke. Numbers will be available, through financial systems, of the number of points scored by practices against these criteria, and it may be considered that evidence that general practices, taken together, receive more quality points in one year than another is itself evidence of quality change in health care.
- 8.58** The new GP contract is a promising source of information on what is done in general practice, rather than just counting consultations (see paragraphs 8.27-8.30). However, many of the indicators will only distinguish between two levels of performance. For example, indicator CHD4 distinguishes between practices with 90 per cent of patients with coronary heart disease and who smoke, whose notes contain a record that smoking cessation advice has been offered within the last 15 months, and those below this threshold. This does not directly yield information on the numbers of patients who are advised to give up smoking, or on the numbers who actually do so. At least initially, practices may gain points by improving the way they record their current clinical practice rather than by changing practice. There is also a possibility that thresholds will be redefined after a few years' experience, with variation at PCT level possible. It may not be easy to find consistent and stable data for national measures of quality gain, but the rich data sources should be explored further.
- 8.59** Work on measurement of changes in health outcome attributable to health care, if the various approaches discussed above lead to results which cover a reasonable range of health care activities, could be used in two different ways. One approach would be to seek to use weights based on the value of health gain from each treatment rather than on its cost. As discussed in paragraphs 6.17-6.19, it would be better to combine different aspects of health care by using weights based on marginal valuation rather than on average costs. This would need a substantial evidence base, and judgement about how to combine marginal valuation and cost weights within the same index if it were not possible to find marginal valuations for all aspects of health care. There are also reasons (eg joint producers – see paragraph 6.24) why cost weights may, overall, be more satisfactory. While a framework for marginal valuations should continue to be explored, it would be unwise to rely on this as the only way to use information about health outcomes.

- 8.60** On an experimental basis (for use in separate publications, rather than the National Accounts, at this stage), it might be helpful to identify some types of health care activity where there is research evidence that marginal valuation and cost weight may be very different. For example, prescription of statins has benefits in reducing the chance of future heart attacks which can be regarded as higher than the financial cost of the drug. There may be other examples where research has shown that expensive treatments have low clinical benefits, and yet they continue to be used. The National Institute of Clinical Effectiveness (NICE) has responsibility for advice in this area, and ONS and DH might like to consider whether some further work with NICE could identify treatments where marginal valuation and cost weight are very different. If so, analysis of the difference in measured output growth, using estimated marginal valuation rather than cost weights, would be of interest.
- 8.61** The alternative to using information on health outcomes to form a set of weights is to use them to define the volume measures for health care in terms of the degree of success of the treatment (the second method in paragraph 6.26), or to use them to apply a separate quality adjustment (the third method). In both cases, the aim is to measure 'change from previous period' in commensurate terms – the absolute level need not be included in the measure. The same approaches apply to the other two dimensions of quality discussed below, speed of access to treatment and patient experience.

Speed of access to treatment

- 8.62** Another aspect of quality identified in paragraph 8.46 is speed of access to treatment. Ideally, this should be measured as the time from first concern about symptoms through to completion of treatment. So far, the main Hospital Episode Statistics (HES) dataset records only the time from a consultant decision to put a patient on a waiting list for treatment, to treatment. Collection of data and analysis of an extended definition of waiting time would be desirable, as a minimum including time spent waiting for an outpatient appointment which is then followed by elective hospital treatment.
- 8.63** The focus in NHS performance measures (see paragraphs 8.7-8.10) is on maximum waiting times. For the National Accounts, it would be more appropriate to reflect the experiences of all patients and to make use of data on mean waiting time, for those in comparable circumstances. It is possible, though, that the quality gain to patients from quicker access to treatment may be non-linear (we are grateful to Professor Barry McCormick, Chief Economic Adviser in DH, for his suggestions on this point). For a condition which is not painful or disabling, say, there may be a gain in utility from waiting two months rather than three, but less gain from waiting one month rather than two, and perhaps disutility if the patient is given less than one week's notice of hospital admission. It might therefore be sensible to set a minimum floor of tolerable waiting time, which would be different for different diagnoses – there is strong management and clinical concern to treat cancer quickly. The data on actual waiting times would be adjusted by subtracting the floor. Other approaches to take account of non-linearity would also be possible. On this basis it may be possible to make early progress in using information on changes in waiting time for elective treatment as one measure of quality gain for NHS patients.

