

SECTION 4

MEASURING INEQUALITY BY HEALTH AND DISEASE CATEGORIES (USING DATA FROM SURVEYS)

The purpose of this section is to show how health and disease can be measured through surveys and questionnaires.

For practitioners trying to decide whether to measure health inequalities in their region with a re-analysis of existing health data, this section offers guidance on the often bewildering quantity and variety of health, disease and 'lifestyle' surveys available, with some explanations on how these data have been acquired and caveats on their applicability and use.

For those wanting to collect new or different data, the section offers a guide to the range of different health and lifestyle questionnaires, the key types of health and disease measures used, the basic structure of questionnaires and surveys and their main strengths and limitations. ([Section 9](#) provides a guide on how to carry out a survey).

The material is divided into three parts. The first (Sections [4.1](#) and [4.2](#)) concentrates on the types of health and lifestyle surveys that are regularly employed in surveys of the general population.

The second part ([Section 4.3](#)) presents a selection of health status questionnaires that have (mainly) been developed for clinical and audit applications. The distinction between the sections is somewhat arbitrary as the shorter of the health status questionnaires are sometimes incorporated into health and lifestyle surveys, or even used as stand-alone instruments in epidemiological applications.

The third section provides a brief discussion on quality of life indexes.

4.1 Health and Lifestyle Surveys

4.1.1 What can be Covered by a Survey?

Most surveys of health and lifestyle aim to cover the topics listed in [Table 4.1](#). Almost all of these topics were included in recommendations of the European 'Health for All' programme's contents of a questionnaire that could be used for cross-national monitoring of health and morbidity in Europe ([Table 4.2](#)). These topics are sufficiently well established. However, there is less agreement when it comes to choosing or devising measures for each topic; and the lack of standardisation makes for difficulties in comparing results.

The following sub-sections give more details of how health and lifestyle surveys tackle three of their main topic areas:

- ◆ Self-reported health and morbidity ([Section 4.1.2](#)).
- ◆ Physiological measurements ([Section 4.1.3](#)).
- ◆ Lifestyle and risk taking behaviours ([Section 4.1.4](#)).

4.1.2 Questions on Self-reported Health and Morbidity

Although self-reported health is often tackled by basic questions on how well (or ill) people feel, there are some more interesting and detailed efforts to get respondents to report recent symptoms. A common format is

Table 4.1. Typical range of contents of a health and lifestyle survey

Typical coverage of health in a large-scale health and lifestyle survey will include

Health, symptoms and disability:

- Perceived health
- Experience of illness and disease (based on checklist of symptoms and reports of any diagnosis supplied by medical professionals)
- Use of health and social care services
- A short battery such as the GHQ12, or parts of the Nottingham Health Profile to measure psychological malaise and well-being
- Reports of any disabilities, their presumed causes and impact on life
- Fitness

Examinations and physiological tests

- Some of the more ambitious surveys also employ nurses to carryout basic physiological tests, blood pressure, BMI and respiratory function.

Questions on lifestyle will at a minimum cover

- Diet
 - Exercise
 - Risk taking behaviour
 - They may also include frequency and extent of social contacts, any caring roles, and support received.
-

Table 4.2. Instruments for health interview surveys recommended for monitoring European Health For All indicators

<i>Health For All indicator (number)</i>	<i>Health indicator general content</i>
Perceived health (2,2)	Percentage of respondents saying that their general state of health is very good, good, fair, poor or very poor.
Temporary disability (4,1)	Average number of days that respondents had to cut down their usual activities because of physical disability/mental conditions
Long-term disability (4,2)	A summary score derived from questions on confinement to bed, chair or house, difficulty in walking, dressing, washing, feeding, hearing, seeing, etc.
Prevalence of selected chronic conditions (4,6)	Physical health Mental health: batteries of questions on, respectively, dementia, mental retardation, and mental disorders
Emotional well-being	General health questionnaire: a summary score derived from questions on worry, concentration, feeling happy, etc.

Source: reproduced from Kunst and Mackenbach [7]. Previously adapted from third consultation to develop common methods and instruments for health interview surveys: report on a WHO meeting.

to present a list of symptoms (or conditions), proceeded by a question such as “*In the past month have you had any problems with . . .*”. This model can be used in both interview and self-completion surveys (Table 4.3).

This method has certain difficulties:

- ◆ Lists of symptoms described in medical terminology may not be understood by everyone.
- ◆ The types of terms that are acceptable to a lay person may not be reliably translated into standard diagnoses or conditions.
- ◆ It does not capture the severity of symptoms.

The obvious alternative is to use open-ended questions that ask people to describe their symptoms and illnesses, but these are very difficult to code.

There have been attempts to validate the results of self-report check lists of symptoms against medical diagnoses and produce formulae for predicting most likely diagnoses from combinations of symptoms. Statistical techniques such as factor analysis have been used for this purpose, but it is unclear whether the results have any medical or epidemiological credibility.

Table 4.3. *Selections of common symptoms and problems typically used in self-completion surveys*

<i>Column A</i>	<i>Column B</i>
Topics for question of type: Have you had any trouble with x in past month	Possible list of conditions for self-report questionnaire
Headaches	Arthritis/rheumatism
Hay fever	Back trouble
Sleeping	Hernia
Constipation	Orthopaedic condition (excl back trouble)
Eyes	Heart disease, angina etc.
Bad back	High blood pressure
Nerves	Stroke, arterial disease
Colds and flu	Bronchitis, emphysema
Feet	Asthma
Tired	Respiratory TB
Kidney/bladder	Other respiratory disease
Joint pain	Stomach ulcer
Concentration	Other gastrointestinal diseases
Palpitations	Genitourinary diseases
Ears	Diabetes
Worry	Gout
Stomach	
Sinus	
Cough	
Faints	

The situation is very different for mental health questionnaires where this type of diagnosis-based validation is well established and the resulting instruments are employed for both screening and diagnosis.

