

**STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE**

CONFERENCE OF EUROPEAN STATISTICIANS

WORKING PAPER No. 6

**Joint ECE/WHO Expert Meeting on Health
Statistics Measurements**
(Ottawa, Canada, 23-25 October 2000)

ENGLISH

**AN INVENTORY OF HEALTH AND DISABILITY RELATED
SURVEYS IN OECD COUNTRIES**

Paper submitted by OECD¹

**DIRECTORATE FOR EDUCATION,
EMPLOYMENT, LABOUR AND SOCIAL AFFAIRS**

Occasional Papers

This series is designed to make available to a wider readership selected labour market and social policy studies prepared for use within the OECD. Authorship is usually collective, but principal writers are named. The papers are generally available only in their original language -- English or French -- with a summary in the other.

Comment on the series is welcome, and should be sent to the Directorate for Education, Employment, Labour and Social Affairs, 2, rue André-Pascal, 75775 Paris CEDEX 16, France. Additional limited copies are available on request.

The opinions expressed and arguments employed here are the responsibility of the author(s) and do not necessarily reflect those of the OECD

Applications for permission to reproduce or translate all or part of this material should be made to:

Head of Publications Service
OECD
2, rue André-Pascal
75775 Paris, CEDEX 16
France

Copyright OECD 2000

¹ Claire GUXE of the National Institute of Public Health, Copenhagen and Gaetan LAFORTUNE, OECD Health Policy Unit, Paris

ACKNOWLEDGEMENTS

The preparation of this paper would not have been possible without a voluntary contribution from the Government of Canada.

The authors also wish to thank the OECD Health Data National Correspondents for their support in identifying and providing access to the national health surveys covered in this inventory.

SUMMARY

There is strong policy interest in monitoring trends on the prevalence of chronic diseases and disability rates, both nationally and internationally, in light of rising life expectancy and population ageing. However, international comparisons of health and disability survey data are difficult because different instruments are used to measure various health and disability dimensions in national surveys.

This inventory examines the comparability of survey instruments used to measure health and disability in various OECD countries. It extends a similar inventory prepared by the Danish Institute of Public Health for Eurostat in 1999. Some 30 surveys from 23 OECD countries are reviewed in detail and compared. These include a mix of cross-sectional and longitudinal surveys, general health and disability-specific surveys, and surveys covering the entire population and some targeting the elderly population only. The domains covered include selected health conditions (chronic physical conditions, mental health and pain) and various activity limitations (ADL, IADL, limitation in work and social activities, and general limitations in usual activities). The main finding is that, beside a few items related to the prevalence of chronic conditions (both generally and for a few important diseases) and general activity limitations, current differences in measurement instruments limit the comparability of data only to those countries that are using the same instrument (e.g., SF-36, EuroQol-5D, HUI-3 or the WHO-Europe long-term disability list). The main problem is not “what” is being measured in various surveys (since the health dimensions and activity limitations tend to be fairly common) but rather “how” specifically these health conditions and limitations are measured. Unless progress is achieved in using some common instrument(s) to measure these health and disability dimensions, cross-survey (and cross-country) comparisons will remain limited.

RÉSUMÉ

On constate un intérêt politique marqué à mesurer l'évolution de la prévalence des maladies chroniques et de l'invalidité, tant au niveau national qu'international, en raison de l'accroissement de l'espérance de vie et du vieillissement de la population. Toutefois, il est difficile présentement de faire des comparaisons internationales des données des enquêtes nationales sur la santé et sur l'invalidité parce que les instruments de mesure varient d'une enquête à l'autre.

Cet inventaire a pour but d'examiner la comparabilité des instruments de mesure de la santé et de l'invalidité dans les enquêtes nationales de différents pays de l'OCDE. Il prolonge un inventaire similaire préparé par l'Institut danois de santé publique en 1999 pour le compte d'Eurostat. Quelques 30 enquêtes de 23 pays de l'OECD sont examinées et comparées en détail. Celles-ci comprennent des enquêtes transversales et longitudinales, des enquêtes générales de la santé et d'autres portant spécifiquement sur l'invalidité, et des enquêtes couvrant toute la

population tandis que d'autres se concentrent uniquement sur la population âgée. Les domaines couverts incluent divers problèmes de santé (les maladies chroniques physiques, la santé mentale et la douleur) et différents types de limitations d'activités (les activités de la vie quotidienne de type ADL, les activités instrumentales de la vie quotidienne de type IADL, les limitations dans le travail et les activités sociales, et les limitations générales dans les activités normales). La principale conclusion de cet inventaire est que, mis à part quelques items reliés à la prévalence des maladies chroniques (à la fois générale et pour un petit groupe de maladies importantes) et les limitations dans les activités normales, les variations existantes dans les instruments de mesure limitent la comparabilité des données uniquement aux groupes de pays qui utilisent le même instrument (SF-36, EuroQol-5D, HUI-3 ou la liste de questions sur l'invalidité à long terme de l'OMS-Europe). Le problème principal n'est pas ce qui est mesuré dans les différentes enquêtes (puisque'il y a beaucoup de points communs dans les dimensions de la santé et des limitations d'activités qui sont mesurées), mais plutôt comment ces conditions de santé et ces limitations d'activités sont mesurées. A moins que des progrès soient réalisés dans l'utilisation d'instruments de mesure communs dans les enquêtes nationales, les comparaisons entre pays resteront limitées.

TABLE OF CONTENTS

INTRODUCTION	5
SCOPE AND FOCUS OF THIS REVIEW	7
IDENTIFICATION OF SURVEYS	8
DEFINITION OF DISABILITY	13
A BRIEF HISTORY OF DISABILITY MEASUREMENT	15
Efforts to Promote International Standardization	16
ANALYSIS OF COMPARABILITY BETWEEN SURVEYS	20
Overview of Coverage	20
1. Health conditions	22
2. Activity limitations	26
SUMMARY AND CONCLUSIONS	42
Main Findings	42
Approaches to improve the comparability of health status and disability-related data	43
REFERENCES	48
ANNEX 1: SELECTED GENERIC HEALTH MEASUREMENT INSTRUMENTS	ERROR! BOOKMARK NOT DEFINED.
ANNEX 2: LISTING OF SELECTED HEALTH STATUS AND DISABILITY-RELATED ITEMS INCLUDED IN 30 NATIONAL SURVEYS (AND THE EUROPEAN COMMUNITY HOUSEHOLD PANEL)	ERROR! BOOKMARK NOT DEFINED.

INTRODUCTION

1. Population ageing, along with concerns about the economic and social exclusion of persons of all ages with disabilities, lie behind the strong interest in monitoring trends on the prevalence of chronic health conditions and disability rates in many OECD countries, and to compare policy and programme experiences internationally. Policy-makers and researchers in OECD countries use indicators of health status and disability for a number of different purposes, including:

- i. For monitoring social and health progress over time, within and across countries, including through the construction of summary measures of population health;
- ii. For health and social policy planning, such as making costs projections for various care needs;
- iii. For evaluating the outcomes of various health policies and health care interventions.

2. In particular, there has been in recent years a strong interest across OECD countries in monitoring past trends and making projections of disability rates, particularly among the elderly population, because of their potential impacts on social and health care programmes and related costs (Jacobzone *et.al.*, 2000).

3. This report deals with the measurement of health status and disability in national surveys. It does not address the more specific issue of how disability is defined as an eligibility criteria for various disability-related programmes and benefits².

4. In recent decades, a growing number of countries have expanded their effort to use and improve their national health interview surveys, as an important source of information on the overall level and distribution of morbidity and disability in the population. Most OECD countries currently have questions in their national health surveys on self-rated general health and on the prevalence of chronic conditions and disability. Over the last few years, *OECD Health Data* has started to collect statistics on self-rated general health (specifically, the proportion of people reporting their health to be “good/very good/excellent” combined). Data on a few specific chronic conditions (e.g., congenital anomalies) have also been collected, but these have been gathered through national or international disease-specific registers. Data on disability have not yet been included in *OECD Health Data* because of a lack of international comparability of the measurement instruments in national surveys. The expert meeting on ageing and disability held by the OECD on 9-10 December 1999 highlighted the fact that policy discussions on disability-related issues are seriously hampered by the lack of internationally comparable data.

5. The purpose of this inventory is to examine more closely the comparability of survey instruments used to measure health and disability in OECD countries. It builds on an inventory of “Survey data on disability” in EU countries prepared for Eurostat (Gudex *et.al.*, 1999), which compared disability-related items in 15 surveys from 11 European Union countries. This Eurostat inventory showed that although there are significant variations in instruments used to measure health and disability, it is possible to compare data on at least a few “core” items related to chronic illness and general limitations in daily activities across a significant number of countries.

2. A separate OECD project on “Policies to support and integrate the working-age disabled” will collect information on the definitions of eligibility criteria for various disability-related programmes.

6. This OECD wide inventory covers 30 recent or planned national surveys from 23 OECD countries. It is designed to help fill gaps in *OECD Health Data*, by identifying items for which there are now reasonably comparable data across a fairly large number of OECD countries. It is also designed to promote informed discussions on the use of common items (or modules) to measure health status and disability in upcoming national surveys, for the purpose of improving international comparisons.

7. The structure of this report is the following. We first review briefly the scope and focus of this inventory. This is followed by a brief description of the 30 national surveys included in this report (along with the European Community Household Panel); all the details about the formulation of questions in each of these surveys are provided in Annex 2. Before examining the comparability of measurement instruments across these surveys, we discuss briefly the definitions of disability suggested at the international level and those effectively used in various countries. We also review quickly the development of disability measurement instruments over time and the move towards the development of more generic health measurement instruments, as well as past and current efforts to promote the international standardization of these instruments. The bulk of the report then examines the comparability of survey instruments for each health and disability item covered in this inventory. The final section summarizes the main findings of this review and draws some conclusions on ways to promote more comparable health and disability data.

SCOPE AND FOCUS OF THIS REVIEW

8. Health-related surveys in OECD countries typically include questions on health conditions and the consequences of health problems on functional and activity limitations (i.e., “disability”). The items covered in this inventory of national surveys are, accordingly, divided in two categories:

- i. health conditions; and
- ii. activity limitations.

9. As indicated in Box 1, these domains cover a range of health conditions and activity limitations. The items on health conditions include both the assessment of physical and mental conditions, with a focus on questions related to physical health problems. This inventory does not review the comparability of the question on self-rated general health, as this indicator is already included in *OECD Health Data* with the proper documentation on sources and methods. The items on activity limitations (or “disability”) cover a wide spectrum of activities, with a focus on items related to activities of daily living (ADLs) and general (unspecified) activity limitations due to health problems³.

10. The focus of this inventory is on chronic conditions and long-term disability (as opposed to acute conditions or short-term disability), although it is often proving difficult to make a clear distinction between the two. This is particularly the case for measurements of mental health, which tend to focus more on measuring acute (or current) conditions than chronic problems.

Box 1. Coverage of health and disability dimensions in this report

1. Health Conditions

- 1.1 General questions asking about the prevalence of chronic conditions but which do not list particular conditions (General prevalence)
- 1.2 Items which list certain conditions (Specific conditions/illnesses)
- 1.3 Mental health, which includes both emotional/psychological well-being (items relating to anxiety, depression, etc) and cognitive ability (items relating to memory, thinking, etc.)
- 1.4 Pain

2. Activity Limitations

This is a broad domain, which moves from the more specific ADL items to more complex activities and general items about activity limitations:

- 2.1 Self-care: mostly relates to traditional ADL items, i.e. washing, dressing, feeding
- 2.2 Mobility: includes the ability to walk, to climb stairs, and to get in and out of bed (transfer)
- 2.3 Communication/senses: relates to the ability to see, hear and talk

3. The term “item” is used here to refer to the individual questions and response formulations used in surveys.

2.4 IADL (Instrumental activities of daily living): items that specify activities necessary to live independently in the community such as shopping, specific housework chores, preparing meals and handling personal finances

2.5 Work/social activities: items that focus on limitations in work or social activities; 'work' here includes housework but only where it is not further detailed, in which case it would come under IADL

2.6 General items on limitations in usual activities: where the activities are not further specified beyond a description such as limitation in 'daily (usual) activities', with or without a list of broad examples (e.g. work, household chores, social and leisure activities)

IDENTIFICATION OF SURVEYS

11. Building on the surveys covered in the 1999 Eurostat inventory, a search was extended to include other recent or planned health surveys conducted in OECD countries, mainly through contacts with OECD Health Data national correspondents. A total of 30 national surveys from 23 countries are included in this inventory (see Table 1) along with a review of the European Community Household Panel. It is important to keep in mind that this list of surveys is far from being exhaustive. In many of the 23 countries included in this inventory as well as in countries not covered in this report, other surveys have been conducted (or are planned) which may provide information on health and disability-related issues included in this report. As a result, the national coverage of various health and disability measures noted in this paper should be viewed as a minimum, since other survey data sources may also be available.

