

# Measuring disability in the general population using sample surveys and censuses

Roger Thomas

Survey Methods Unit, National Centre for Social Research

## Introduction

“Disability” is a popular survey topic in Britain because it is an area where the interests of central government departments, local government and advocacy groups for the disabled intersect. There is therefore a demand for trustworthy measures of the prevalence, the incidence, the nature and the severity of disabilities suffered by people living in the community, for use in censuses and surveys. Some major national surveys which include coverage of disability are shown in Table 1.

Most of the major sources of statistical information about disability shown are either continuous or regular periodic, though detailed questions on disability are not necessarily asked every year. Disability is also covered, though not usually in depth, in various other surveys carried out *ad hoc* for government and other sponsors. A question referring to disability was included in the 1991 Census of Population and again in 2001.

Important uses of statistics from these sources include:

- monitoring the general health of the population;
- formulating and monitoring disability benefit policies;
- assessing needs for care in the community and the role of paid and unpaid carers;
- assessing the impact of disability on employment and employability.
- assessing the impact of disability on quality of life.

The focus of much disability research is upon the elderly, a large and growing population subgroup amongst whom rates of severe physical and mental disability are relatively high. Other sources are primarily concerned with the population of working age and with the measurement of employment disability. There are growing concerns about mental illness as a major cause of functional disability - that is, disability that affects ability to carry out specific types of activity or to cope with particular situations in everyday life. When it comes to defining and measuring disability the approach of those with a public health orientation is not always consistent with those favoured by users whose prime interest is in, for example, eligibility for benefits, care in the community and demand for social services, or conforming to provisions of the Employment Disability Acts or requirements to provide housing suitable for the disabled.

Advocacy groups for disabled people and their carers are generally not in a position to commission large-scale surveys capable of producing sound national statistics, or even to carry out careful reanalysis of existing data sets, but tend to search what is published for findings and statistics that are supportive of their campaigns. They are often puzzled to find what appear to be inconsistencies between sources that they have assumed to be authoritative.

Table 1: National surveys which have attempted to measure disability

<b>Survey and sample size</b>	<b>Brief description</b>
The OPCS* Disability Surveys (published in 1986)	Population screen and follow up, estimating and comparing levels of disability, calibrating severity. Adults and children in general and institutional populations covered.
ONS General Household Survey (continuous 20,000 persons annually.)	Personal interviews with adults. Has included at least one question on general health and chronic illness in adults most years since the early 1970s. Special coverage of disability, caring etc in some years.
Health Survey for England (HSE) (continuous 17,000 persons annually)	Personal interview and health examination survey conducted for the Department of Health. Questions on chronic illness and disability. Special coverage of general health, including disability questions, in 1996.
National health surveys for Scotland and Wales.	Carried out regularly. Scottish survey similar to English HSE, Welsh survey uses postal methodology. Cover general health and disability. Welsh survey uses SF-36 instrument. <sup>+</sup>
Health Education Monitoring Survey	Carried out periodically for the Health Education Authority. Coverage of general health follows that of the GHS.
Family Resources Survey (FRS). 25,000 annually)	Carried out for Department for Work and Pensions <sup>o</sup> . Covers general health and disability with reference to eligibility for and receipt of benefits. Has attempted to replicate OPCS Disability Surveys.
Labour Force Survey (LFS). 60,000 households annually.	For ONS. Main source of statistics on reasons for non-participation in labour force. Covers disability as it affects employment and employability, sometimes using proxy household informant.
House Condition Surveys	Sponsored periodically by the DETR <sup>e</sup> and the Scottish and Welsh Offices. Cover disability as it affects housing needs, sometimes in considerable detail.

\* Now the Office for National Statistics (ONS)

<sup>+</sup> See below.

<sup>o</sup> Formerly the Department for Health and Social Security (DHSS)

<sup>e</sup> Previously the Department of the Environment, now the Office of the Deputy Prime Minister

### **Approaches to measuring disability**

Disability takes many different forms and measuring it satisfactorily turns out to be quite difficult. One reason for this lies in the inherent complexity of the concept. Another lies in the

differing administrative and political agendas of the users of census and survey data on disability, which lead them to focus on different aspects of “disability” and to seek to use the information in different ways.

In spite of their differing agendas, the major surveys and major users of survey results that have been referred to all require to measure disability in a way that is technically adequate. Six different criteria for assessing survey questions as measuring instruments can be distinguished and are listed below.