4.1.3 Physiological Measurements and Surveys

If resources permit, researchers may try to incorporate a small set of physiological measurements in a survey. For example, interview-based surveys employ nurses to conduct physiological measurements on respondents:

- ◆ Levels of refusal can be quite high, and some surveys do not attempt to collect these data from older people (over 75 is a common cut-off point).
- ◆ Exercise can be both costly and difficult to arrange, especially when the measurements require technology that is not readily portable.
- ◆ Blood pressure and respiratory function are two measurements most often collected, but conventional interviews and self-completion methods can record details of height, weight and waist measurements.

- ◆ Height and weight are often used to generate a combined measure of physiological health in the form of the body mass index (BMI).
BMI = weight (kg)/height (m)² (Table 4.4).

Table 4.4. Recommended values for body mass index (BMI) Royal College of Physicians (RCP) (1983)

	Males	Females
Underweight	20.0 and below	18.6 and below
Acceptable/normal	20.1–25.0	18.7–23.7
Mildly overweight	25.1–29.9	23.8–28.5
Obese	30.0 and above	28.6 and above

In theory, physiological measurements should not suffer from the problems of standardisation. However in practice, although the instruments themselves are technologically standardised, there is often disagreement on what are acceptable or problematic values. The guide values for BMI, one of the better established measures, are a case in point. In U.K., these are periodically reviewed and amended; internationally, different thresholds apply to different countries. What to take as critical threshold values for other indicators, such as volumetric flow measures for respiratory conditions, is the subject of much debate in the medical literature.

4.1.4 Surveys Measuring Health Behaviours (Lifestyle)

Although the usual outcome variable in a survey is a measure of ill health or morbidity, behaviour with respect to established risk factors to health are commonly employed in surveys. The usual focus is on the ‘favourite four’:

- ◆ Alcohol
- ◆ Diet
- ◆ Exercise
- ◆ Tobacco

Alcohol

Heavy alcohol consumption is associated not only with a higher risk of liver damage but also with heart disease, high blood pressure, cerebral vascular disease and other problems.

Alcohol consumption is measured in ‘units’, that correspond to standard amounts of pure alcohol in different alcoholic drinks. In principle, a unit corresponds to the amount of alcohol in a half-pint of ordinary strength beer or lager, a small glass of wine, or one measure of spirit. A person’s weekly consumption of units is graded into ‘Safe Drinking’, ‘Moderate Drinking’ and

'Heavy Drinking'. This criteria has changed several times during the last twenty years.

The usual source of data is from a self-completion survey where respondents are asked to report on how many glasses of beer, lager, spirits and wine they have consumed during the last week. Clearly a person's self-assessment may be quite different from the above definition (e.g. what is 'ordinary' strength beer, how much is a glass?). In general, there is a tendency to under-estimate,³ and the accuracy of both the recall and the assessment will vary between people.

Diet

In principle, measurement of a person's diet should include measurement of their intake of calories and protein and vitamins, etc., and consumption compared to standard values. This is an intensive and onerous data collection procedure (for both respondents and researchers) and is therefore usually only used for particular groups (e.g. young children).

In practice, therefore, except in specialised nutrition surveys, data collection and analysis are rarely that detailed. Instead, the more usual source of data is the self-report of the kinds and amounts of foods eaten during a fixed period. Comprehensive recording of foods consumed would also require an extensive data collection instrument, in which respondents would be asked to record whether or not they have consumed any one of a list of food items and how often during the reference period. Usual practice is to concentrate on 'marker' foods such as fruits, vegetables, bread and sugar.

There are all kinds of difficulties in interpreting survey responses because of variability in the quality and size of portions, as well as the usual problems of accuracy and recall. There are several ways for improving the usefulness of the data that is collected from surveys. For example:

- ◆ Respondents can be asked about their shopping habits and these data can be used to demonstrate inequalities in access to different kinds of food;
- ◆ Respondents can be asked about their cooking and eating habits.

Exercise

It is clear that some exercise is good for health, with considerable evidence linking lack of exercise to morbidity. However, apart from a physiological

³ Thus, the national annual consumption estimated from national surveys of drinking behaviour is much lower than the volume estimated to have been sold based on customs and excise figures.

measure, units of energy expended, which is impractical to measure, there is no agreed standard set of measurements.

Some researchers use changes in pulse rate during exercise as a surrogate process measure, but this is also constrained to specific contexts such as in a gym. Others rely on self-reports of exercise taken during the previous week. The main difficulty lies in combining reports of different forms of exercise into a single measure.

Tobacco

Tobacco is associated not only with a higher risk of lung and other cancers, but also with heart disease, other cardio-vascular morbidity, and increased vulnerability to many other conditions.

Measurement should focus on nicotine intake but, of course, it is difficult to measure directly, especially when one wants to estimate levels of passive smoking.

The usual source of data is from a self-completion survey where respondents are asked to report on how many cigarettes or pipes of tobacco they have smoked during the last week. Although cigarettes (and cigars and tobacco) are clearly graded according to their nicotine content, there is no way of taking into account the extent to which smokers only smoke part of a cigarette. And, as with the above examples, there are the perennial problems of accuracy and of recall.

4.2 Regular Health and Lifestyle Surveys

Several surveys are routinely conducted to measure the health of the population of England. The major ones include:

- ◆ HSE ([Box 15](#)).
- ◆ Surveys of Psychiatric Morbidity in Great Britain.
- ◆ GHS.
- ◆ The Omnibus Survey.