12. Table 1 shows that in about half of the OECD countries, national authorities have decided to conduct both general health surveys (covering the entire population) as well as disability-specific surveys (with the latter often focussing on the elderly population). In other countries, only a general health survey is conducted, often including (short or long) modules on disability-related items. Most of the surveys are conducted regularly, with the time interval between surveys normally varying from 1 to 5 years. Sample size and sampling frames vary across the surveys, in particular on whether or not children and people living in institutions are included. The main characteristics of each survey are reviewed briefly in Annex 2⁴.

13. In addition to these national surveys, reference to the European Community Household Panel – an annual EU-wide survey conducted since 1994 which contains a short module on health and disability – is made where appropriate. The European Community Household Panel provides a source of comparable data for up to 14 EU countries on some of the general items included in this inventory.

14. Before reviewing in detail the comparability of instruments used in national surveys to measure the various health and disability items, it is useful to discuss briefly first issues related

4 . This report does not focus on comparability issues arising from variations in sample designs and other survey methodologies, with the exception of issues related to population coverage by age group.

to the definition and measurement of disability and the evolution of disability measurement over time.

Table 1: Summary of national surveys included in the inventory

Country & Survey no.	Survey title	Contact	Population	Data collection	Internet address
Australia AL01	National health and nutrition survey	Australian Bureau of Statistics	All; includes institutions	1995; every 5-6 years	
Australia AL02	Survey of disability, ageing and carers	Australian Bureau of Statistics	People with disability; ≥60 yrs; carers; includes institutions	1998; every 5 years	
Austria AT01 *	Microcensus survey on physical disabilities	Austrian Central Statistical Office	All; includes institutions	1995; approx. every 10 yrs	
Belgium BL01 *	Health interview survey	Scientific Institute of Public Health	All	1997, 2001	www.iph.fgov.be/epidemie/pien/index4.htm
Canada CA01	National population health survey	Statistics Canada	All; includes institutions	1998/9; every 2 yrs (longitud.)	www.statcan.ca/health_surveys
Canada CA02	Health and activity limitation survey	Statistics Canada	All; includes institutions	1991; every 5-10 yrs (planned for 2001)	
Czech Republic CZ01	Health interview survey	Institute of Health Information and Statistics	≥15 yrs	1999; every 3 yrs	
Denmark DK01 *	Danish health and morbidity survey	Danish Institute for Clinical Epidemiology	≥16 yrs	1987, 1994, 2000	www.dike.dk
Finland FIN01 *	Health behaviour among Finnish adult population	National Public Health Institute	15-64 yrs	1999, 2000 (annual)	
France FR01 *	National health interview survey	INSEE	All; excludes institutions	1970, 1980, 1991, 2000-2	
France	National disability	INSEE	All; includes	1998-	

FR02 *	interview		institutions	2001	
Germany DEU01	National health examination and interview survey	Robert Koch Institute	18-79 yrs; excludes institutions	1998	
Iceland ICE01	Health and living conditions in Iceland	Statistics Iceland & University of Iceland	18-75 yrs; excludes institutions	1998; irregular	
Ireland IRE01	National health and lifestyle survey	Department of Health and Children	≥18 yrs; excludes institutions	1998; every 3 yrs	
Italy IT01	Survey of health conditions and use of health services	National Institute of Statistics	All; excludes institutions	1999-2000; every 4 yrs	

Table 1: Summary of national surveys included in the inventory (continued)

Country & Survey no.	Survey title	Contact	Population	Data collection	Internet address
Japan JP01	Household survey of physical and mental health	Department of Health Statistics	≥12 yrs	2000	
Korea KR01	National health interview survey	Korean Institute for Health and Social Affairs		1995; every 3 yrs	
Netherlands NL01 *	Health interview survey	Statistics Netherlands	All; excludes Institutions	Annual	
New Zealand NZ01	New Zealand health survey	Statistics New Zealand	All; excludes institutions	1996-7	www.moh.govt.nz
New Zealand NZ02	New Zealand disability survey	Statistics New Zealand	All; excludes institutions	1996	
Norway NR01	Health interview survey	Statistics Norway	All; excludes institutions	1995; every 10 yrs	
Portugal PR01 *	Health interview survey	Ministry of Health	All; excludes institutions	1998/9; every 3 yrs	
Spain SP01 *	Survey on disability, impairment and health status	National Institute of Statistics	All	1999	
Sweden SWE01 *	Survey of living conditions	Statistics Sweden	16-84 yrs	Annual	
Switzerland SZ01	Swiss Health Survey	Office fédéral de la statistique	≥15 yrs; excludes institutions	1997; every 5 yrs	
England UK01 *	Health Survey For England	Department of Health	≥2 yrs; England only	1995, 1996 (annual)	www.mimas.ac.uk/surveys/hse/hse.info
United Kingdom UK02 *	Labour Force Survey	Office for National Statistics	Men 16-64 yrs; women 16-59 yrs; GB + N. Ireland	1998/99 (annual)	
Great Britain	Disability in Great Britain survey	Office for National Statistics	All; excludes institutions	1996	

UK03					
United States US01	National health interview survey	National Center for Health Statistics	All; excludes institutions	1998; annual	www.cdc.gov/nchs/nhis.htm
United States US02	National long-term care survey	Duke University	≥65 yrs; includes institutions	1999; every 5 yrs (longitud.)	www.cds.duke.edu/nltes_intro

* indicates surveys covered in the 1999 Eurostat report

Note: In addition to these national surveys, relevant information from the European Community Household Panel – a survey covering 14 European countries which contains a brief module on health and disability-related matters – is included in this inventory.

DEFINITION OF DISABILITY

15. As it stands, the most commonly used general definition of disability is the one proposed by the WHO in the 1980 International Classification of Impairment, Disability and Handicap (ICIDH-1):

“A disability is any restriction or lack of ability (resulting from an impairment) to perform an activity in the manner or within the range considered normal for a human being” (WHO, 1980, p. 143)

16. In ICIDH-1, disability therefore was defined as the consequences of diseases and other impairments on activity limitations. The ICIDH is currently under revision. One of the objectives of the current revisions is to put more emphasis on the use of more positive language, such as “activity” and “participation” rather than “disability” and “handicap”. The proposed revisions to the ICIDH also include a broadening of the definition of “disability”, to include not only problems related to “activities” but also those related to the “body systems/structure” (previously called “impairment”) and those associated with “participation” (previously called “handicap”) (WHO, 1999, p.12). Although the definition of disability in ICIDH-2 might therefore be broadened, it is clear that a core part of the definition will continue to be centred around activity limitations.

17. To define “disability” as an “activity limitation” is an important conceptual first step; however it raises issues about what “activities” and what levels of “limitation” should be used in implementing the definition.

18. A review of operational definitions of “disability” used by national authorities and researchers in various OECD countries shows that, generally, they have been consistent with the 1980 ICIDH definition of disability as being an activity limitation, although there are important variations in the specific definitions used (Table 2). At least two broad approaches have been used in operationalizing the definition of disability (Mathers, 1997):

- the first approach – the most commonly used -- has been to define disability as a limitation in basic activities of daily living (also referred more narrowly as “functional limitations” or “functional disability”). This definition of disability has tended to focus primarily on problems related to elderly people, although it has been used as well for population-wide estimates of disability;
- the second approach has used a broader definition of disability, based on people reporting limitations in usual activities generally (unspecified). This definition has been used mainly to provide estimates of the prevalence of disability for the entire adult population.

19. In some countries (like Australia and Canada), these two different approaches have, in the past, been combined to provide a single estimate of the adult population with a disability; in other countries (like the United Kingdom) researchers have provided estimates of the disabled population separately using each of these two different approaches⁵.

Table 2: Examples of definitions of 'disability' in OECD countries

Australia	Disability defined as having one or more of the following conditions lasting more than 6 months: any physical functional limitations (loss of sight, hearing, speech, dexterity), any mental, cognitive or emotional problems, any other chronic conditions or impairments that restrict everyday activities (Australian Bureau of Statistics, 1998)
Canada	Disability defined as having one or more of the following conditions lasting more than 6 months: any ADL limitations, any mental or cognitive problems, any limitations in daily activities (at home, at school, at work, etc) due to physical or mental problems (Statistics Canada, 1991)
Finland	Disability defined as limitations in work or other daily activities (Valkonen <i>et.al.</i> , 1994)
France	Disability defined as a limitation in usual activities generally (work, school, etc). In addition, 'severe' disability defined as people confined to their homes (including all people living in institutions) (Robine et Mormiche, 1993)
Germany	Disability based on reported days that people are unable to perform their usual activities due to ill-health or injury (Bruckner, 1997)
Korea	Disability defined as one or more ADL limitations (Lee, 1997)
Netherlands	Disability defined as people having one or more functional or ADL limitations, based on the OECD long-term disability list (data also broken down by main areas of functional limitations – mobility, seeing, hearing) (RIVM, 1998)
New Zealand	Disability defined as having one or more functional limitations (Ministry of Health, 1998)
UK	Disability defined as having a general limitation in usual activities due to health problems (Bebbington and Darton, 1996)

5 . In addition to these two broad approaches, a number of other approaches have been used to define and measure disability. A 1998 REVES paper provides a list of the various types of disability and health expectancy measures that have been estimated in various OECD countries (REVES, 1998).

	Disability defined as having a problem or a limitation related to 13 physical and mental health conditions and ADLs (Bebbington, 1992)
USA	Disability defined as limitations in both major and secondary activities ('major' activities include able to go to work, go to school, do housework, while 'secondary' activities are activities such as going to church and recreational activities) (Crimmins <i>et.al.</i> , 1997)

20. Different approaches to defining and measuring disability obviously result in different prevalence rate estimates. For instance, in the case of Great Britain, the more encompassing definition, based on general “limiting long-standing illness”, provided an estimate of 21% of the adult population (16 years and over) being “disabled” in 1985 (using the 1985 General Household Survey). On the other hand, the alternative definition of disability, based on a more specific measurement of 13 ADL limitation and health condition items, resulted in an estimate of 14% of the adult population 16 years and over with any disability in Great Britain in 1985, based on the 1985 OPCS surveys of disability (Bebbington, 1992). The latter estimate was even lower for more severe levels of disability calculated on the basis of the various scales attached to each ADL.

21. Even when only ADL-based measures are used to estimate disability prevalence, significant differences have been reported depending on which specific ADL instrument is used (Wiener and al., 1990). These different results highlight the importance of specifying how disability is defined and measured.

A Brief History of Disability Measurement

22. Because of its importance for social and health care planning, there has been a proliferation of instruments to measure disability in OECD countries over the last 40 years. The early development of disability indicators focussed on measuring basic functional ability, with a focus on limitations in activities of daily living (ADL); one of the prime examples was the ADL scale proposed by Katz *et.al.* (1963). These ADL instruments were originally aimed at assessing the severe disablement commonly found among institutionalised patients and the elderly population (measuring, for example, independence in bathing, dressing, moving around the house and eating).