1. **Validity**, meaning that the question instrument is consistently and correctly interpreted by respondents across the whole population and measures only “disability”, and not other extraneous factors. Relevant technical issues here are:
  - *Face validity* – Do questions designed to measure disability appear relevant and plausible in terms of their wording, response scoring etc?
  - *Criterion validity* – Do members of groups recognised to suffer from disability score corresponding and consistently higher on a proposed measure than other individuals?
  - *Construct validity* – Does the proposed measure of disability show the expected pattern of correlation with other, conceptually related, measures?
2. **Reliability**, meaning that the questionnaire instrument does not generate large amounts of random response error and thus give highly variable results even where the true incidence and severity of disability remains constant, fogging measurement of differences between groups and of change over time.
3. **Freedom from gross or relative bias**, meaning that the question instrument does not exaggerate or understate the amount of disability either overall, or as between sub-groups (eg elderly and non-elderly).
4. **Sensitivity**, meaning that the question instrument detects disability where it exists even at low levels and distinguishes consistently between the more and the less disabled across the whole range of severity.
5. **Consistency over time**, so that summary statistics of prevalence and severity of disability can validly be compared when measures are taken at different points in time.
6. **Portability**, meaning that the instrument provides strictly comparable results on prevalence and severity of disability when used in different survey contexts.

The degree of rigour with which these criteria need to be applied varies according to the uses of the data. For example some users require only a measure which will roughly distinguish and order broad groups in terms of average level of disability suffered, for the purposes of a one-off survey. They need not be overly concerned about precise definition, consistency over time, sensitivity to small differences in disability status, bias in relation to objective measures, or random measurement error. But other users require to *define* disability in a *consistent* way and to monitor levels of prevalence and severity in a *reliable* and *sensitive* way across different population subgroups and different types of disability, and to detect *small changes* in this level *over time*, perhaps using *different survey vehicles*. Requirements of the second type are very much more difficult to satisfy than requirements of the first type.

In practice relatively few questions have been thoroughly and objectively tested against criteria 1-6. The fact that they may have obtained response distributions which are stable, or show regular trends from year to year, is not *in itself* sufficient proof that they satisfy the measurement criteria set out above. To apply the criteria we ideally require to use special question testing and comparison methods. Such methods are available, but their application costs some time and money - and of course the results that are obtained on the performance of measures may be disappointing.

**Questions asking directly about (limiting long-standing) disability**

---

General Household Survey (also HSE, FRS etc)

Do you have any long-standing illness, disability or infirmity? By long-standing I mean anything that has troubled you over a period of time or that is likely to affect you over a period of time?

Yes 1  
No 2

**If yes**  
What is the matter with you?

Does this illness or disability (Do any of these illnesses or disabilities) limit your activities in any way?

Yes 1  
No 2

Census 1991

Do you have any long-term illness, health problem or handicap which limits your daily activities or the work you can do?  
Include problems which are due to old age.

Yes 1  
No 2

---

This type of question depends upon self-classification (or, in some applications, classification by another household respondent). Most respondents seem able to answer and the questions are cheap and quick to administer. The GHS version is currently used on the GHS, HSE, HEMS, FRS and on other health surveys, though the second question (“What is the matter with you?”) is often omitted. These questions have been widely used over a long period, but in versions which differ from one another in ways that could affect the levels of disability reported. Nevertheless, through long and widespread usage and the publication of time series they are sometimes assumed to be “standard”, portable, general-purpose measures of chronic ill health and disability.

Even with standard versions of the questions, cognitive question testing has shown that respondents may interpret the key words (including “disability”) in inconsistent ways. The criterion for identifying disability is vague and measurement is insensitive and cannot identify, for example, persons who are *severely* disabled and therefore unable to live independently without support. When the question is used on different surveys (for example, the GHS and the HSE), survey context affects the rates of disability reported, indicating a lack of portability. Also, time trends seem anomalous, suggesting that measurement may be inconsistent over time (persons counted as “disabled” in 1999 who might not have been counted as “disabled” twenty or ten years ago.)

---

### Labour Force Survey (Spring 1997 onwards)

Do you have any health problems or disabilities that you expect will last for more than a year?

**If yes**

Does this health problem affect

- the kind of paid work that you might do or
- the amount of paid work that you might do?

---

This question is expressly designed for measuring the narrower concept of “employment disability” and is not suitable for more general applications. It involves hypotheticals (“...the kind of work you might do...”), which may lead to inconsistencies of interpretation.