HSE is a major tool for monitoring the nation's health. It has been conducted annually since the early 1990s and since 1995 the sample has included measurements of the heights of children aged 2–15 as well as adults. The survey consists of an interview and some basic physiological measurements. Every year the survey covers blood pressure, anthropometry, cigarette smoking, alcohol consumption and self-assessed general health. Each year there are also a set of special themes relating to particular aspects of physical functioning, symptoms and lifestyle. The data are available from the Data Archive at the University of Essex [[140](#)].

Box 15**The Health Survey for England**

From 1991 to 1994, the survey focussed on cardiovascular disease and associated risk factors. In 1995 and 1996 the major focus of the survey was atopic disease (asthma, eczema, etc.), accidents and disability. In 1997, the survey report focussed on the health of young people aged 2–24, and combined data (in report only) from 1995–1997. In 1998, the focus was on cardiovascular disease. In 1999, the focus was on ethnic groups and in 2000, older people and social exclusion. In 2001 the focus was on non fatal accidents, disability, respiratory conditions, atopic conditions, fruit and vegetable consumption. In 2002 the focus was on the health of children and young people, the health of infants aged under one, and their mothers.

The Survey of Psychiatric Morbidity in Great Britain was conducted by Office for Population Censuses and Survey (OPCS) during 1993–1994. The survey was repeated in 2000. The report *Psychiatric Morbidity Among Adults Living in Private Households, (2000)* (is available from the ONS). It presents the findings of a survey of psychiatric morbidity among adults aged 16–74 living in private households in Great Britain that was carried out between March and September 2000. It was commissioned by the Department of Health, the Scottish Executive and the National Assembly for Wales. It is one of a series of surveys of different population groups which began in 1993 and is a repeat of the first of these surveys, which covered 10,000 adults aged 16–64 living in private households. The data for all these surveys is deposited at the Data Archive.

Both the GHS and the Omnibus Survey frequently have supplements relating to aspects of health and lifestyle.

GHS is a multi-purpose survey conducted by the Social Survey Division of the ONS. The survey started in 1971 and has been carried out continuously since then, except for a break in 1997–1998 and 1999–2000. The GHS is carried out for a number of government departments and provides information for planning and policy purposes, including to monitor progress towards achieving targets. Data are collected on housing, employment, education, health and family information from approximately 9000 households in Great Britain. The GHS has a modular structure and other subjects, e.g. elderly, smoking, drinking, contraception, hearing and childcare are covered periodically, and new topics are introduced from time to time. The data are available from the Data Archive at the University of Essex [140].

The Omnibus is a multi-purpose survey developed by ONS for use by non-profit organisations. The Omnibus allows customers to receive results

quickly, yet retaining the hallmarks of high quality – a random probability sample, high response rates and advice on questionnaire design. The survey is carried out in two of the three months each quarter. Approximately 1800 adults are interviewed each survey month, and the average response rate is 65%.

Other health and lifestyle surveys

In addition to these, a variety of large-scale major health and lifestyle surveys have been conducted over the past twenty years, funded by commercial, charitable and government agencies at both national and sub-national levels. Their data still have considerable potential for reanalysis. The Data Archive holds the individual level data for many of these. Interview schedules for these surveys often exceed 50 pages in length. They frequently include short copyrighted health-status questionnaires such as the GHQ or parts of the SF-36.

4.3 Health Status and Symptom Report Questionnaires

4.3.1 Introduction

This section discusses a major growth area in health questionnaire development – the many hundreds of instruments (questionnaires and checklists) often intended for medical assessment, screening and outcomes measurement. Though some cover very similar themes to the health and lifestyle surveys, there is a difference of emphasis, as most of these are intended to provide a health score, or diagnostic assessment. Consequently, they may not be ideal for epidemiological purposes; they may have to be completed by a medical professional, and not be appropriate for use in interviews, and they may be too long for self-completion. Moreover, many are strongly copyrighted and can only be used with the author's or publisher's permission. This may be given freely when the purpose is to preserve the integrity of the instrument and ensure that a body of comparable data is developed. Some instruments can only be used on payment of licence fees. In both instances, it may be difficult to obtain details of the scoring systems of the instruments without approaching the authors or publishers.

Fortunately, there are several detailed and wide-ranging reviews that can assist in the selection of a suitable measure [5,141,142]. Ann Bowling has produced some of the most comprehensive reviews in this area *Measuring Health* [4] and *Measuring Disease* [143] give details and references for several hundred of the more widely used measures.

Instruments tend to cluster into two overlapping groups. The first is the *generic model* which seeks to establish standard universal measurements

of health that can be applied to any group in any setting. The interest in standardisation may arise from a simple concern to minimise confusion and duplication of effort, or, it may be part of a drive to establish common databases for the monitoring and planning of healthcare (Box 16).

Box 16

Gold Standard for the Measurement of Health

“To meet the needs of the 1990s, information about general health outcomes must be added to the nation’s healthcare database. Minimum standards of comprehensiveness should be adopted to monitor the health of the general population and evaluate healthcare policies. A core set of generic health outcome measures should be standardised and adopted to compare the relative burden of medical and psychiatric conditions and relative treatment benefits. It is now practical to include a standardised core set of general health measures across applications (e.g. general population surveys, clinical trials) while supplementing this core according to the needs of a given study. The resulting comparisons would greatly advance understanding of the interpretation of health measures for all applications. Adoption of a standardised core set of health measures should be a high priority” (Source: *SF-36 Manual*, ch. 11 p.18).

This approach assumes there are universal constructs of physical, mental and social health across similar cultures and societies. If an instrument is a valid measure of these constructs then measurements can be compared between different settings and between groups with different medical conditions.