23. Later, Lawton & Brody (1969) introduced the notion of Instrumental Activities of Daily Living (IADL) to cover a broader range of activities, including activities required to live independently (such as the ability to manage personal finances, do housework and shopping). These IADLs tend to be more complex and demanding than basic ADLs. They have been used, among other things, to measure less severe levels of disability. One of the disadvantages of IADLs for the purpose of international comparisons is that the performance of these activities tends to be affected by gender-specific roles in various cultures. By comparison, basic ADL items are more likely to reflect pure functional ability.

24. Over the years, a number of other instruments have been developed to measure similar types of ADLs and IADLs. McDowell and Newell (1996) reviewed a sample of 16 such ADL and IADL scales.

25. Somewhat more recently, there has been a strong development of more generic health measurement instruments which contain disability-related components along with items on physical and psychological health. These instruments are referred either as “generic health status measures” or as “measures of health-related quality of life”. Prime examples of such generic instruments include the SF-36 questionnaire (and its abbreviated versions, such as SF-12), the EuroQol-5D instrument and the McMaster’s Health Utilities Index (HUI-3). These instruments are used to provide health profiles which describe health status in a set of scores (e.g., for physical health and mental health in the case of SF-36 or SF-12) or health indices which can summarize health status in a single number (in the case of EuroQol-5D and HUI-3). These generic health measurement instruments are increasingly being used in national surveys to measure health and activity limitations, either as a complement or as a substitute to disability-specific instruments.

Efforts to Promote International Standardization

26. One of the first attempts to standardize disability measures across countries was undertaken by the OECD in the late 1970s, as part of a broad programme to develop social indicators. This effort led to the so-called “OECD long-term disability list” (McWhinnie, 1982). The instrument proposed by the OECD at that time focussed on the measurement of long-term disability, based on an ADL scale covering mobility, self-care and communication items (Table 3). The main objective of the OECD long-term disability list was to allow better international comparisons of disability across countries, through its implementation in national surveys. The instrument was implemented, for a period of time at least, in 8-10 OECD countries. Some countries, like the Netherlands, continue to use this set of questions in their health surveys, while other countries continue to use only selected items. Various assessments of the instrument showed that its focus on basic ADL limitations made it more valid as a measure of severe levels of disability, and most relevant to assess activity limitations for people over 65 (McDowell and Newell, 1996).

27. A more recent attempt to standardize disability measures was undertaken by WHO-Europe, in collaboration with Statistics Netherlands, as part of an on-going effort to standardize methods and instruments in health interview surveys (WHO-Europe, 1996). WHO-Europe (along with Statistics Netherlands) made recommendations for standardized instruments to measure both short-term disability and long-term disability. As for the OECD list, the WHO-Europe “long-term disability list” is designed to measure disability through ADL-type limitations, covering the same key basic activities related to mobility, self-care and communication (Table 3). This recommended instrument has now been incorporated in some recent surveys in European countries (e.g., Belgium, Czech Republic, Portugal as well as in the 1995 Health Survey for England with some adjustments). Although the proposed instrument does not seem to have been assessed yet to the same extent as its OECD predecessor, it is also likely to prove to be most relevant in measuring relatively severe levels of disability more frequently found in the population 65 and over. The current phase of the WHO-Europe EUROHIS project is attempting to complete the set of recommended instruments for national health surveys, by proposing common instruments related to the measurement of chronic physical conditions, mental disability as well as overall measures of quality of life.

28. The Euro-REVES Network (Réseau espérance de vie en santé), with support from the European Commission, also recently released a set of recommendations for survey instruments

in five areas (i.e., physical and sensory functional limitations, ADL limitations, limitations in usual activities, self-rated general health and mental health), as part of an effort to develop a consistent set of health expectancies across European countries (Euro-REVES, 2000). The Euro-REVES recommendations build on previous recommended instruments by WHO-Europe. In the area of self-rated general health and for some aspects of mental health (emotional well-being), the Euro-REVES recommended instruments are the same as those recommended previously by WHO-Europe. In the area of disability, Euro-REVES has proposed to break down the WHO-Europe “long-term disability list” into two components: some of the questions would be used to measure “physical and sensory functional limitations”, while the other questions would be used to measure “ADL restrictions” (see footnote 5 to Table 3). This distinction has been justified on the grounds that “functional limitation and activity restriction refer to different levels of disability leading to different types of consequences on daily life, and towards which different public health actions can be addressed”. Euro-REVES recognizes however that “the distinction [between functional limitation and activity restriction] is not straightforward, especially because existing measurement instruments are most of the time combining these two levels and are relying on apparently similar questions”⁶.

29. The Euro-REVES network also proposes to include four severity levels in functional and ADL limitations (as opposed to three for most of the WHO-Europe questions), by making a clear distinction in the ability to function or carry out basic activities of daily living with or without special aids (such as glasses or hearing aids).

30. In addition, the Euro-REVES network proposes a global (single) question on disability in usual activities, which can be administered to a population of all ages. The proposed instrument is a variation on the question used in many surveys to measure the general prevalence of disability or “limiting long-standing illness” (see section 2.6 below).

6 . In the present report, the categories “mobility” and “communication” under activity limitations are roughly equivalent to what the Euro-REVES network has called “physical and sensory functional limitations” while the category “self-care” is roughly equivalent to their more narrow definition of “ADL restrictions”.

Table 3: Reference instruments for measuring functional and ADL disability*

Katz (1963)¹	OECD (1982)³ (*Long term disability 10 items minimum core set)	WHO-Euro (1996)⁴ and Euro-REVES (2000)⁵
Dressing	Dress and undress*	Dress
Transfer from bed and chair	Get in and out of bed*	Transfer from bed
		Transfer from chair
Bathing		Wash hands and face
Toileting		Get to and use the toilet
Feeding	Can you cut your own food*	Feed, including cutting up food
Continence		Continence
Nagi (1976)²		
Standing for long periods		
Lifting or carrying weights	Carry an object of 5 kilos for 10 meters*	
Going up and down stairs	Walk up and down one flight of stairs without resting*	Stairs (optional)
Walking	Walk 400 meters without resting*	Locomotion
Stooping, bending or kneeling	Bend down (when standing) and pick up shoe	Retrieval (optional)
Using hands and fingers		
Reaching with either/both arms		
	Move between rooms*	Mobility
	Speaking*	Speaking (optional)
	Hear normal conversation with another* Hear normal conversation with 3 or 4 other persons	Hearing
	Read ordinary newsprint* See the face of someone from 4 metres	Seeing
	Run 100 meters	
	Cut your toenails	
	Bite and chew on hard foods	

* Adapted from Robine and Jagger (1999).

¹”Do you perform” without supervision, direction or personal assistance.

²”Do you have any difficulty.....” ; coded as no difficulty, some difficulty, great difficulty.

³”Can you” ; coded as yes without difficulty, with minor difficulty, major difficulty, unable to do.

⁴”Can you” ; coded as without difficulty, with some difficulty, only with someone to help.

⁵ The last seven items relate to the Euro-REVES recommendations to measure “physical and sensory functional limitations” while the first five items relate to their recommendations for measuring “ADL restrictions” (with some adjustments).

31. In addition to these efforts to promote the use of standard disability instruments, there have also been a number of academic efforts to promote the use of common generic health measurement instruments internationally. These include the EuroQol-5D instrument in Europe (which has been incorporated in the 2000 Health and Morbidity Survey in Denmark, the 2000 Health Survey in Finland, the 2000 Health Survey in Hungary, the 1998 national Health and Lifestyle Survey in Ireland and the 1996 Health Survey for England), the SF-36 instrument and its abbreviated versions initially developed in the U.S. (which have now been included in surveys in Australia, Denmark, Germany, Italy, New Zealand and the United Kingdom) and the Health Utilities Index Mark 3 (incorporated in Canadian health surveys as well as in other non-OECD countries). Table 4 provides an overview of these three leading generic health measurement instruments, while Annex 1 provides the detailed formulation of questions. The three instruments include items on some basic functional ability (e.g., mobility) and other aspects of physical and emotional well-being. They vary however in their coverage of health and disability dimensions, with the HUI-3 putting more emphasis on measuring functional health, while the SF-36 and EuroQol-5D place a greater emphasis on measuring limitations in daily activities. There is also a notable difference in the number of items used in each instrument to measure different activities, with EuroQol being the most parsimonious in using only one item by dimension covered, while the SF-36 and the HUI-3 use several items to measure mobility and other functional/activity limitations.

Table 4: Overview of three generic health measurement instruments

	EuroQol-5D	SF-36	HUI-3
Number of questions	5	36	31
Skip Patterns	No	No	Yes
Reference Period	Today	Last four weeks	Usual
Number of Dimensions	5	8	8
Dimensions	Mobility (1)	Mobility (9) & Self-care (1)	Mobility (7)
	Self-care (1)		Dexterity (4)
			Vision (5)
			Hearing (5)
			Speech (4)
	Anxiety/depression	Emotional well-being	Emotional well-being

	(1)	(5)	(1)
			Cognition (2)
	Pain/Discomfort (1)	Pain (2)	Pain (3)
		Vitality/energy/fatigue (4)	
	Usual activities (1)	Role limitations due to physical problems (4)	
		Role limitations due to emotional problems (3)	
		Social functioning due to physical or emotional problems (2)	
		General health perceptions (6)	

Note: Numbers in brackets indicate the number of items related to each dimension.

ANALYSIS OF COMPARABILITY BETWEEN SURVEYS

32. Let's turn now to reviewing the comparability of the health and disability-related items across the 30 national survey included in this inventory (and the European Community Household Panel).

Overview of Coverage

33. The specific items used in each survey are listed in annex 2. Table 5 summarises the broad health status and disability-related components covered in each survey. The specific items used in each survey are listed in Annex 2. Traditional ADL components are included in most surveys, with all but 3 surveys including self-care and mobility items, and 24 surveys including the ability to communicate. Just under half of the surveys (14) include IADL items, while about two-thirds of the surveys include items on work/social limitations or a general item on limitations in usual activities (the European Panel also includes a general question on "limiting long-standing illness").

34. Items asking about the presence of specific health conditions (23 surveys) are more common than a broad item on the general prevalence of any chronic condition (11 surveys). Here as well the European Panel includes a general question on the prevalence of chronic conditions. Three-quarters of the surveys (23) include some items on mental health, and half include questions related to pain.

35. We turn now to a more detailed review of the comparability of survey items in each domain.

Table 5: Health and disability-related domains included in selected national health surveys

	Health conditions				Activity limitations					
	General prevalence	Specific conditions	Mental	Pain	Self-care	Mobility	Communication	IADL	Work/Social	General item
AL01	No	√	√	√	√	√	No	No	√	√
AL02 ¹	No	√ ²	√	No	√	√	√	No	√	√
AT01	√	√	No	No	√	√	√	√	√	No
BL01	√	√	√	√	√	√	√	No	√	√
CA01	√	√	√	√	√	√	√	√	√	√
CA02	No	No	√	No	√	√	√	√	√	√
CZ01	√	√	√	No	√	√	√	No	No	No
DK01	√	√	√	√	√	√	√	No	√	√
FIN01	No	√	√	√	No	No	No	No	√	No
FR01	No	√	No	No	√	√	√	√	√	No
FR02	No	No	√	No	√	√	√	√	√	√
DEU01	No	√	√	√	√	√	No	No	√	√
ICE01	No	√	No	No	√	√	No	√	√	No
IRE01	No	√	√	√	√	√	√	No	No	√
IT01	No	√	√	No	√	√	√	√	√	√
JP01	No	No	√	√	No	No	No	No	No	√ ³
KR01	No	√ ⁴	No	No	√ ⁵	√	√	No	√	No
NL01	√	√	√	√	√ ⁶	√	√	No	No	√
NZ01	√	√	√	√	√	√	No	No	√	√
NZ02	No	No	No	No	√	√	√	√	√	No
NR01	√	√	No	No	√	√	√	√	√	√
PR01	No	√	No	√	√	√	√	No	No	No
SP01	No	√	√	No	√	√	√	√	√	√
SWE01	√	No	√	√	√	√	√	√	No	√ ⁶
SZ01	√	√	√	√	√	√	√	No	No	√
UK01	√	√	√	√	√	√	√	No	No	√
UK02	√	√	√	No	No	No	√	No	√	√
UK03	No	No	√	√	√	√	√	√	No	√
US01	No	√	√	√	√	√	√	√	√	√
US02 ⁷	No	√	√	No	√	√	√	√	No	No

1. Only asked to people with disability or ≥60 years
2. Only asked if a disability has already been identified
3. Specifies due to strain, stress or pressure
4. Respondents are asked if they have 'Diseases of respiratory system', etc, but conditions are not specified
5. It is assumed here that 'Activity of daily living' includes self-care items
6. Only for persons aged ≥55 years (NL01) or old-age pensioners (SWE01)
7. Only for persons aged ≥65 years

1. Health conditions

1.1 General prevalence of chronic conditions

36. Eleven of the 30 surveys (covering 10 different countries) include a general item similar to 'Do you have any long-standing/chronic illness or health problem? Yes/No'. In addition, the ECHP also includes a similar general question on long-standing or chronic conditions, providing a source of data for up to 14 European countries (at least 5 of these countries also use a similar instrument in their national surveys – Austria, Belgium, Denmark, the Netherlands and the United Kingdom – thereby providing at least two sources of survey data).