### **Activities of Daily Living (ADLs) and Instrumental Activities of Daily Living (IADLs)**

These measures focus on *severe functional limitations leading to dependency* and do not in their present form aim to detect mild disabilities or disabilities of a kind that do not cause dependency. They comprise batteries of questions (about 10 ADL questions and about 9 IADL questions) designed to measure difficulty experienced in carrying out specifically described activities. They therefore do not require respondent to interpret terms like “disability”, nor to make general judgements about their condition. No information is collected on the cause or nature of the impairment(s) which lead to functional disabilities, or on implications for employment or employability.

The questions have high face validity and they conform to the general principle that to provide consistent and reliable information questions should focus on the concrete and the specific. However, there is rather little formal evidence on reliability, portability etc. It is possible to interpret the results either via some overall disability scoring system, or by looking at the responses to particular items separately.

The instruments obviously require much more questionnaire space and interview time than the question approaches described above. They have been included only occasionally in the GHS personal interview and other surveys, generally in questionnaire sections addressed to elderly people only. There is no particular reason why some more or less adapted version should not be applied to younger people, though of course the proportion found to have difficulties (disabilities) would in their case be very much lower.

### Example questions: ADLs

How difficult is it for you to -	Not difficult	Quite difficult	Very difficult	Impossible
Get up and down stairs or steps on your own?	1	2	3	4
Wash yourself all over?	1	2	3	4

### Example questions: IADLs

How difficult is it for you to -	Not difficult	Quite difficult	Very difficult	Impossible
Prepare a snack for yourself?	1	2	3	4
Deal with personal affairs (eg paying bills, writing letters)?	1	2	3	4

### **Purpose-designed disability sift instruments**

These include the Disability Sift Questionnaire which formed part of the OPCS Disability Surveys carried out in 1984-85, the Family Resources Survey Disability Follow-up question set (developed on the basis of the OPCS model) and the Health Survey for England Disability question set, which was influenced by the OPCS instrument and by a model developed by the World Health Organisation.

The instruments comprise batteries of questions to detect and determine the nature and severity of functional disabilities from which a person may suffer. They take a few minutes to complete. They are akin to the (I)ADLs in that they focus on operationally defined functional disabilities (i.e. the concrete and the specific). Many versions have been used and instruments may comprise anything from 13 to about 30 items. Some have built-in scoring systems to measure severity. They aim to cover physical and mental disabilities comprehensively in terms of several (in the case of the OPCS instrument 13) functional areas, including locomotion, personal care, seeing, hearing, communication, digestion and metabolism, mental faculties, etc. In the OPCS case they comprised the sift stage of more elaborate surveys of the condition of disabled persons.

### **Short Form 36 (SF-36)**

This is a 36-item questionnaire instrument produced as the outcome of a systematic and methodologically sophisticated research and development programme carried out in the USA. Considerable work has been done to check and improve the reliability and validity of the

instrument. A version slightly modified for use in the UK is available. The instrument may take up to 15 minutes to administer (by interview or by self-completion) and is intended to provide a comprehensive health profile, with interval scales measuring level of health on each of 8 dimensions. Of these, “Physical functioning” and “Role limitation (physical)” appear to reflect aspects of physical disability, while the “Social functioning”, “Role limitation (social)” and “Mental health” scores appear to reflect aspects of mental or emotional disability.

The instrument is concerned with symptoms and their effects on the life of the individual and does not deal with the nature of underlying diseases or impairments. Unlike the approaches so far described it does not provide interpreted scale points, but instead scores on factors representing dimensions of health, which enable any individual to be placed relative to other individuals (population norms are now available).

The SF-36 thus does not directly and in itself provide answers to questions like “How many people are (severely) disabled?”; but on reflection it is evident that imposition of cut-off points to identify “the disabled” is arbitrary, though no doubt essential for some administrative purposes.

### **EuroQol 5-dimensional Quality of Life classification of health states (EQ5D)**

This instrument was developed through a European collaboration, the aim of which was to provide a measure for evaluating the impact of health interventions on health-related quality of life (HRQL). High priority was given to providing a measure that can be generalised across different types of health problem and to keeping the instrument short, simple and easy for individuals to complete to describe their own health state; in fact it can usually be completed in well under five minutes.