Most Multi-dimensional Health Status Profiles (MDPs) and indices, such as quality of life years lost, are grounded in this generic approach.

The second direction of development for health status measures could be described as the *condition specific model*. Unlike the generic approach it is not concerned to establish global standards and may question the validity of such standards. Rather, it aims to develop instruments that reflect the priorities and perspectives of groups with particular conditions, or those who are in receipt of similar kinds of healthcare. Some of these instruments will cover specialised topics, others measure more general health status. But the choice of domains, the associated definitions and selection of items, take account of specific interests.

The two approaches differ on their preferred strategies for developing and testing instruments. Evidence on the properties of the instruments, such as correlations with other measures, and demographic variations in

Table 4.5. Advantages and limitations of generic measures

Advantages of generic measures

- More likely to have been extensively developed
- Psychometric properties may be better known
- More likely to have supporting baseline/normative data from population and patient surveys
- Constructs may be sufficiently general to be widely relevant
- Provide an overview of health status/impact of illness from a single instrument
- Relatively cost-effective approach, provided the domains are independent and relevant

Limitations

- Can be hard to tell precisely what they mean and measure; this can present difficulties when selecting instruments and interpreting results
 - May miss critical factors for individual conditions, interventions or patient groups
 - May collect superfluous data; only a subset of domains are relevant to most applications
 - May be oversensitive to effects of confounding factors
 - May not provide sufficiently specific information for a number of administrative and audit applications such as: evaluating quality of care, monitoring variations in protocols and interventions, or planning the details of service development
-

scores, may be interpreted differently. The generic model may place more emphasis on psychometric criteria in the testing and developing of measures (Table 4.5), while the condition-specific model approach may stress qualitative methods of testing for validity and acceptability, alongside psychometric criteria.

The instruments

This section briefly introduces some of the major instruments under the following headings, indicating which are most likely to be suitable for use in surveys:

- 4.3.2 Measures concentrating on physical symptoms, functioning and general health.
- 4.3.3 Measures of mental health and mental illness.
- 4.3.4 Measures of social adjustment.
- 4.3.5 Measures of social support.
- 4.3.6 Measures of children's health.
- 4.3.7 Measures for older people.
- 4.3.8 Measures for people with specific conditions:
 - cancer.
 - asthma and respiratory conditions.
- 4.3.9 Multi-dimensional Health Profiles (MDPs).
- 4.3.10 The Health Index.

This is only a selection of the available instruments, which have been selected because, in our view, are the most likely to be suitable for measuring health inequalities. The criteria used to make the selection include the following:

- ◆ Brevity (number of items)/time and ease of administration questionnaires it should be suitable for postal use.
- ◆ Suitable for use by non-specialist interviewers.
- ◆ Wording and topic coverage that is suitable for self-completion (e.g. they should not induce anxiety or distress).
- ◆ Availability of published results – though this could range from one specialised study to a large body of publications with supporting databases of results from prior studies.
- ◆ Some evidence of psychometric properties and other tests of validity.

Not all of our selections meet all these criteria. When possible, we will note major drawbacks, but our knowledge is limited to only a sample of the literature on most of these instruments. Moreover, there are simply not enough publications on a number of the measures to be able to make all these judgements.

For more detail, we strongly recommend the two books by Ann Bowling ([Section 4.3.1](#)). Coverage of instruments produced post-2000 may be patchy.

Finally, we must emphasise that the following should be treated as a list of suggestions rather than recommendations.

4.3.2 Measures Concentrating on Physical Symptoms and Functioning

Functional status measures

These are some of the oldest health status questionnaires and are designed to monitor levels of disability as they affect the performance of basic activities such as feeding, dressing and bathing. They are also known as indexes of Activities of Daily Living (ADL), or as measures of dependency. They have been mainly developed for use in institutional settings and are often intended for rapid assessments by health professionals. More recent versions cover tasks more associated with living in the community, such as shopping and cooking.

Two of the best known in this group – the Barthel Index [144] and the WHO Functional Scale (WHO, 1979) – are only suitable for use by medical professionals. However, both are quick and relatively easy to complete, and provide simple overall assessments (e.g., the WHO instrument rates people on a 5-point scale). They, or similar measures of dependency, could be included in standard electronic records for older people needing nursing or

residential care, or high levels of domiciliary care. The Royal College of Physicians and The Royal College of General Practitioners have recommended the Barthel Index for screening elderly people in general practice.

Several instruments exist in multiple versions – some intended for professional assessment and some for self-completion. It is worth considering a self-completion version of the Functional Status Index [145]. This is predominately intended for use with adults suffering from arthritis, but even the short 18-item version covers the full range of topics associated with an ADL measure.

There are several questionnaires to assess functional status in the general population that have epidemiological potential. The Sickness Impact Profile (SIP) [146], though originally intended to measure changes in behaviour due to sickness and the outcomes of care, has been suggested for use in health surveys, though the full 136-item schedule is too long for self-completion. The Functional Limitations Profile (FLP) [147] is a modified version of SIP intended for use in the U.K. rather than the United States. Patrick et al. [148] have also developed the Lambeth Disability Screening Questionnaire, a checklist which is available in two versions; one for households, including questions such as: “does anyone in the household have difficulty with . . .?”, and another for individuals. Both are intended to be used to identify levels of disability in a population, as well as a screening instrument to identify individuals or households with problems.

Currently, functional status is as likely to be measured by a domain taken from a multi-dimensional instrument, than a single topic measure. The Medical Outcomes Survey (MOS) instruments (of which the SF-36 is the best known) have sections on functional status which are discussed later. The Multi-dimensional Functional Assessment Questionnaire (MFAQ) – developed for the Older Americans Resources and Services Program (Duke University paper 1972–1975) – is a multi-dimensional health measure, but its first section (intended for interview) is a wide ranging checklist for assessing function in older people.