37. There are at least two main variations in the wording of the question used in various surveys for this item. First, “long-standing/chronic illness” is sometime defined as a condition lasting for 6 months or more, or lasting 1 year or more. Sometime it is not defined at all in terms of duration. It is not known to what extent these variations in the reference period affect the comparability of data across surveys.

38. A second source of variation relates to the explicit reference to mental health problems in the question. For instance, the European Panel explicitly asks whether respondents suffer from any “chronic physical or mental health problem”, while most other national surveys refer generally to any “chronic health problem”, without specifying whether these are physical or mental. The explicit reference to possible mental health problems in surveys such as the ECHP might be expected to increase (at least slightly) the prevalence of chronic conditions compared with surveys that do not include an explicit mention of such conditions. No one however appears to have done a direct comparison of the impact of the explicit mention of mental health problems with similar questions which do not.

39. Overall, a recent source of comparable survey data on the general prevalence of chronic conditions would seem to be available for at least 20 OECD countries (the 14 EU countries which have administered the European Community Household Panel and Canada, Czech Republic, New Zealand, Norway, Sweden and Switzerland).

1.2 Specific conditions/illnesses

40. While it is useful to have a general indication of the prevalence of chronic conditions in the population, health researchers and policy-makers are also interested to know more specifically the main chronic conditions that people suffer from. In countries like Canada, recent data shows that the reported prevalence of certain chronic conditions, such as asthma and diabetes, have tended to increase over the last few decades (Statistics Canada, 2000).

41. Table 6 shows the 9 chronic conditions most commonly asked in the 23 surveys which contain a check list of specific conditions⁷. Nearly all of these surveys enquire about: high blood pressure, heart disease, asthma or chronic bronchitis (either separately or together in one item), and diabetes, thereby providing a source of survey data on the prevalence of these conditions.

Table 6: List of specific chronic conditions commonly asked in national health surveys

7. The Korean Health Interview Survey has not been included as the conditions do not seem to be sufficiently specified.

	High blood pressure	Heart disease ¹	Asthma/chronic bronchitis	Diabetes	Stroke	Arthritis ²	Allergy ³	Epilepsy	Stomach / duodenal ulcer
AL01	√	No	√	√ ⁴	No	√ ⁴	√	No	√
AL02 ⁵	√	√	√	√	√	√	No	√	√
AT01	√	√	√	√	√	√	√	No	No
BL01	√	√	√	√	√	√	√	√	√
CA01	√ ⁴	√ ⁴	√ ⁴	√ ⁴	√ ⁴	√ ⁴	√ ⁴	√ ⁴	√ ⁴
CZ01	√	√	√	√	√	√	No	√	√
DK01	√	√	√	√	√	No	√	√	√
FIN01	√ ⁴	√ ⁴	√ ⁴	√ ⁴	No	√ ⁴	No	No	√ ⁴
FR01	√ ⁴	√ ⁴	√ ⁴	√ ⁴	√ ⁴	√ ⁴	No	No	√ ⁴
DEU01	√	√	√	√	√	√	√	√	√
ICE01	√	√	No	√	√	No	No	No	No
IRE01	√ ⁴	√ ⁴	No	√ ⁴	√ ⁴	No	No	No	No
IT01	√	√	√	√	√	√	√	√	√
NL01	√	√	√	√	√	√	No	√	√
NZ01	√	No	√	√ ⁴	No	No	No	No	No
NR01	√ ⁴	√ ⁴	√ ⁴	√ ⁴	√ ⁴	√ ⁴	√ ⁴	√ ⁴	√ ⁴
PR01	√ ⁴	No	√ ⁴	√ ⁴	No	No	√ ⁴	√ ⁴	No
SP01	√ ⁴	√ ⁴	√ ⁴	√ ⁴	√ ⁴	√ ⁴	√ ⁴	√ ⁴	√ ⁴
SZ01	No	No	No	No	No	No	√	No	No
UK01	√	√	√	√	√	√	√	√	√
UK02	√	√	√	√	No	√	√	√	No
US01	√ ⁴	√ ⁴	√ ⁴	√ ⁴	√ ⁴	No	√ ⁴	No	√ ⁴
US02 ⁶	√	√	√	√	√	√	No	√	No

1. Items asking about angina, heart attack (myocardial infarction), coronary heart disease, any/other heart disease
2. Items specifying rheumatoid arthritis or osteoarthritis (arthrosis)
3. Includes items asking about allergy (not specified further), hay fever, food allergies, skin allergy
4. The question specifies that the condition should have been diagnosed by a health professional
5. Only to people with disability or ≥60 years
6. Only for persons aged ≥65 years

42. There are, however, three important variations in the formulation of items across surveys:

- i) Surveys vary as to whether or not they specify that the conditions should have been diagnosed by a health professional. It can be expected that surveys that include the additional criterion that the condition should have been verified by a health professional (such as the National Population Health Survey in Canada and the National Health

Interview Survey in the United States) would report lower prevalence rates than those which do not impose that condition (since not all people suffering from a chronic condition may seek verification and help from a health professional).

- ii) Various conditions are sometimes combined together into one group, preventing a more specific decomposition. For instance, some surveys include emphysema with asthma and bronchitis in one item, while others ask about these separately. Another example is that while most surveys ask separately about angina and heart attack, some combine angina, heart attack, arrhythmia and/or other heart disease as part of one response category only. For the purpose of cross-survey comparisons, some grouping of chronic conditions might therefore be required.
- iii) As is the case for the general question on chronic conditions, the reference period varies between surveys, with some surveys referring to conditions the respondent currently has (AL01, AL02, CA01, CZ01, FR01, PR01, SP01, SZ01, US02), others referring to conditions the respondent has ever had (AL01, DEU01, IRE01, SP01, US01), while others combine these and refer to conditions the respondent currently has or has ever had (DK01, IT01).

43. Despite these variations, the chronic nature of the conditions listed here (excluding possibly stroke and ulcer) and the likelihood of health service contact for treatment of these conditions make it more likely that the data on the prevalence of these various chronic conditions would be comparable across surveys. The reliability of these survey data might also be tested by comparing results with other sources of morbidity data such as administrative records or disease-specific registers.

1.3 Mental health

44. Modules on mental health are increasingly included in national health surveys, with three-quarters of the surveys covered in this inventory including items related to some aspects of mental health and their impact on activity. This reveals a growing interest in measuring mental well-being in population surveys, as an important complement to the measurement of physical health conditions. However, reflecting the traditional emphasis on physical health conditions, international databases such as *OECD Health Data* now only include a few proxy indicators of mental health conditions (e.g., suicide rates).

45. Defining precisely the boundaries of measures of mental health has proven to be difficult. As McDowell and Newell note, “there have been many attempts to specify what is being measured – to distinguish, for example between “distress” and “disorder” and between “psychological”, “emotional” and “mental” well-being... [these] attempts have not always been successful” (McDowell and Newell, 1996, p. 178). In this report, the term “mental health” is used to include both psychological/emotional problems (e.g., nervousness, anxiety and depression) and cognitive functioning (e.g., ability to think and remember).

46. In general, modules on mental health in national surveys have not been designed to diagnose specific mental disorders (such as dementia or Alzheimer disease) or psychological/emotional distress (such as depression). Very specific instruments are available in clinical settings to screen for or diagnose specific mental disorders (these include the Mini-Mental State Examination, the Alzheimer’s Disease Assessment Scale, the Dementia Scale, etc.). Rather, the instruments to measure mental health in national surveys are designed to provide a

general indication of the mental health conditions in a population and the likelihood of mental problems.

47. Table 7 shows that more than half of the surveys in this inventory include the measurement of various aspects of psychological/emotional well-being, such as nervousness, depression or the level of vitality/energy, through one instrument or another. The mental health component of the SF-36 questionnaire (5 questions) has been used to measure mental health in a number of national surveys (Australia, Denmark, Germany, Italy, New Zealand and the 1996 Health Survey for England). The General Health Questionnaire (GHQ-12; Goldberg, 1972) – an instrument designed specifically to detect psychological/emotional problems -- has also been used in some national surveys in Europe (e.g., Belgium and England)⁸. In addition, surveys using the EuroQol-5D and HUI-3 also measure psychological/emotional well-being through a single question on anxiety/depression (in the case of EuroQol-5D) and general feelings (in the case of HUI-3).

Table 7: Coverage of mental health and pain items in national health surveys

	Mental: emotional			Mental: cognitive		Pain	
	Nervous	Depressed	Energy	Think	Memory	General	Specific
AL01 ¹	√	√	√			√	
AL02 ¹	√	√	√				
BL01 ²	√	√	√	√		√	√
CA01	√	√	√	√	√	√	
CA02	√	√		√	√		
CZ01	√	√	√				
DK01 ¹	√	√	√			√	
FIN01	√	√					√
FR02					√		
DEU1 ¹	√	√	√			√	√
IRE01	√	√				√	
IT01 ¹	√	√	√				
JP01	√	√	√				√
NL01	√		√				√
NZ01 ¹	√	√	√			√	
PR01							√
SP01				√	√		
SWE01	√						√
SZ01			√				√
UK01 ^{1,2}	√	√	√			√	
UK02	√	√					
UK03	√	√		√	√	√	
US01	√	√			√		√

8. The GHQ-12 has also recently been recommended by the Euro-REVES network as a standard instrument to measure mental health in Europe (Euro-REVES, 2000).

US02			√		√		
------	--	--	---	--	---	--	--

1. Indicates the use of the SF-36 (or SF-12) mental health instrument
2. Indicates the use of GHQ-12

48. Differences in item formulations and in scoring procedures limit the comparability of survey responses on mental health (psychological/emotional problems) only to those surveys using the same instrument. Based on the present review, these would include 7 surveys (covering 6 countries) using the SF-36 instrument, a few surveys using the GHQ-12 and other surveys using the EuroQol-5D and HUI-3.

49. Fewer surveys covered in this inventory (only 8) include items related to cognitive functioning, such as the ability to think, to concentrate, to make decisions or to remember. The measurement instruments for these cognitive skills also tend to vary from one survey to another. In some cases, these items are part of a broader, more generic health measurement instrument (such as HUI-3). The current use and degree of harmonization of measures of cognitive functioning in national surveys are lagging behind the measurement of emotional well-being.