Five dimensions of HRQL are distinguished and the individual’s health is assigned to one of three levels (“No problem”, “Some problems”, “Severe problems”) on each dimension. The “Mobility” and “Self-Care” dimensions seem to reflect quality of life aspects of disability and each scale point is interpreted (e.g. “Some problems with mobility”). The nature of the underlying disability and impairment is not referred to. The originators of the system also provide a “health utility score” for each health state, which supposedly allows individuals with particular combinations of disabilities to be ranged on a common interval scale measuring how “good” or “bad” their health-related quality of life is overall (according to the averaged collective judgements of their fellow-citizens). Alternatively, the five dimensional health state descriptions can be summarised on some other principle, such as “Some problem” versus “No problems”, or “No severe problems” versus “At least one severe problem”.

The drive for brevity and generality makes the characterisation of health states in EQ5D rather crude and lacking in detail. While in theory any of 243 possible health state descriptions might apply in a particular case, in practice nearly all individuals assign themselves to one of about 12 health states. The instrument seems to be best at distinguishing individuals whose HRQL is severely impaired by disability and in this area it seems to be reasonably reliable. It typically classifies about half the general population as having “No HRQL problems” and this probably matches common perceptions, even though fine distinctions are not made.

## Discussion

The concept of “disability” is complex and difficult to measure, partly because disability can take many different forms, partly because subjective ideas and standards of what constitutes “disability” vary. Disability, objectively considered, is a continuum along which all individuals in the general population are ranged (though not evenly), from the Olympic athlete at one end to the comatose patient at the other. At one extreme there is a large part of the population that a reasonable person aware of their age and normal activities would certainly classify as “not significantly disabled”. At the other extreme there is a smaller but very substantial group that most reasonable observers would regard as significantly disabled. In numeric terms this group is dominated by the elderly and would be more so were it not for the universal tendency to set the threshold of disability higher for the elderly (a young man of 16 who is physically unable to play football would be regarded by many as “disabled”, but not a man of eighty who cannot play football). Another hazy area is disability due to mental diseases or conditions. Between the extreme groups lies a third large group whose disabled/non-disabled status is a matter of judgement.

Therefore locating individuals in relation to a line between “(significantly) disabled” and “not (significantly) disabled” is bound to be to some extent arbitrary and contentious, as is assessment of degree of disability. “Disabled” status in turn has significant implications for the way an individual is treated, such as recognition that he or she cannot be expected to compete in the open labour market, or is entitled to benefits and support. Political and economic judgements may reasonably enter in, such as limits on the proportions of individuals that the country can afford to excuse from seeking employment and make eligible to receive disability benefits.

For administrative purposes and because users of the data often wish to quote numbers or proportions of people who are “disabled”, there is a strong demand on surveys to estimate numbers of persons falling above and below “disabled”, “severely disabled” or other cut-offs. Putting the responsibility on survey respondents to make these distinctions leads to inconsistencies. External criteria must therefore be invoked to draw cut-offs between discrete categories. One conceptual criterion used is “Ability to look after oneself” (cf. ADLs and IADLs); another is “Ability to undertake paid employment”. However, these concepts in turn are quite difficult to operationalise in a consistent and satisfactory way and the way in which they are applied tends to vary and change over time according to current standards of what levels of activity and what quality of life people (of different ages) are entitled to expect.

There are several important ways in which the approaches to measuring disability currently used in surveys and census differ from one another. For example, there are differences of aim and outcome between:

- focused attempts to measure disability as such *versus* approaches which treat disability as a facet of general health;
- approaches which aim to provide a single measure of disability valid across the whole spectrum *versus* approaches which provide a multidimensional disability profile;
- approaches which treat disability according to a medical “disease” model *versus* approaches which focus on health-related quality of life;

- the direct classificatory approach, in which respondents' interpretations of terms such as "disability" are critical, *versus* approaches which seek to apply functional criteria and infer disability status;
- simple one-question approaches *versus* approaches which rely on scores derived from responses to batteries of questions;
- approaches that aim to assign individuals to interpreted scale points, such as "Moderately disabled" or "Able to do X but not Y", *versus* approaches which yield scores on named dimensions, but where the position of an individual can be interpreted only relatively or in terms of a population norm.

None of the questions and instruments currently used in surveys and census in the UK has a complete methodological test pedigree in terms of the six criteria listed earlier in this paper, though measures such as the SF-36 are much more advanced in that direction than, for example, the single-question measures. In particular, questions that refer to "(Limiting) long-standing illness, disability or infirmity" seem to owe their current popularity to cheapness and simplicity of application, to precedent and to publication of results in accessible sources, rather than to proven merits as survey measures.