Many condition-specific instruments have short sections on functional status that could be employed in population surveys, especially with older people. One of the most relevant is the Stanford Health Assessment Questionnaire (HAQ), [149], originally intended for use with adults suffering from arthritis, but equally useful as a short general assessment of the difficulty of performing everyday tasks.

Measures of pain

A wide range of instruments are available both for measuring pain in general and pain associated with specific, often chronic, conditions.

Several measures concentrate on pain associated with musculoskeletal disorders. A few of the better known examples include the Back Pain

Classification Scale [150]; sections of the Arthritis Impact Measuring Scale (AIMS2) [151]; and the Oswestry Low Back Pain Disability Questionnaire.

More general coverage is provided by the lengthy McGill Pain Questionnaire, while the Self-rating Pain and Distress Scale [152] is intended as a brief measurement of mood and behaviour changes associated with acute pain.

Several of the MDPs (Section 4.3.9) contain items referring to pain. These include all versions of the MOS questionnaires, from the 149-item version to the SF-6, the Dartmouth Coop Function Charts and the Nottingham Health Profile.

4.3.3 Measures of Mental Health and Mental Illness

There are many measures of mental health and illness, mostly concerning depression. They are often well established, widely validated and theoretically based.

There are several groups of measures. One set is concerned with cognitive functioning, especially in older people, and include tests of memory and general knowledge. The 10-item Mental Status Questionnaire [153] has been widely used and has been influential in the development of subsequent instruments such as the Comprehensive Assessment and Referral Evaluation Schedule CARE [154]. Even in its short form, CARE is probably too long for survey use; it is also rather specialised, with a focus on distinguishing dementia from depression in the elderly. However, there are several shorter instruments intended to establish the incidence of memory problems and the onset of dementia that might be practical for surveys. These include the community version of the Geriatric Mental State Test [155] and various versions of the Mental Status Questionnaire [156–159]. The Hodgkinson abbreviated Mental Test Score and Mental Test [160] are two further short tests of memory and confusion.

There are several widely used screening instruments for depression whose shorter forms appear quite regularly in population health surveys. The Carroll Rating Scale for Depression [161] is a self-administered version of the Hamilton Rating Scale for Depression [162]. Better known are the General Health Questionnaire [163], which, despite its title, is a measure of psychiatric symptoms, predominately depression, the Beck Depression Inventory [164] and the Zung Self-rating Depression Scale [165]. Anyone planning to use extracts from these instruments in surveys should note that permission may be required to reproduce the questions.

Slightly more general measures of mental health in the community have been produced. The 38-item Mental Health Inventory [166] (developed as part of the MOS programme) is the longest of these. It has been used in several large surveys to measure psychological distress and well being. However, unless one is intending to dedicate an entire survey to mental health issues, it may be preferable to use shorter mental health

assessments, such as the section on mental health in the SF-36. Most Multi-dimensional Profiles include short sections on mental health.

4.3.4 Measures of Social Adjustment

Measures of social adjustment are closely related to measures of mental health. They attempt to record the extent of ‘feeling at home’ in the local community and the ability to cope with immediate social surroundings. They are frequently, but not exclusively, intended for use with older people, but other measures (or parts of longer measures) are concerned with the problems of social activity and integration for people with severe physical conditions.

A few instruments are designed exclusively to measure social adjustment. The Katz Adjustment Scales [167] concentrate on the social adjustment of psychiatric patients following treatment, but parts of it could be more widely used. The Crichton Royal Behaviour Scale is intended for use in residential care and requires an assessor who knows the person being assessed. Again, most of the longer measures of health status include sections on social adjustment, which could be extracted (with permission) for specific applications. For most purposes, it may be helpful to combine such an extract on social adjustment with a measure of social activity or social support.

4.3.5 Measures of Social Support

While measures of social adjustment record an individual’s perceptions of their surroundings, measures of social support focus on the links between the individual, their family and the community. They measure the level and quality of contact with friends and relatives. Predictably, they can be very culturally specific, including questions on memberships of community organisations such as clubs, societies and churches that are unlikely to be relevant to all groups of respondents. The HSE has contained a standard set of items on social support since 2000.

4.3.6 Children’s Health

There is a well-established tradition of various forms of mental health, intelligence and psychiatric screening tests for children. There are also a few instruments recording the symptoms of children with specific conditions, especially childhood cancers [168] and asthma [169,170]. However, there are very few measures for children and most are completed by proxy.

Bowling in 1997 [4], noting the lack of measures for children, comments that “there have been attempts to develop scales for children . . . , but these are (still) relatively under-used”. She goes on that “Medical Outcomes Trust in the U.S., together with the Children’s Health and Quality of Life Project team have recently developed child health forms for age groups 5–15 and

infant pre-school, which are parent completed, and middle childhood–adolescent, which are completed by the child” [171].

4.3.7 Measures for Older People

Most of the functional status measures and Instrumental Activities of Daily Living (IADL), covering more social topics than ADL scales, are predominately intended for use with older people, though not necessarily validated. Indeed, many of the older functional status measures are better regarded as measures of dependency, or the need for nursing support, rather than measures of general health or health concerns.

Widely used examples of questionnaires covering the impact of health and disease on the life of older people include the already mentioned FLP and the Lambeth Disability Screening Questionnaire (Section 4.3.2), whose 25 items cover mobility both at home and outside, specific physical functions and impact on daily activities. The 22 items of the individual version cover a similar range of topics. Both versions may be distributed by post for self-completion or may be used in interviews.