1.4 Pain

50. Pain is another important aspect of health that is increasingly being measured in national surveys (Table 7, last columns). About one-third of the surveys covered in this inventory try to measure pain (or discomfort), mainly through the use of a generic health measurement instrument (such as EuroQol-5D, SF-36 or HUI-3). A few surveys might be comparable on that basis, possibly as part of a broader range of health and disability conditions measured through these generic instruments. In other cases, questions are asked about pain in specific parts of the body, such as back pain. Other question formulations also attempt to build a link between pain and limitations in work or other activities (see section 2.5 below).

2. Activity limitations

51. National health surveys typically include a number of items related to activity limitations due to health problems as a way to measure the prevalence of various types of disability in the population. Modules on activity limitations range from a detailed list of ADL items, to fewer questions on IADLs and limitations in work or other social activities, to the use of a single question on general activity limitations. Not surprisingly, disability-specific surveys tend to include a greater number of specific activity limitations than more general social or health surveys.

ADL items

52. The use of ADL items has become commonly used to measure functional and activity limitations, particularly as they relate to the elderly population. Disability measures based on ADL limitations have been found to be a good predictor of the use of a wide range of health care services, such as admission to a nursing home, use of paid home care and hospital services (Wiener and al., 1990). Recent OECD work on trends in disability among the elderly used the fairly common definition of people reporting at least 1 or more ADL restrictions to estimate trends in “severe disability” (Jacobzone *et.al.*, 2000).

53. Consistent with the ADL items included in the 1982 OECD and the 1996 WHO-Europe long-term disability lists, three main types of ADL are typically measured in national surveys: self-care activities (typically covering the ability to wash, dress and feed oneself); mobility (which includes the ability to walk, climb stairs and get in and out of bed); and communication (the ability to see, hear and speak) (Table 8). As noted before, the Euro-REVES network has recently proposed to make a clearer distinction between items related to “physical and sensory functional limitations” (covering more or less the items referred here as “mobility” and “communication”) and those related to “ADL restrictions” (covering the items referred here as “self-care”).

Table 8: Items related to functional and ADL limitations in national health surveys

	Self-care			Mobility			Communication		
	Wash	Dress	Feed	Walk	Stairs	Bed	See	Hear	Speak
AL01	√ ¹	√ ¹		√	√	√			
AL02 ²	√	√	√	√	√				
AT01	√ ¹	√ ¹				√		√	
BL01 ³	√	√	√	√	√	√	√	√	
CA01	√ ⁴	√ ⁴	√ ⁴				√	√	√
CA02	√	√	√	√	√	√	√	√	√
CZ01	√	√	√	√	√	√	√	√	√
DK01	√ ¹	√ ¹		√	√	√	√	√	√ ⁵
FR01	√	√	√	√	√	√	√		
FR02	√	√	√	√	√	√	√	√	√
DEU01	√ ¹	√ ¹		√	√				
ICE01		√	√		√				
IRE01	√ ¹	√ ¹				√		√	
IT01	√	√	√	√	√	√	√	√	√
NL01 ⁶	√	√	√	√	√	√	√	√	
NZ01	√ ¹	√ ¹		√	√				
NR01	√	√	√	√	√	√	√	√	
PR01	√	√	√	√	√	√	√	√	√
SP01	√	√	√		√			√ ³	√
SWE01 ⁷	√				√	√	√	√	
SZ01		√	√	√		√	√	√	√
UK01	√	√	√	√	√	√	√	√	√
UK02									√
UK03	√	√	√	√	√	√	√	√	
US01	√	√	√	√	√ ⁸	√			
US02 ⁹	√	√	√	√	√	√	√	√	√

1. Washing & dressing combined in one item
2. Self-care and mobility items only asked if people reported being disabled, based on a broad definition of disability
3. Self-care and mobility items (excluding SF-36 items, which are administered to all respondents) only asked to people 60 years and over, or people 15 years and over who answered positively to previous screening questions on disability
4. Washing, dressing and eating combined in one item
5. Interviewer assessment
6. Self-care, Stairs & Bed items only asked to people 55 years and over
7. Washing & Bed items only asked to people reporting long-standing illness, problems with sight or mobility, or aged 75-84 years
8. Specifies "without using any special equipment"

9. Only for persons aged 65 years and over

54. Although the measurement of this core group of activities is very common, various ADL instruments are used to measure these activities (or functional limitations) in different surveys, thereby preventing (or at least complicating) comparisons of estimates across surveys. There are at least four main variations in the ADL instruments used in various national surveys:

i) Number and type of activities measured

55. The total number of ADL items range from 2 (in the case of the very concise EuroQol instrument) to 10 or more items (in the case of the WHO-Europe long-term disability list and the SF-36). Breaking this down by type of activities, some surveys include, for instance, only one or two questions related to self-care, while others include four questions or more.

56. Differences in the number of items included in a questionnaire is likely to affect disability prevalence estimates. As a general rule, the more activities that are being measured, the larger will be the number of people identified as disabled, if disability is defined as having one or more ADL limitations (Wiener *et.al.*, 1990; Rodgers and Miller, 1997; Madden and Wen, 1999). For the purpose of international comparability, it is therefore important to specify as much as possible the core group of items that should be included in the calculation of ADL-based disability, if no common instrument is used.

ii) Qualifiers/Scales (with/without help or special aids)

57. The use of qualifiers, such as with/without difficulty, with/without help or with/without special aids, are designed to measure the severity of the activity limitation (or disability).

58. A first important definitional issue is whether to count as disabled all the people who simply report having “some difficulty” in performing an activity (usually considered a “moderate” level of disability), only those who report having “a lot difficulty” or only those who rely on someone’s help or are not at all able to perform the activity (more “severe” levels of disability).

59. A second important source of comparability problems in ADL-based estimates is the differences in the qualifiers/scales used in various surveys. There is broad agreement now that the measurement of ADL activities (such as self-care) should distinguish clearly between those people who are able to do the activities on their own compared with those who can only do it with someone’s help (i.e., the inclusion of a “need for human assistance” scale). There is however less agreement on whether response scales for functional and ADL limitations should systematically distinguish between those who can do the activity without special aids and people who can only do it with special aids. The current practice in most surveys (but not all) is to allow the use of common aids, such as glasses and hearing aids, in asking people about their ability to do the activity. The Euro-REVES network recently recommended however a more systematic distinction in the measurement of functional and ADL limitations of people who can do the activity without or with special aids only (Euro-REVES, 2000).

iii) Question formulation (e.g., capacity versus performance)

60. While most surveys use a question formulation related to the “capacity” to do the activity (‘can you...?’) for ADL items, some surveys use different wording. These include:

- a 'performance' formulation, i.e. 'Do you...?' (in the French version of FR02)
- a 'limit' formulation (used in the SF-36 items), i.e. 'Does your health now limit you in these activities?' (AL01, DK01, GR01, NZ01)
- a 'difficulty' formulation, i.e. 'Do you find [this activity] difficult to do?' (AL02), 'Do you have any difficulty with...?' (CA02, IRE01, SP01, UK03), 'How difficult is it for you to...?' (ICE01, US01, US02).

61. Different item formulations may result in different answers. The issue of whether to use a 'capacity' wording (can you...?) versus a 'performance' wording (do you...?), and the resulting impact on prevalence estimates, has been examined in several studies (Box 2). In general, it seems that a module using a "capacity" formulation might overestimate the actual ability of a respondent to perform such an activity by about 15% to 20%.

Box 2. Measuring Capacity versus Performance in ADL Limitations

There are at least two ways of formulating questions on functional or activity limitations. One is to ask what a person "can" do while the other asks the person what he "does" do.

Anderson *et.al.* (1977) reported that capacity scores from self-completed questionnaires were approximately 20% higher than performance scores obtained through interviews in their study of 1324 San Diego residents. While scores were potentially confounded by the different data collection methods, the difference was attributed to the capacity versus performance formulations. Patrick *et.al.* (1981) also attributed lower (by 15-20%) disability prevalence estimates to the use of capacity formulations. More recently, in a study involving a very small number of physically disabled children, Young *et.al.* (1996) found that on average 17% fewer children 'did do' an activity that they reported they 'could do'.

Van der Wiel *et.al.* (1999) compared responses to capacity and performance items from 272 inhabitants of Leiden in the Netherlands, aged 85 years and over. Good agreement between capacity and performance items were found for slightly more complex activities such as 'shopping', 'telephone use' and 'walking outdoors', while poor agreement was found for basic activities such as 'dressing', 'washing' and 'transfers'. A "capacity" formulation items also identified less disability related to housework activities (such as cleaning and doing the laundry) than a "performance" formulation: 20% of the respondents claimed to be capable of doing these activities, but did not perform them on a regular basis (these respondents were typically men with higher incomes).

Branch and Meyers (1989) noted that the 'do you' formulation assumes that the activity is one that people absolutely need to do or would prefer to do themselves. This formulation appears to be most suitable to measure basic ADLs, such as self-care. However, for more complex IADL-type activities (such as doing housework), the 'can you' formulation is more suitable, since people may be able to perform these activities, but may prefer not to do these themselves. This is particularly relevant for those activities, such as housework, which tend to be gender-based.

iv) *Population coverage*

62. While some surveys ask the same ADL questions regardless of people's age or regardless of their reported health conditions, several others only asked a set of ADL questions to people above a certain age or people who have identified themselves as suffering from a disability (broadly defined). Amongst the latter group of surveys, the age threshold for asking ADL items varies from 55 years of age (in the case of the Dutch Health Interview Survey) to 60 (in the case of the Belgian Health Survey) to 65 (in the case of the French Health Interview Survey). Other surveys using ADL scales are designed to be administered solely to the elderly population (for instance, the National Long-Term Care Survey in the U.S. covers only the population 65 years and over). For the purpose of international comparisons, a larger number of countries might therefore be compared on ADL-based disability indicators for the population 65 years and over.

63. These are some of the main general variations between the various ADL scales presently used in national surveys. There are also some more specific variations in the instruments used to measure self-care, mobility and communication ability in different surveys. These are briefly reviewed.

2.1 Self-care

64. Twenty-five surveys include self-care items that are either the same or at least roughly similar to those recommended by WHO-Europe (Table 8). The most common are the ability to dress/undress (24 surveys from 19 countries) and to wash (23 surveys from 18 countries); over half of the surveys (18 surveys from 14 countries) also include at least one question on the ability to feed oneself.

65. Beside the type of variations in ADL instruments mentioned above, a few notable variations relate more specifically to how self-care activities are measured across surveys.

i) Combination of items

66. Seven surveys combine 2 or 3 activities into one item (mostly those using the SF-36 question combining the ability to wash and dress). For the purpose of cross-survey comparisons, the core group of self-care activities (washing, dressing, eating) might therefore be combined together into one category of activities. The disadvantage of this approach would be that it would generate comparability problems with those surveys that only ask questions related to one or two of these activities (e.g., the Swedish Survey of Living Conditions).

ii) Qualifiers/Scales

67. As already noted, an important difference in how surveys measure ADL activities such as self-care relates to qualifiers used in assessing the ability to perform the activity.

68. A fairly common approach – recommended in the WHO-Europe instrument - is to use a three-point response scale, along the following lines: “Can you dress/wash/feed yourself? Yes, without difficulty/Yes, with some difficulty/Only with help”. This range of responses allows the construction of a disability scale ranging from 'no disability', to 'moderate' levels of disability, to more 'severe' levels of disability (i.e., people requiring assistance).

69. Different scales are used however in different surveys. A distinction can be made between surveys that use a “difficulty” scale versus those that use a “need for human assistance”

scale. Among those surveys that use a “difficulty” scale, some surveys (e.g., those using the OECD list or the SF-36) make a distinction between “no”, “minor” or “major” difficulty, while others only specify “with difficulty” or “without difficulty”. Some surveys use items where the level of difficulty is specified within the question (e.g. Can you...without difficulty? or 'do you have any difficulty...?), with a simple “Yes/No” answer. This is equivalent to a ‘with or without difficulty’ response formulation. The use of a two-point or three-point scale for assessing the degree of “difficulty” in performing an activity can influence disability estimates (Box 3).