8.64 Other information on speed of access to diagnosis or treatment could also be studied for potential use as measures of quality change over time. There is management information about the distribution of waiting times for admission to hospital from accident and emergency departments, the proportion of patients seeing a GP within 48 hours, and emergency ambulance response times. In each case data collection is designed to identify the numbers of patients who exceed a maximum time period and may not be suitable to identify the full distribution. Issues of non-linearity also suggest that the unmodified mean may not be the most useful measure. But it would be helpful if ONS and the health departments would consider further whether measures of quality change could be developed in these areas.

Patient experience

8.65 The final dimension of quality identified in paragraph 8.46 is patient experience. This is measured by individual Trusts through a national patient survey programme, using questions on five dimensions of care: access and waiting; quality and co-ordination of care; information and choice; relationships with staff; and environment. Results for England are available at national aggregate level from the Healthcare Commission. All these areas are relevant to the measurement of quality, and we understand are being considered by the York/NIESR project. Care will need to be taken to be sure data analysis is meaningful as a measure of quality change over time.

8.66 **Recommendation 8.5:** we regard the measurement of quality change in health care as a difficult area, but have a number of suggestions for work which should be taken forward. The results of research commissioned by DH from the University of York and National Institute for Economic and Social Research will also be important. We recommend that:

- a) a number of dimensions of quality should be measured, with results weighted together by marginal social valuation: more work would be required to underpin these weights;
- b) a range of expertise should be used to develop quality measures, including public health medicine, epidemiology, health service management, health informatics and health economics;
- c) ONS and the health departments should assess options for collecting new information on health outcomes resulting from NHS treatment, with particular consideration to the needs ONS has for measurement of change over time, rather than cross-sectional data sets which are useful to health departments for other purposes;
- d) ONS and the health departments should consider studies of changing treatment patterns for particular major disease groups to assess whether these could provide useful estimates of improved health outcomes resulting from changes in clinical practice;

- e) ONS and the health departments should explore the data set on quality standards in general practice, resulting from the new GP contract, to see whether this could be the basis for a measure of quality change;
- f) ONS and the health departments should consider whether, with advice from the National Institute for Clinical Effectiveness, it might be possible to identify treatments where marginal valuation and cost weights are very different, and explore the difference in output growth resulting from use of estimated marginal valuation instead of cost weights;
- g) ONS and the health departments should develop a measure of quality change based on speed of access to elective treatment, using the Hospital Episode Statistics data set and taking account of non-linearity, with further developments if new measures of total waiting time are introduced;
- h) ONS and the health departments should explore whether measures of quality change could be developed from information sources for time taken for admission to hospital from accident and emergency departments, time before seeing a general practitioner and ambulance emergency response times; and
- i) ONS and the health departments should explore whether measures of quality change over time could be based on the national patient survey programme which measures aspects of patient experience.

Inputs and Deflators

8.67 As discussed in paragraphs 5.11-5.17, the flow of data for the compilation of estimates of government spending on Health is complex. Working with the review team, ONS has made significant progress over the last year in improving the methods used for calculating current price Health expenditure, including reclassifying some areas of spending (like health care in prisons) between areas of COFOG. Further work to document the data flow and check all relevant classification issues, as set out in chapter 5, would be helpful.

8.68 Improvements in the methods used to deflate current spending, for use in productivity calculations, were reported by ONS in *Public Service Productivity: Health* (ONS 2004c). Table 8 of that article shows the impact that different deflators have on estimated productivity change. Further detail on the improved deflation sources and methods can be found in ONS's technical paper *Sources and methods for Public Service Productivity: Health*.

- 8.69** There is more work to be done to ensure the health deflators meet in full the quality criteria proposed in recommendation 5.9 (see Table 5.1). The ONS article points to uncertainties on the best approach to deflating expenditure on prescription drugs, and more work is needed in this area, and on capital services. There would also be benefit in a more disaggregated approach to measuring changes in skill mix as part of the labour deflator, whether the direct or indirect method is used (see paragraphs 5.62-5.65). Current data on pay and price changes relate primarily to England, and should be developed to take into account any differential in price movement in other parts of the United Kingdom. The deflators for some items of expenditure are not appropriate, for example indicators of price movements in the hospital sector are used to deflate expenditure on administration of the Department of Health. Expenditure on health care in prisons, provided by the armed forces and in nursing homes within Health also requires further review of relevant pay and price indicators for this spending.
- 8.70** **Recommendation 8.6:** we recommend that ONS should work with the four health departments to improve the deflators for current price expenditure on Health, and the matching expenditure weights.