The MFAQ (Section 4.3.2) could be regarded as a general MDP, but as already noted, its first section which is intended for interview, is an extensive checklist for assessing function in older people.

Some of the MDPs (including the SF-36) are criticised for focussing too heavily on relatively high levels of physical mobility and activity, such as sports performance, and fail to cover the sorts of everyday tasks, that may be problematic for older people. It is important not to assume that general health status questionnaires will validly represent the health status of older people. This is an area in which it is particularly important to pilot an instrument before use.

Several of the instruments mentioned earlier, under ‘mental health’, were designed to distinguish dementia from depression in older people. These include the community version of the Geriatric Mental State Test [155] and several versions of the Mental Status Questionnaire [156–159]. The Hodgkinson abbreviated Mental Test Score and Mental Test [160] are the two further short tests of memory and confusion that are mainly administered to older people.

4.3.8 Questionnaires for People with Specific Conditions

There are questionnaires to record symptoms, and experiences of coping with both treatment and symptoms, for most major medical conditions. By their very nature they are unlikely to be suited to general population use. However, a population survey might want to include questions on the symptoms of some of the more widespread conditions such as asthma. These instruments may help map inequalities in the outcomes of care, symptom management and impact on life.

Caution is required, as most of these instruments are usually administered in controlled settings, often as part of a care regime. Although most are sufficiently short and clear to be used in wider surveys, it is unclear whether they are ideal or appropriate for this purpose. It may also be unethical to use some of the instruments relating to cancer and mental health, without providing respondents with counselling or some form of support.

Measures for people with cancer

There are many questionnaires to monitor the health status of people with cancer, most of which include items on the side-effects of interventions. They have a long history; the Karnofsky Performance Status Scale, a very limited assessment of ability to perform normal activities, was published in 1949.

Some of the shorter instruments include:

- The European Organisation for Research on Treatment of Cancer (EORTC) has developed core questionnaires with 30 and 36 items and various modules relevant to specific cancers [172–174].
- The Rotterdam Symptom Checklist has been recommended by an MRC working party and exists in various forms mostly with 30–40 items [175].
- Fayer and Jones produced a diary card for cancer sufferers that is sometimes referred to as the Medical Research Council Scale, which records details of vomiting, activity, mood, anxiety and overall health [176].

It is not always clear whether such measures are intended to test for side-effects of therapies in clinical trials, or to monitor and treat in individual patient care.

Some of the more widely used short instruments for self-completion (or interview) include:

- The Cancer Functional Living Index [177];
- The short version of the Cancer Rehabilitation Evaluation System (CARES) [178].
- The Ontario Cancer Institute/Royal Marsden Scale [179].
- The Cancer Quality of Life Index [180,181].

Measures for people with asthma and respiratory conditions

At least 10 major instruments have been developed for measuring health status and quality of life of asthmatics in clinical trials. Most of these are too long and elaborate to use in routine care or population surveys.

For example, two of the more patient-centred instruments are the Living With Asthma Questionnaire [182] and the Asthma Quality of Life

Questionnaire [183]. Both were developed from interviews and group work with patients.

In the group work to develop the first of these, Hyland's participants reported effects of their condition that could be grouped under 11 headings: social/leisure, sport, holidays, work and other activities, sleep, colds, mobility, use of medications, effects on others, comments on doctors, dysphoric states and attitudes to asthma. Physical symptoms were excluded from the final 68-item questionnaire, as the aim was to develop an instrument that concentrated on disability.

Asthmatics who participated in Marks' groups were asked to consider the impact of their asthma on most aspects of their life, including: usual functioning, moods, feelings and relationships, and the consequent restrictions on life and overall well being. Though they were not asked about physical symptoms, these were mentioned spontaneously in discussion. Participants mentioned: physical symptoms, physical incapacity, control issues, emotional states, role fulfilment, social interaction and general health perceptions. Based on the suggestions, principal component analysis suggested six core domains: breathlessness (shortness of breath and reports of physical activities that induced breathlessness), concerns (mainly general concerns for health and future, but also feelings of lack of control and dependency on medication), control (understanding, feelings of coping and being in control), cough (sputum, congestion and cough), mood (a combination of fatigue and anxiety), social (restrictions on social life and sport, feelings of underachievement). These are all covered in the resulting 20-item questionnaire.

Shorter versions are being developed. Though too long for regular use in routine care, these instruments give an indication of the full range of outcomes that might be considered, the interrelation between the different outcomes, and an idea of what is most important to patients. They are typical of instruments designed to be administered to people with specific medical conditions.

4.3.9 Multi-Dimensional Health Status Profiles

MDPs represent a relatively recent approach to health status measurement. They cover several dimensions of health that might have previously been measured with separate instruments. (Box 17).

MDPs range in length from more than 140 to less than 10 questions and all but the shortest provide information on at least six different dimensions of illness and health. There are between 10 and 20 of these instruments in reasonably common use and most of these major instruments are listed in Table 4.6. As far as length/cost of administration/completion time is concerned, they can be considered in three groups: those with less than 15 questions (normally used alongside other instruments); those with 15–50 questions; and those with more than 50 items.

Box 17

A multi-dimensional picture of an individual’s health

It takes 5–10 min to complete the 36 items on the SF-36 questionnaire, and there are several scoring schemes that provide a health profile such as the following based on the instrument’s nine domains.

Physical functioning	95%
Social functioning	70%
Energy/vitality	40%
Physical impact on social role	80%
Emotional impact on social role	60%
Mental health	65%
Experienced pain	80%
General health	65%

It is typical of the sort of picture provided by an MDP questionnaire.