Box 3. Variations in ADL limitations in the Netherlands: The effect of various scales

Picavet and van den Bos (1995) examined the impact of differences in survey methodology on the prevalence of ADL-based disability in surveys of elderly people in the Netherlands. They concluded that seemingly minor differences in the structure and wording of the questionnaires resulted in major differences (up to 15.6 percentage points) in prevalence estimates of disability. These differences were mainly associated with the severity level of the disability implied by the wording of questions and answers.

Nine surveys targeted at the Dutch elderly were found to use disability items with response categories that included levels of severity (e.g. 'with difficulty', 'needing help' or 'unable to carry out the activity'). Two particular differences were studied with respect to ADL items: the distinction between minor and major difficulty and the distinction between 'unable to' and 'needing help' to carry out the activity as the most severe level of disability.

Where a distinction between minor and major difficulty was made (6 surveys), more people reported either 'minor' or 'major' difficulty (approximately 24%) compared to the number reporting simply 'difficulty' (approximately 19%), for almost all ADL items. Picavet and van den Bos concluded that the response category 'with difficulty' does not simply cover both minor and major difficulty, but instead represents something in between; as a result, the prevalence estimates obtained through these two different types of item formulation are not strictly comparable.

Where a distinction was made between 'unable to' (2 surveys) and 'only with help' (4 surveys) as the most severe level of disability, considerable variations were also seen in the proportion of respondents suffering from the most severe disabilities. These terms are therefore not interchangeable.

70. Consistent with the WHO-Europe recommended instrument, many surveys offer a "need for human assistance" response ('Only with help'), often as the most severe limitation (in some cases, they also offer an "unable to" category as the most severe restriction). A number of surveys however make no reference at all to help from others in their items (e.g., AL01, DEU01, ICE01, NZ01, SP01); they typically include "a lot of difficulty" as the most severe limitation.

71. Jette (1994) reported some large discordance between results coming from 'difficulty' scales and those coming from 'need for human assistance' scales. Comparing prevalence estimates of disability in a sample of 1818 adults aged 70 and over, Jette found that measures using a 'need for human assistance scale' produced lower estimates of disability, ranging from 1.6% persons disabled in walking to 8.6% disabled in bathing. By comparison, the use of a 'difficulty' scale produced estimates up to 5 times greater, especially for disability in walking (8.1%) and getting in and out of bed/chair (13.4% versus 3% obtained through a "need for human assistance" scale). The proportion of people reporting "difficulty" in bathing (12.2%) was also significantly higher than those reporting a "need for assistance" (8.6%). Differences across scaling methods were smaller in areas such as dressing and eating.

2.2 Mobility

72. Twenty-four surveys include mobility-related items (Table 8). The most common activities measured are the ability to climb stairs (21 surveys from 17 countries), to walk (19 surveys from 15 countries) and to get in and out of bed (18 surveys from 16 countries).

73. There are at least two types of variations regarding mobility items.

i) Intensity of activities

74. Beside differences in the number of activities that are being measured, there are also some variations in the intensity of the specific activity measured. This is particularly the case with respect to the ability to walk where there are significant variations across instruments regarding walking distance (ranging from a few steps, to 100 or 200 metres or more). Some instruments also include the specification to walk a certain distance “without resting” while others don’t.

ii) Qualifiers

75. As discussed for self-care activities, mobility items in different surveys vary with respect to whether or not they specify only “with/without difficulty” or also “with/without help”. This raises again the issue of comparability of data, particularly between the majority of surveys that use a “need for human assistance” scale versus those which use a ‘difficulty’ scale.

76. In addition, there are differences in the treatment of special aids across surveys. Most surveys do not specify whether the person is able to walk or climb stairs with special aids if necessary. Only 3 of the 16 surveys which include an item on walking ability specify 'with a walking stick [or other aids] if necessary' (IT01, UK01, UK03). On the other hand, the National Health Interview Survey in the US specifies whether people can walk and climb stairs 'without using any special equipment'. Allowing the use of various technical aids should obviously result in a lower percentage of respondents being defined as disabled compared with those surveys which do not allow the use of those technical aids.

2.3 Communication

77. Twenty surveys include seeing, hearing and speaking items more or less similar to those recommended by the OECD and WHO-Europe (Table 8). The most common items relate to hearing (18 surveys from 16 countries), seeing (16 surveys from 13 countries), while 12 surveys from 10 countries also include items related to speaking ability.

78. Here as well there are two types of variations in the instruments used to measure these abilities.

i) Specific reference situation

79. Various questions can be used to measure eyesight in population surveys. The most common way is to ask people whether they “can see well enough (with glasses or contact lenses if necessary) to recognize a face at a distance of four metres”. This specific item was included in both the OECD and WHO-Europe list; it is designed to measure the ability to see at a distance. The OECD list also included another item which continues to be used in several surveys related to the ability “to read ordinary newsprint”; this question is designed to measure eyesight at a

close distance. These two different questions measure different aspects of the ability to see; they cannot be taken as being interchangeable.

ii) *Qualifiers/Scales*

80. Some surveys specify 'without difficulty' in the questions related to seeing and hearing (DK01, SP01, SWE01, UK03, US02), with a simple “yes/no” response category, while others offer various levels of difficulty (no/some/a lot of difficulty) in the answer categories (FR01, FR02, IRE01, NL01, SZ01). This variation might create again problems of data comparability.

81. As noted before, most surveys do not try to make a distinction between people who are able to see or hear with or without a special aid, by including in the question “with glasses/contact lenses/hearing aids if necessary”. The Irish National Health and Lifestyle Survey appears to be an exception as it stands in specifying “without a hearing aid” in measuring hearing capacity.

Summary on ADL Items

82. This inventory confirms that ADL scales are commonly used to measure various activity limitations, in particular as they relate to the elderly population. Table 8 has shown that there is a core group of activities that are being measured across most surveys. These include activities related to self-care, mobility and communication.

83. However, a number of different ADL instruments are now being used in national surveys, resulting in a lack of data comparability. These ADL instruments vary notably in terms of the number of specific activities being measured and the scales used in assessing the degree of activity limitations (e.g., “difficulty” versus “need for human assistance” versus “need for special aid” scales). As a result, the only ADL-based disability estimates that are strictly comparable across surveys (and countries) are those using the same instrument. Based on this inventory, these would include the group of countries using:

- the WHO-Europe “long-term disability” instrument (Belgium, Czech Republic, Portugal and the 1995 Health Survey for England);
- the SF-36 instrument (Australia, Denmark, Germany, Italy, New Zealand and the 1996 Health Survey for England);
- the EuroQol-5D instrument (Denmark, Ireland and the 1996 Health Survey for England).

2.4 IADL

84. Instrumental Activities of Daily Living (IADL) comprise a number of activities beyond ADL which are considered important for people to live independently in the community. Although reported limitations in IADL may reflect somewhat less severe levels of disability than limitations in ADL, such limitations may nonetheless express a need for social support.

85. Fourteen surveys (from 11 countries) include at least some IADL items, with the majority asking about the ability to do housework, shopping and prepare meals (Table 9). Other IADL items are also used in different surveys, complicating cross-survey comparisons, particularly when IADL ‘disability’ is defined as being limited in at least one of these activities. The greater the number of IADL activities, the more likely it is that someone will be limited in at least one of these activities.

86. Beside the number of IADL items, another difference relates to the intensity of the activity measured across surveys. For instance, some surveys refer to “light” housework while others mention “heavy’ work.

87. On the other hand, nearly all IADL items include a ‘need for human assistance’ scale (in either the question or the response categories), except for ICE01 and SP01 which only measure the degree of difficulty.

88. Taking the above into account, a dozen surveys (from 11 countries) might provide reasonably comparable data of IADL limitations, if these focus only on the more common activities measured across surveys (i.e., housework, shopping, meal preparation).

Table 9: IADL and more general activity limitation items in OECD surveys

	IADL			Work	Social	General items
	Housework	Shopping	Meal preparation			
AL01				√	√	√
AL02				√ ¹	√ ¹	√ ¹
AT01	√	√	√		√ ²	
BL01				√	√	√
CA01	√	√	√	√ ²		√ ²
CA02	√	√	√	√ ²		√ ²
DK01				√	√	√
FR01			√			
FR02	√	√	√	√		√
DEU01				√	√	√
ICE01	√			√		
IRE01						√ ²
IT01	√	√	√	√	√	√
KR01					√ ³	
NL01						√
NZ01				√	√ ²	√ ²
NZ02	√	√	√	√		
NR01	√	√				√ ²
SP01	√ ²	√ ²	√ ²		√ ²	√
SWE01 ₄	√	√	√			√
SZ01						√
UK01						√
UK02				√		√
UK03	√	√	√			√ ²
US01	√	√		√	√	√
US02 ⁵	√	√	√			

1. Work item asked if the respondent is disabled and employed; Social and General items asked if disabled and aged ≥15 yrs
2. Item specifies restrictions due to condition lasting at least 6 months (CA01, CA02, NZ01) or 1 year (SP01)
3. Only asked if a person has mobility problems
4. Only asked if old-age pensioners or if the person has a long-standing illness, reduced sight, mobility problems or aged 75-84 years (IADL items)
5. Only for persons aged ≥65 years

Illustrative types of questions

IADL:

Housework: “Can you do your own heavy/light housework/cleaning/laundry” or, turn the other way around,

“Do you need help with...?”

Shop: “Can you do your own shopping?” or “Do you need help with...?”

Meals: “Can you prepare your own meals?” or “Do you need help with...?”

Work: “Does your health/illness/chronic condition restrict/limit/interfere with your work?”

Social: “Are you restricted in your social activities due to your health/illness/chronic condition?”

General: “Are you restricted in your daily activities due to your health/illness/chronic condition?”

2.5 Limitations specific to work and social activities

89. Seventeen surveys (from 13 countries) include items that specifically ask about limitations in work and/or social activities (Table 9). Seven of these surveys (6 countries) use the SF-36 items (AL01, AL02, DK01, DEU01, IT01, NZ01 and UK01), which include eight items about the effect of health problems (both physical and psychological) on work and social activities and an item on the effect of pain on work (including housework). Most other surveys use a more general formulation along the lines 'Does your health affect your social activities/work?', with a simple “Yes/No” answer. One survey (AT01) refers to the need for help in carrying these activities, while three surveys enquire about the level of difficulty (ICE01, SP01 and US01).

90. With respect to the reference period, except for those surveys using SF-36 items, which refer to the previous 4 weeks, limitations in work and/or social activities relate to the current situation.

91. The comparability of survey instruments to measure limitations in work or other social activities appear to be limited to those using a similar instrument (e.g., either the SF-36 or a more general question formulation with a “Yes/No” response).

2.6 General item on limitation in usual activities (also referred as global activity limitation)

92. A general question (or a few questions) on limitations in usual activities is designed to provide a summary measure of people who report being disabled because of health problems. In most countries where data based on this indicator is available, between 15% and 25% of the adult population report being limited in their usual activities because of health problems. Because the question does not specify the type of activities, it can be administered to a population of all ages. However, the general (unspecified) nature of the question also means that responses to it are more subjective and more subject to reporting biases than responses to more specific questions on well-defined activity limitations.

93. Twenty surveys (from 16 countries) include at least one question on general limitations in usual activities (Table 9; last column). The European Community Household Panel also includes a general item on “limiting long-standing illness”, thereby broadening the number of countries for which comparable data might be available. The most common formulation refers to

whether or not usual activities are limited (or hampered/restricted/affected) due to health problems (BL01, CA01, CA02, IRE01, NL01, NR01, SZ01, UK02, UK03)⁹.