Triangulation and Productivity Measurement

- 8.71** We welcome the ONS article *Public Service Productivity: Health*, published on 18 October 2004. It is helpful to put on public record a range of estimates to illustrate the different ways in which aspects of productivity can be estimated, and to be explicit about known problems of measurement, as a guide to those who interpret the figures.
- 8.72** It is possible to measure the productivity of health care by dividing national accounts inputs at constant prices by outputs at constant prices, as in the October article. This measure cannot be regarded as a total picture of productivity. We therefore also welcome the section on triangulation. Future such articles might also publish new sources and methods which would eventually be considered for incorporation in the National Accounts output measure, as well as providing corroborative evidence on Health productivity.
- 8.73** Fruitful areas of investigation for further triangulation material include the following:
- a) Account should be taken of the changing skill mix of health care staff: a change in the overall level of skills gives a legitimate expectation of a change in the quality of output, for which evidence can be sought; other shifts in skill mix may indicate more, or less, effective use of support staff to bring about change in the productive time of those with more specialist skills. Good examples of this include the changing balance between numbers of different grades of hospital doctor and reductions in junior doctors' hours as a result of the working time directive.
 - b) The migration of treatments from expensive settings (eg an acute hospital) to cheaper ones (eg GP's surgery) should be examined.
 - c) Changes in the technology of treating particular diagnoses should be examined, particularly in areas where such developments move the output from one category to another, eg the substitution of drugs for invasive surgery.

- d) The levels of use of scarce resources should be investigated, such as intensive care beds, ambulances or operating theatres; more intensive use would generally increase productivity, though a very high rate of use may be as inefficient as a very low rate, as time is wasted in hunting for an alternative, and/or patients have adverse outcomes.

8.74 Some Health expenditure is expected to lead to outputs in a different time period. This is the case for prevention and public health services, education and training of health personnel, and research and development in medicine and health care. None of these should be expected to produce immediate value in terms of health care activity. Prevention aims to reduce the requirement for health care, with benefits accruing over many years. Staff training (much relevant expenditure remains within Health, despite the recent reclassification) may increase the productive capacity of the health service with a lag, as new staff take up posts. Or it may increase the quality of health care in different ways if current staff are trained in new ways of delivering care or improving outcomes. Research should improve the effectiveness of health care spending, but the link may be indirect and with a considerable time lag before there is sufficient body of evidence to change health care practice.

8.75 Analyses of productivity should investigate how far current expenditure may lead to future output, and whether current output is reflecting lagged past investment. This will require more work to build a model of past expenditure and its results. It may also be fruitful to explore the relationship between input and output for spending on:

- arms' length bodies such as NICE, the NHS University, the Healthcare Commission and the Health Protection Agency; and
- modernisation programmes, which incur expenditure which is designed to improve service provision over many years (see for example DH's Technical Note describing how efficiency savings set out in Spending Review 2004 will be made <http://www.dh.gov.uk/assetRoot/04/09/29/44/04092944.pdf>).

8.76 Recommendation 8.7: we recommend that ONS should continue to publish health productivity articles and extend the range of sources and issues explored in them.

Satellite accounts

8.77 ONS has been developing health accounts for the United Kingdom. Currently, this work publishes estimates of total UK health expenditure on a fairly regular annual basis. It has also developed an experimental set of UK health accounts, based on the international framework promulgated by the OECD. These present an analysis of the source of financing for health activities, of the producer or provider of those activities, and of the purpose for carrying out the activities (eg to cure an existing illness, to prevent illness in the first place).

- 8.78** A true health satellite account would examine the production of health goods and services and would have a framework-based analysis of inputs used in the production of such goods and services. The framework for a satellite account is permitted to differ from that of the National Accounts in order to add to analytic power. This gives the possibility of developing a satellite account framework for health which encompasses alternative ways of measuring inputs and outputs, as well as concepts which are not wholly part of the National Accounts, for example the material on triangulation discussed in the preceding section. However, whilst there is an international framework for health accounts, this does not extend to a satellite account as described in this section.
- 8.79** The development of a health satellite account would be greatly assisted by the creation of input-output tables for the health function, which would make use of the planned improvements of service sector data as recommended in the Allsopp Review.
- 8.80 Recommendation 8.8:** we recommend that ONS should consider developing the framework for health accounts further, developing full satellite accounts including health production accounts.