MDPs are *profiles* (Sections 7.2 and 7.4.3) in that they separately report results for each of their constituent domains. Many oppose attempts to collapse these scores to a single number and turn the profile into an index, but there is a great deal of interest in ways of doing this. Much of this comes from healthcare purchasers, planners, administrators, policy makers and associated researchers. They require an index with a single score for overall health that can be used for purposes such as comparing the health gain produced by different services, units and interventions. In one respect,

Table 4.6. Some of the more widely used multi-dimensional health profiles

<i>Name of instrument</i>	<i>Number of items</i>	<i>Number of scored dimensions</i>
Duke–UNC Health Profile	63	4
Duke Health Profile (17-item selection from Duke–UNC Profile)	17	4 (6)
Nottingham Health Profile Ptl	38	6
Sickness Impact Profile and Functional Limitations Profile (U.K. version of SIP)	136	12 (2)
Dartmouth Co-op Function Charts	9	9
McMaster Health Index Questionnaire	68	3
Medical Outcomes Survey (MOS) Full 149-item Functioning and Well-Being Profile (FWBP)	149	35 (8)
MOS FWBP-C (condensed version of full instrument)	113	20 (4)
MOS short form – 36-item instrument	36	9
MOS short form – 20-item instrument	20	6
MOS short form – 6-item instrument	6	6
EuroQol (6-item profile)	6	6

indexes and profiles are not radically different, as most profiles use the principles of indexation to combine component/item scores within each domain. Nevertheless, this type of indexation seems more widely acceptable than collapsing the different domains into a single number, partially because the single value index assumes that the same weightings for different aspects of health will be valid for all groups of respondents.

Domain coverage

Most MDPs cover a similar range of core topics, though they may describe them rather differently. The domain names of several MDPs are listed in [Table 4.7](#). Domains with similar names may be measured in quite different ways, with different sets of components and different abilities to measure change or record extreme values. It is common to find them containing items that are culturally specific, or more appropriate to people of certain ages or health status. [Table 4.7](#) lists the domain coverage of the main instruments, but, copyright permitting, it may be best to mix and match sections from different instruments. When doing so, it is worth remembering that factor analysis or related statistical techniques may have been used to maximise the independence of domains and that sections from several instruments are unlikely to have the same degree of independence. However, this is unlikely to pose serious problems unless the domain scores are to be combined into an overall rating of health.

It is worth considering using MDPs when:

- ◆ There is a need for self-reports of health and the impact of illness in non-clinical terms.
- ◆ These need to cover a relatively standard range of topics.
- ◆ There is a need to supplement conventional bio-medical assessments.
- ◆ There is a need for 'generic' measures to produce information for comparing different services and interventions.

MDPs may not be useful when:

- ◆ Resources are very limited and very specific information is required.
- ◆ Data of immediate relevance to services are required for audit and assessment.
- ◆ There is a need for a single index value of health.

Most self-report multi-dimensional instruments are likely to be more sensitive to non-health service influences than more specific or clinical measures, so it can be difficult to decide what part of any effects are due to health services or their lack.

The relationship between what these instruments measure and lay perceptions of health and health concerns are rarely clear. Before

Table 4.7. Domain coverage of some major MDPs

<i>Name of instrument</i>	<i>Number of items</i>	<i>Methods of application</i>	<i>Scored dimensions</i>	<i>Number of scored dimensions</i>
Duke–UNC Health Profile	63	I S	Symptom status, physical function, emotional function, social function	4
Duke Health Profile (17-item selection from Duke–UNC Profile)	17	S	Physical health, social health, mental health, general health	4 (6)
Nottingham Health Profile Ptl	38	I S	Physical mobility, pain, sleep, energy, social isolation, emotional reactions	6
Sickness Impact Profile and Functional Limitations Profile	136	I I A S	Categories: ambulation, body care and movement, mobility, household care and management, recreation and pastimes, social interaction, emotion, alertness, sleep and rest, eating, communication. Dimensions: physical, psychosocial	12 (2)
Dartmouth Co-op Function Charts	9	I A S	Physical condition, emotional condition, daily work, social activities, pain, change in condition, overall condition, social support, quality of life	9
McMaster Health Index Questionnaire	68	I T S	Physical functions (physical activities, mobility, self-care, communication); emotional functions (self-esteem, personal relationships, critical life events, thought about the future); social functions (general well-being, work, social role performance, material welfare, family participation and friendships)	3

Table 4.7. (Continued)

<i>Name of instrument</i>	<i>Number of items</i>	<i>Methods of application</i>	<i>Scored dimensions</i>	<i>Number of scored dimensions</i>
Medical Outcomes Survey (MOS) full 149-item Functioning and Well-Being Profile (FWBP)	149		Physical functioning, satisfaction with physical ability, mobility, role limitations due to physical health, role limitations due to emotional problems, unable to work due to health, unable to do housework due to health, social activity limitations due to health, sexual problems, satisfaction with family life, overall happiness with family life, marital functioning, anxiety (I and II), depression/behavioural emotional control (I and II), positive affect (I and II), feelings of belonging, cognitive functioning, current health, prior health, health outlook, health concern, resistance to illness, health distress, energy/fatigue, sleep quantity, optimal sleep, sleep disturbance, sleep adequacy, sleep somnolence, sleep shortness of breath or headache, snoring	35 (8)
MOS FWBP-C (condensed version of full instrument)	113		20 scales and 4 summary indexes	20 (4)
MOS Short form – 36-item instrument	36		Physical functioning, role functioning/physical, role functioning/emotional, energy/fatigue, emotional well-being, social functioning, pain, general health, health change	9
MOS Short form – 20-item instrument	20		Physical functioning, role functioning, social functioning, mental health, health perceptions, pain	6
MOS Short form – 6-item instrument	6		Physical functioning, role functioning, social functioning, psychological distress/well-being, health perceptions, pain	6

considering their use in population surveys it is important to test whether the items accurately reflect the concerns of major users of healthcare, such as older people. The utility of a general questionnaire may be diminished if the phenomena of interest tend to be concentrated in large sub-groups of the population, for which a general instrument may not be well suited.