94. There are at least three sources of variations however in the specific item formulation used across surveys.

i) Qualifiers

95. A first important source of variations across surveys relates to the response formulation. Three types of response scales are offered to the general question on activity limitations:

- a number of surveys offer a simple “Yes/No” answers to the question “are you limited in your usual activities due to a health problem?” These surveys include the two Canadian health surveys, the New Zealand Health Survey (NZ01), the Spanish survey, the Health Survey for England (UK01) and the U.S. National Health Interview Survey (US01).
- other surveys offer a “difficulty” scale, typically ranging from “Yes, severely/Yes, to some extent/No, not limited”. These include surveys using the EuroQol-5D instrument (e.g., Denmark, Ireland), the Norwegian Health Interview Survey, the Swedish Survey of Living Conditions, as well as all the national results based on the European Community Household Panel.
- a few other surveys offer a “chronic” scale, with responses ranging from “Continually/At intervals/Not or seldom”, to measure to what extent the limitation in usual activities tends to be chronic or temporary. Such a response formulation is used in the Belgian and Dutch Health Interview Surveys.

96. The effect of using a “difficulty” scale versus a simple “yes/no” scale on the comparability of results across surveys hasn’t been fully examined yet. Ideally, such a comparison should be done directly in a single study, using the same population sample and the same survey method. It is possible however to obtain a general sense of the comparability of results, by comparing the national estimates from the European Community Household Panel (using a “difficulty” scale) with those from countries which also conduct a national survey using a simple “yes/no” scale, such as the Health Survey for England. The data from these two different item formulations appear to be generally consistent, with 25% of people in the United Kingdom reporting being hampered in their usual activities due to health problems in the 1996 ECHP (with 8% reporting being severely hampered and 17% being hampered to some extent), compared with 26% of people responding “yes” to the question about “limiting long-standing illness” in the 1996 Health Survey for England. If anything, one might have expected higher “disability” rates from the ECHP item formulation, given that its scale is designed to capture both severe and more moderate limitations in usual activities. Further work is required to examine whether and how estimates from a binary (yes/no) response need to be adjusted to be comparable with estimates from a “difficulty/severity” scale.

ii) Number of questions and degree of specificity

9 . The Euro-REVES network recently proposed the use of a new global activity limitation instrument, in national surveys, that is: “For the past 6 months or more have you been limited in activities people usually do because of health problem? Yes, strongly limited/Yes, limited/No, not limited” (Euro-REVES, 2000).

97. Another important source of variation across surveys relates to the number of questions used to measure usual activity limitations generally, and the specification of the health problems causing the activity limitation. Surveys which use the SF-36 items (AL01, AL02, DK01, DEU01, IT01, NZ01) include 4 questions related to the effect of physical health on activity limitations, and 3 items concerning the consequences of any emotional problems on limitations in usual activities. This compares with only one general question on activity limitations for most other surveys. The results coming from more specific question formulations are not likely to be comparable with those coming from a single (more general) item.

98. Some surveys (such as CAN01 and CAN02) specify broad categories of activities in their question formulation, such as activities “at home”, “at work” or “at school”. This formulation may not pose significant problems of comparability with those surveys that do not specify any activities, if the estimates for the various broad types of activities can be added into a summary estimate of overall activity limitations: (as is done in Canada).

iii) Reference period

99. There is also some variation in the reference period mentioned in the question. While most of the surveys refer to the current situation, an alternative formulation found in those surveys using the SF-36 asks about limitations in usual activities during the past month.

100. Finally, a more minor source of variation concerns the treatment of medication in the item formulation. The U.K. Labour Force Survey (UK02) appears to be unique in stipulating that general activity limitations should be answered ignoring the effect of any medication or treatment. Such a specification can be expected to increase the percentage of people reporting “limiting long-standing illness” from that survey.

101. In summary, the comparability of a general question on limitations in usual activities is limited by the use of different response scales. Further work is required to determine to what extent estimates based on a “difficulty/severity” scale for this question (mainly used in European countries and in the European Panel) are comparable with estimates from a simple binary (“yes/no”) response scale (presently used mainly in non-European countries like Canada, New Zealand and the U.S.).

SUMMARY AND CONCLUSIONS

Main Findings

102. The main finding of this review of 30 surveys from 23 OECD countries (plus the European Community Household Panel) is that, beside a few general items related to the prevalence of chronic conditions and general limitations in usual activities, differences in measurement instruments in domains such as mental health and specific activity limitations seriously limit the comparability of data to only those sub-group of countries using the same measurement instrument. Table 10 summarizes the degree of comparability of different survey items, based on this inventory and on information already collected in OECD Health Data with regard to self-rated general health.

Table 10. Comparability/availability of items addressing health and disability in national surveys

Items	Comparability/availability
Self-rated general health ¹	+
Chronic conditions: General (yes/no)	+
Chronic conditions: Disease-specific	+
	(only for a core group of diseases)
Mental health: Psychological/Emotional	-
Mental health: Cognitive	-
Pain	-
Disabilities: ADL	-
Disabilities: IADL	-
Disabilities: General (yes/no)	+

+: reasonably comparable/available -:comparability/availability limited

1. The comparability of survey questions on self-rated general health hasn't been covered in this report; the information on this item is available in the "sources and methods" section of *OECD Health Data*.

103. The main problem in terms of comparability is not "what" is being measured in the various surveys, since there tends to be a fairly common group of health domains and activity

limitations measured across surveys. Rather, the problem for comparability purposes lies in “how” specifically these health conditions and activity limitations are being measured. This is probably best illustrated by the review of activities of daily living (ADL) items in different surveys. While most surveys use some ADL items to measure disability rates in the population (in particular amongst the elderly population), the use of various ADL scales (e.g., the OECD long-term disability list, the WHO-Europe long-term disability instrument or the SF-36 instrument) limits the comparability of data only to those group of countries using the same instrument.

104. Based on this inventory, it would seem possible to now collect reasonably comparable data for a fairly large number of OECD countries for the following health and disability indicators:

- i. General prevalence of chronic conditions (Yes/No)
- ii. Prevalence of a few specific chronic conditions (e.g., high blood pressure, heart disease, asthma and chronic bronchitis, and diabetes), although the reliability of survey data on these specific conditions should be tested by comparing survey estimates with other sources of morbidity data
- iii. General activity limitations (Yes/No)

105. Even for these few (mainly general) items, the comparability of data would be improved by moving towards the use of common instruments in existing or new national (and international) surveys.

Approaches to improve the comparability of health status and disability-related data

106. There is keen interest in many OECD countries to monitor the health status of the population and disability rates, both nationally and across countries, to promote informed policy discussions on the impact of population ageing and the integration of persons with disabilities in social and economic activities. Three broad approaches might be pursued to achieve greater international comparability of health status and disability-related indicators:

- i. developing and administering international surveys
- ii. encouraging the use of common instruments in national surveys
- iii. applying some “post-harmonization” methods (to adjust for variations in survey instruments)

i. Development and administration of international surveys

107. Probably the surest way to obtain internationally comparable data on health status and disability is through the development and administration of a common international health-related survey across all participating countries. But this is also the most ambitious and resource-intensive approach. The European Community Household Panel (although it contains only a few general questions on health and disability issues) provides an example of such an international survey which has been developed and administered at the European level. The development of an *ad hoc* module on disability to be included in the European Union Labour Force Survey of 2002 will provide another source of comparable data on a few disability-related items for the working-age population at the European Union level.

108. At the OECD level, the International Adult Literacy Survey provides an example of a co-operative international effort involving national statistical agencies, research institutions and the OECD Secretariat to develop and administer a common survey, thereby providing comparable data in the area of education, training and literacy skills across the 20 participating countries. Such a survey however requires a substantial and sustained commitment of time and resources from all partners.

ii. Encourage the use of common instruments in national surveys

109. A “second best” approach to developing more comparable health status and disability indicators relies on the willingness of national statistical agencies to use some common instrument(s) to measure at least a few key aspects of health and disability in upcoming national surveys.

110. At the European level, a number of projects have recently been completed or are presently underway to further develop and recommend common instruments to measure health and disability in national surveys. These projects include the past and current phases of the EUROHIS project, led by WHO-Europe and funded by the European Commission, and the recent Euro-REVES network project to develop a consistent set of health expectancies in Europe (also funded by the European Commission). Table 11 summarizes the instruments recommended thus far through these two European projects for selected health and disability indicators. While these European projects are proposing the use of these instruments, the ultimate success of these efforts to promote data comparability will depend on their actual implementation in national surveys. We have noted that the 1996 WHO-Europe recommended instruments to measure mental health and long-term (ADL) disability have been used in a few recent surveys in Europe, while the Euro-REVES recommendations on the measurement of disability (both ADL-related limitations and general disability) are too recent to have been implemented.

Table 11. Recommended instruments for selected health and disability items in Europe: WHO-Europe (1996 and current EUROHIS project) and Euro-REVES (2000)

Items	Recommended instrument(s)
General health status	WHO-Europe (1996): “How is your health in general? very good, good, fair, bad, very bad”
Chronic conditions: General (yes/no)	WHO-Europe (under development in EUROHIS)
Chronic conditions: Disease-specific	WHO-Europe (under development in EUROHIS)
Mental health: Psychological/Emotional Problems	WHO-Europe (1996) and Euro-REVES (2000): General Health Questionnaire– 12 items (see Goldberg, 1972)
Disabilities: ADL (including both functional and activity limitations)	WHO-Europe (1996): Long-term disability instrument (13 questions plus 3 optional) Euro-REVES (2000): proposes adjustments to WHO-Europe long-term disability instrument

	(e.g., breakdown of items to measure “physical and sensory functional limitations” versus “activity restrictions”, and greater range of severity levels)
Disabilities: General (yes/no)	Euro-REVES (2000): “For the past 6 months or more have you been limited in activities people usually do because of a health problem? Yes, strongly limited/ Yes, limited/No, not limited”

111. In addition, the present inventory has identified three leading generic health status measurement instruments which have been used to measure key aspects of health status and activity limitations in more than one national survey. These are:

- i. the SF-36 instrument (or its abbreviate versions), which has been included in national surveys in different parts of the OECD (Australia, Denmark, Germany, Italy, New Zealand and the 1996 Health Survey for England);
- ii. the EuroQol-5D instrument, which has been included in the surveys of some European countries (Denmark, the 2000 Health Survey in Finland, the 1998 National Health and Lifestyle survey in Ireland and the 1996 Health Survey for England as well).
- iii. the HUI-3, which has been included in health surveys in Canada as well as in a few other non-OECD countries (e.g., Singapore).

112. These generic health measurement instruments have become increasingly used as a way to provide summary health profiles (in the case of the physical and mental health components of the SF-36 or SF-12) or health indices which can summarize health status in a single number (in the case of EuroQol-5D and HUI-3). Table 4 provided an overview of the coverage of these leading generic instruments. The main differences between these instruments relate to the number of specific health and disability dimensions measured and the emphasis placed on measuring functional health status as opposed to measuring general activity limitations.

113. A first important strategic issue in moving towards the use of common instruments across national surveys is whether to encourage the use of *existing* measurement instrument(s) or whether to develop and recommend *new* instrument(s). As a general rule, given the proliferation of health measurement instruments in recent years and the increasing number of recommendations for “standard” instruments, it would seem preferable and more expeditious to build as much as possible on current leading instruments rather than start with the construction and recommendation of new ones, unless it can be shown that such new instrument(s) would be superior in measuring certain aspects of health or easier to implement in national surveys.

114. A number of guiding criteria can be used to select recommended measurement instruments whether these are generic instruments or more specific ones. In addition to meeting the basic criteria of validity and reliability, the choice of “standard” measurement instruments to be recommended for use in national surveys might be based on the following three criteria:

- saliency: the instrument should measure what are considered to be important dimensions of health and disability across countries;
- conciseness: the instrument needs to achieve a balance between measuring the multi-dimensional nature of health and/or disability and the practical requirement to be as concise as possible to facilitate implementation in national surveys;
- cross-country comparability: the instrument should contain questions/items that are the least susceptible to being influenced by social and cultural differences across countries.