Most MDPs have been subjected to basic psychometric assessments, both during their development and subsequently. The testing tends to concentrate on the capacity to measure change over time (as required for use in clinical trials), but it may have some relevance to epidemiological applications. More details of the tests can be found in [Section 7](#).

4.3.10 A Single Health Rating – the Health Index

Several of the MDPs described above provide scoring systems to aggregate the separate domain scores into an overall health rating. The approach is very attractive to policy makers because it provides a single measure on which to evaluate the gains that can be achieved by healthcare interventions and hence to devise standards for cost effectiveness. The difficulty is how to translate the overall health rating into a metric that can be compared with the resource inputs of different interventions. We consider two approaches: Disability Adjusted Life Expectancy (DALE) and Quality Adjusted Life Years (QALYs).

Disability Adjusted Life Expectancy

Measures of life expectancy can be adjusted to take account of disability, hence DALE or DALYs.

Measures that adjust life expectancy for disability have been used since the beginning of the 1970s when both Japan and the U.S. presented data on Life Expectancies Adjusted for Disability (LEADS). The Japanese calculations illustrated the effect of various adjustments ([Table 4.8](#)).

Table 4.8. Average life expectancy and average loss of years due to health impairments

Age	Year	Average life expectancy (a)	Loss due to illness (b)	Losses due to mental or physical disabilities (c)	Average loss (b + c – overlap) (d)	Average life expectancy minus average loss (a–d)
0	1966	70.88	2.00	0.94	2.80	68.08
	1970	71.93	2.35	0.98	3.15	68.78
15	1966	57.38	1.83	0.90	2.59	55.20
	1970	58.45	2.13	0.91	2.89	55.56
65	1966	13.81	0.78	0.47	1.18	12.63
	1970	13.97	0.91	0.51	1.35	12.62

The adjustments attenuated improvements (compare columns (a) and (d)), but the differences were obviously small. It is interesting to note that at the time, the LEADS were abandoned as an overall measure of health status.

With increasing longevity, these measures have become more widely used, and the OPCS has produced a manual on how these are calculated and how to interpret them [184].

These measures are similar to the DALYs used by the World Bank as a tool for prioritising and targeting healthcare interventions in many developing countries. The problem for using such an approach is that relying only on disabilities as a measure of ill health is relatively restricted, and health means different things to different people. Moreover, the logic of the measures has been criticised by Williams [185] as not taking into account the possible effectiveness of interventions; he and other economists have therefore searched for a method that explicitly focuses on the possible outcomes of healthcare interventions.

Quality adjusted life years

Their approach is to base the algorithm for converting the overall health rating into a metric on people's preferences for different states of (ill)health. They have conducted surveys where people evaluate different states of (ill)health. Various methods – for example, time trade-off in which people are asked to compare the value of being in state A for $n(a)$ years compared to state B for $n(b)$ years are used to convert these evaluations into a 'Q' or quality value, that is constrained to be between 0 (worst health) and 1 (best health). These 'Q's are used for weighting or 'adjusting' the years during which people are in those states of (ill) health, and the assessment of the cost effectiveness is then made by comparing the number of QALYs that are saved through the intervention.

This approach raises a number of problems. First, health means different things to different people (Box 18). Even in the very first study of this kind, doctors and patients had different views [186]. The QALY procedure is supposed to be 'democratic' because people are asked to rate health status. In fact, they are constrained by the design of the questionnaire instruments [187].

Box 18

Health means different things to different people

For example, Blaxter [188] identifies four dimensions of health:

- ◆ Unfitness or fitness.
- ◆ Disease and impairment or their absence.
- ◆ Experienced illness or freedom from illness.
- ◆ Psycho-social malaise or well being.

In addition, should we value patients or groups of patients equally, the QALY gives decreasing weights to older people (because they have lower life expectancies) and, therefore, less years that can be gained or saved through the intervention. However, there is survey evidence that the young and old alike place more weight on parents with young children and less on the old and very young [189].

People who suffer from illness generally adapt; and even if they do not adapt, their perceptions change. Some people may be more independent than others; in particular, older people, when asked, will still claim to be satisfied with their health [189] even though 'objectively' health deteriorates with age. This is because of the way in which the index is constructed by giving a 'floor' of 0 that most (although not all) interpret as equivalent to death, which means that the quality scores tend to be compressed towards one (the 'ceiling').

The final value of the QALY is calculated by adding together the years of life left and each year being weighted by 'Q' and then discounted on the basis that future years are worth less than current years. Typically, a 3% discount rate is used. There is little survey evidence to support this. High discount rates lead to a bias against educational or other interventions at young ages.

Planning context

Given the move towards public participation, is it appropriate to introduce an index as a basis for decision making where only a limited number of 'experts' are conversant with the criticisms?⁴ The costs per QALY or cost per DALY figures are based on average costs. Such figures are sub-optimal, and may not be easily transferred from one context to another (let alone from one country to another). Marginal costs of an additional intervention may be different from average costs. Given the other problems of interpretation, the public may well prefer disaggregated figures to overall DALY or QALY results.

⁴ Of course, one cannot expect everyone to be conversant with the statistical calculations and technical procedures underlying many such indexes; but we are referring here to the issues of valuation (discussed under the theoretical heading above).