115. An on-going dilemma in choosing survey instruments is to strike a balance between comprehensiveness (in terms of covering all important aspects of health and activity limitations in sufficient detail) and conciseness (to facilitate survey implementation). Another important dilemma or constraint faced by national statistical agencies in moving towards the adoption of common measurement instrument(s) in national surveys is the need to maintain the continuity of existing time series. This is obviously not a constraint for new surveys. In the case of surveys that have been conducted for some years, an interesting approach is the one that has been adopted in the 2000 Health and Morbidity Survey in Denmark, whereby a new leading measurement instrument (EuroQol-5D) has been administered for at least some sub-samples of the population, in addition to the items used in previous Danish health surveys (including the SF-36 instrument which has been included in both the 1994 and 2000 Danish surveys). This “juxtaposition” approach allows both the maintenance of existing (national) time series and greater international comparability with countries using the same common measurement instrument(s).

116. This approach obviously faces its own limitation, since there are limits to the overall number of questions that can be included in national surveys. Conciseness will be a definite advantage for promoting the diffusion of leading instruments across national surveys.

117. An important issue related to promoting the use of common measurement instruments across countries is to address the problems associated with finding the equivalent translation and adaptation of questions and answers in various languages and for various cultures. In the past, not enough attention has been paid to the equivalent translation of recommended instruments for use in different countries.

iii. Application of post-harmonization methods

118. The third possible method for more comparable data across surveys is to try to use some “post-harmonization” methods, to adjust for existing variations in survey instruments and also to try to correct for any systematic reporting biases.

119. In this report, we have reviewed some of the available evidence on the effect of various item formulations on the comparability of estimates. For instance, we have noted that the effect of a “capacity” formulation for ADL items may lead to estimates of the ability of people to carry out these activities about 15% to 20% higher than those obtained from a “performance” formulation. We have also reviewed the influence of using various scales (e.g., “difficulty” versus “need for assistance” scales) in measuring ADL-based limitations (with evidence from a study of elderly people in the U.S. suggesting that for some activities, such as walking and

getting in and out of bed, a “difficulty” scale may produce disability estimates up to five times greater than a “need for human assistance” scale).

120. It may be possible to make some adjustments to survey estimates that use different item formulations, to make them more comparable with data from other surveys. However, this “post-harmonization” approach will be limited to areas where sufficient studies have been done about the precise effect of differences in item formulations, to come up with an acceptable adjustment factor. In cases where an adjustment factor is derived from the study of the response pattern of one group of the population in one country, questions can be raised about the extent to which these results can be extended (generalized) to other population groups in that country and to populations in other countries. To become fully acceptable, these adjustment factors would need to be validated across various populations through direct studies of the effect of differences in item formulations.

121. “Post-harmonization” efforts often lead to a search for the “lowest common denominator” across surveys. This is particularly the case for the various scales used to assess the degree of activity limitations, where the lowest common denominator is often to collapse responses in a simple “Yes/No” category. In some cases, there is no evidence that such a move towards the “lowest common denominator” affects the comparability of results, while in other cases there is some evidence that it might.

122. This report has not addressed another important issue regarding the cross-country comparability of survey data, that is, the possibility of systematic reporting biases due to social or cultural differences in response patterns across countries. Over and above differences in survey questions, serious reservations have been expressed by WHO-Geneva and others about the overall reliability of survey data to allow meaningful comparisons of health and disability across populations. The general concern is that people’s *self-report* of their health may suffer from a number of reporting biases (e.g., differences in health knowledge, expectations and norms, or differences in the availability of disability-related programmes), thereby providing some misleading information about their *true* health status. WHO-Geneva has started to use different statistical techniques (e.g., factor analysis, Item Response Theory model) to try to correct for reporting biases at the individual and population level, by re-scaling survey responses (Sadana and al., 2000). Although the initial results of these attempts to improve the comparability of survey data across populations have been disappointing, WHO-Geneva intends to pursue its effort to “calibrate” survey data to make them more comparable across populations.

123. In general, any “post-harmonization” methods will be facilitated if instruments used in different surveys are as similar as possible, so that the instruments are indeed measuring more or less the same health condition or activity limitation. Therefore, efforts to harmonize input (survey questions) and output (adjustments to variations in survey questions or for reporting biases) are complementary and need to be carried out simultaneously, if we are to improve the international comparability of health and disability data.

REFERENCES

- Anderson, J.P., Bush, J.W. and Berry, C.C. (1977)
Performance versus capacity: a conflict in classifying function for health status measurement. Presented at the American Public Health Association Meeting, Washington DC, 1977.
- Australian Bureau of Statistics (1998)
Disability, ageing and carers: summary of findings.
- Bebbington, A. (1992)
Expectation of life without disability measured from the OPCS disability surveys. Health expectancy: First workshop of the International Healthy Life Expectancy Network (REVES).
- Bebbington, A. and Darton, RA (1996)
“Healthy life expectancy in England and Wales: recent evidence”. Discussion Paper 1205.
- Branch, L.G. and Meyers, A.R. (1987)
“Assessing physical function in the elderly”. *Clinics in Geriatric Medicine*: 3(1): 29-51.
- Brooks. R. and the EuroQol Group” (1996)
“EuroQol: the current state of play”. *Health Policy*: 37: 53-72.
- Bruckner, G. (1997)
“Health expectancy in Germany”. Paper presented at the REVES 10 meeting of the Network on Health Expectancy.
- Cambois, E. and Robine, J.M. (1994)
« Théorie et analyse des principales enquêtes santé portant sur l’incapacité dans les pays occidentaux ». Equipe INSERM *Démographie et Santé*, Montpellier : pp. 1-33.
- Crimmins, E.M., *et.al.* (1997)
“Trends in disability-free life expectancy in the United States, 1970-90”. *Population and Development Review*: 23(3): 555-72.
- Euro-REVES (2000)
(edited by J-M Robine, C. Jagger and V. Egidi), *Selection of a Coherent Set of Health Indicators: First Steps Towards a User’s Guide to Health Expectancies for the European Union*, Montpellier, France.
- Feinstein, A.R., Josephy, B.R., and Wells, C.K.(1986)
“Scientific and clinical problems in indexes of functional disability”. *Annals of Internal Medicine* 105: 413-20.
- Furlong, W., Feeny, D., Torrance, G.W., *et. al.* (1998)
Multiplicative Multi-attribute Utility Function for the Health Utilities Index Mark 3

(HUI3) System: a technical report, McMaster University Centre for Health Economics and Policy Analysis Working Paper No. 98-11, December .

- Goldberg, D.P. (1972),
The Detection of Psychiatric Illness by Questionnaire, Oxford University Press, London.
- Gudex, C., Christensen, S. and Rasmussen, N. (1999)
“Survey data on disability: final project report”. Eurostat: *Population and social conditions* 3/1999/E/no. 20, November.
- Hibbett, M.J., Jagger, C., Polge, C., Cambois E. and Ritchie, K. (1999)
“Mental health. Cross-European mental health indicators. A dream or reality?” *European Journal of Public Health*: 9: 285-9.
- Jacobzone, S., *et.al.* (2000)
“Is the Health of Older Persons in OECD countries improving fast enough to compensate for population ageing?”, *OECD Economic Studies*, no. 30. pp. 149-190.
- Jette, A.M. (1994)
“How measurement techniques influence estimates of disability in older populations”.
Social Science and Medicine: 38: 937-42.
- Jenkinson, C., Wright, L. and Coulter, A. (1993)
Quality of life measurement in health care: a review of measures and population norms for the UK SF-36. Health Services Research Unit, University of Oxford.
- Katz, S., Ford, A.B., Moskowitz, R.W., *et.al.* (1963)
“Studies of illness in the aged. The Index of ADL: a standardized measure of biological and psychosocial function”, *JAMA*, 185, pp. 914-919.
- Lee. (1997)
“A study of disability-free life expectancy of elderly in Korea”. REVES paper 308.
- Mathers, C. (1997)
“Health expectancies and DALYs: Towards consistency of measures”, Working Party on Social Policy, Ad hoc Meeting of Experts in Health Statistics, 3-5 December 1997, Room Document no. 4, Paris.
- McWhinnie, J.R. (1982)
Measuring Disability. OECD Social Indicator Development Programme, Special Studies No. 5.
- Madden, R. and Wen, X. (1999)
“Towards international consistency in collecting information on ageing and disability”. Report for the OECD Working Party on Social Policy. Australian Institute of Health and Welfare, Canberra, December.
- McDowell, I, Newell, C. (1996)
Measuring Health: a guide to rating scales and questionnaires. 2nd ed. New York/Oxford: Oxford University Press.
- New Zealand Ministry of Health (1998)
Disability in New Zealand – overview of the 1996/97 Surveys.

- Patrick, D.L., Darby, S.C., Green, S., *et.al.* (1981)
 “Screening for disability in the inner city”. *Journal of Epidemiology and Community Health* : 35: pp. 65-70.
- Picavet, H.S.J., van den Bos, G.A.M. (1996)
 “Comparing survey data on functional disability: the impact of some methodological differences”. *Journal of Epidemiology and Community Health* : 50: 86-93.
- RIVM – National Institute of Public Health and Environment. (1998)
Public health status and forecasts 1997 - Health, prevention and health care in the Netherlands until 2015. RIVM.
- Robine, J.M. and Jagger, C. (1999)
 “Developing consistent disability measures and surveys”, paper prepared for the OECD meeting on implications of disability for ageing populations: monitoring social policy changes, 9-10 December, DEELSA/ELSA/WP1/DIS(99)7, Paris.
- Robine, J.M., Romieu, I. and Jee, M. (1998)
 “Health Expectancies in OECD Countries”, REVES paper no. 317.
- Robine, J.M. et Mormiche, P. (1993)
 « L’espérance de vie sans incapacité augmente ». *INSEE Première* 281.
- Rodgers, W. and Miller, B. (1977)
 “A comparative analysis of ADL questions in surveys of older people”. *Journal of Gerontology*. Series B 1977: 52B (special issue): 21-36.
- Smith, L.A., Branch, L.G., Scherr, P.A., Wetle, T., Evans, D.A., Hebert, L. and Taylor, J.O. (1990)
 “Short-term variability of measures of physical function in older people”. *Journal of the American Geriatrics Society*: 38(9): 993-8.
- Statistics Canada. (1991)
Health and Activity Limitations Survey: User guide.
- Statistics Canada (2000)
 “How Healthy are Canadians?”. *Health Reports – A Special Issue*, No. 82-003-XPB.
- Valkonen, T., *et.al.* (1994)
Disability-free life expectancy by level of education in Finland.
- van der Wiel, AB., Gussekloo, J., van Exel, E., Lagaay, A.M., Knook, D.L. and Westendorp, R.G.J. (1999)
 “Measuring disability: the differences between 'can do' and 'do do’”. Abstract of paper presented at the IVth European Congress of Gerontology. 32(2 Suppl.): 666.
- Verbrugge, L.M. (1990)
 “The iceberg of disability”. In: Stahl S.M., ed.: *The legacy of longevity: health and health care in later life*. Newbury Park, CA: Sage: pp. 55-75.
- Ware, J.E. and Sherbourne, C. (1992)
 “The MOS 36 item short form health survey. I. Conceptual framework and item selection”. *Medical Care*: 30: pp. 473-83.

WHO-Europe; CBS (1996)

Health interview surveys: towards international harmonisation of methods and instruments. WHO Regional Publication European series No. 58, Copenhagen.

WHO-Geneva (1980)

International Classification of Impairments, Disabilities, and Handicaps.

WHO-Geneva (1999)

ICIDH-2: International Classification of Functioning and Disability. Beta-2 draft, Short Version.

Wiener, J.M., Hanley, R.J., Clark, R. and Van Nostrand, J.F. (1990)

“Measuring the activities of daily living: comparisons across national surveys”. *Journal of Gerontology*: 45(6): S229-37.

Young, N.L., Williams, J.I., Yoshida, K.K., Bombardier, C., and Wright, J.G. (1996)

“The context of measuring disability: does it matter whether capability or performance is measured?” *Journal of Clinical Epidemiology*: 49(10): 1097-101.