The Journal of TRAUMA® Injury, Infection, and Critical Care

Guidelines for the Conduction of Follow-up Studies Measuring Injury-Related Disability

Ed F. Van Beeck, MD, PhD, Claus F. Larsen, MD, PhD, Ronan A. Lyons, MD, PhD, Willem-Jan Meerding, PhD, Saakje Mulder, PhD, and Marie-Louise Essink-Bot, MD, PhD

Background: **Scientific knowledge on functional outcome after injury is limited. During the past decade, a variety of measures have been used at various moments in different study populations. Guidelines are needed to increase comparability between studies.**

Methods: **A working group of the European Consumer Safety Association conducted a literature review of empirical studies into injury-related disability (1995–2005). We included injury from all levels of severity and selected studies using generic health status measures with both short-term and long-term follow up. The results were used as input for a consensus procedure toward the development of guidelines for defining the study populations, selecting the health status measures, selecting the timings of the assessments, and data collection procedures.**

Results: **The group reached consensus on a common core of health status measures and assessment moments. The group advises to use a combination of EuroQol-5D and Health Utilities Mark III in all studies on injury-related disability. This combination covers all relevant health domains, is applicable in all kinds of injury populations and in widely different age ranges, provides a link with utility scores, and has several practical advantages (e.g., brevity, availability in different languages). For specific types of injury, the common core may be supplemented by injury-specific measures. The group advises a common core of assessments at 1, 2, 4, and 12 months after injury.**

Conclusions: **Our guidelines should be tested and may lead to improved and more consistent epidemiologic data on the incidence, severity, and duration of injury-related disability.**

Key Words: **Injury, Disability, Guidelines.**

J Trauma. 2007;62:534–550.

Disabilities (i.e., reduced levels of functioning resulting from diseases or injuries)[1] are increasingly seen as an important component of a population's health. This has been recognized in the field of injury prevention and trauma care,[2] where the number of survivors of severe injury has rapidly risen.[3] Moreover, many survivors are young people, whose daily activities may greatly and/or permanently suffer from the consequences of trauma.[4] However, comparable and representative epidemiologic data on the incidence, severity, and duration of injury-related disabilities are still scarce and incomplete. Most functional outcome studies in this area have so far focused on adult patients (mostly within the age range of 15–64 years) with severe trauma, such as polytrauma,[5] traumatic brain injury,[6–9] and spinal cord injury.[10,11] Only a few studies have been conducted already on the functional outcome of (the more severe) childhood injuries.[12–14] The functional outcome of injuries among the elderly has so far mainly been studied for patients with hip fractures,[15,16] although some studies on geriatric trauma patients in general have been published.[17] For many types of injury, however, hardly any empirical disability data are available yet. In addition, the available knowledge is difficult to interpret. During the past decades, a variety of measures have been used, which makes a comparison of the available disability estimates rather difficult. Moreover, these studies have focused on a variety of health domains (leading to incomplete information) at various moments in a variety of patient populations (leading to incomparable information). Most of the disability estimates obtained so far can therefore not be used to quantify the impact of injury-related disability on population health.[2] Nor do they allow evaluations of the (cost-) effectiveness of injury prevention and/or trauma care. To stimulate new epidemiologic data collections better fitting these purposes, the European Consumer Safety Association (ECOSA) has established a working group on injury-related disability. This article reports the current progress of this group, which has evaluated the available literature (i.e., empirical epidemiologic studies) on postinjury levels of functioning, and has developed guidelines for future empirical work in this area.

Submitted for publication April 10, 2006.

Accepted for publication November 8, 2006.

From the Department of Public Health (E.F.V.B., W.J.M., M.-L.E.-B.), Erasmus MC, University Medical Centre Rotterdam, the Netherlands; the Department of Orthopaedics (C.F.L.), Hillerod Hospital, Hillerod, Denmark; the School of Medicine (R.A.L.), University of Wales, Swansea, United Kingdom; the Consumer Safety Institute (S.M.), Amsterdam, the Netherlands.

Address for reprints: Ed F. van Beeck, MD, PhD, Department of Public Health, Erasmus MC, University Medical Centre Rotterdam, PO Box 2040, 3000 CA Rotterdam, the Netherlands; email: e.vanbeeck@erasmusmc.nl.

DOI: 10.1097/TA.0b013e31802e70c7

MATERIALS AND METHODS

Demarcation of the Subject

The working group aims to develop standards and guidelines for quantifying the total amount of disability at the

population level. It looks at injury from all causes and from all levels of injury severity, because even minor injuries may lead to a substantial health burden as a result of their frequent occurrence.[4,18] The working group has adopted the broad definition of disability according to the International Classification of Disabilities, Functioning and Health (ICF) of the World Health Organization (WHO).[1] According to the ICF, disability is an overall term that includes all the negative aspects of the following four components: body structures (anatomic body parts), body functions (physiologic and psychological functions), activities (execution of tasks or actions by an individual), and participation (involvement in life situations). It should be noted that, unlike the previous terminology of the International Classification of Impairments, Disabilities and Handicaps (ICIDH), this current disability concept includes bodily impairments, activity limitations, and restrictions in social participation. Each of the four components of the ICF consists of various domains that, in principle, could all be relevant for injury patients.

The working group uses the framework of the ICF to assess whether all relevant health domains for (specific groups of) injury patients are included, when measuring the functional consequences of injury. The choice of a meaningful timeframe for measuring disability is a rather difficult but important issue. The working group distinguishes the following phases:

- Acute treatment phase
- Rehabilitation phase: increasing personal capacity toward preinjury level of functioning
- Adaptation phase: finding a balance between personal capacity and environmental demand for those that do not reach full recovery to preinjury level of functioning
- Stable end situation: reached when no medical or other intervention is expected to improve the condition and no further major adaptations are likely to occur.

The average length of these phases depends highly on the type of injury and there may be overlap between phases. The working group looks at all postinjury phases, and aims to collect data on both the stable end situation of injury patients and the process toward this situation.

It is well known that injuries and their consequences not only affect the victim, but may have large health effects on other persons as well, such as parents and/or other primary caregivers.[19–21] The working group, however, restricts its work to the impact on health of the injury patient.

Literature Review

We conducted a PubMed search aiming to identify studies on injury-related disability published since 1995. As argued above, all postinjury phases and all levels of injury severity were included. Studies into the health effects on people other than the injury victim were excluded. We only searched for studies using generic health status measures (i.e., those that are applicable to all diseases and injuries). The information collected with generic measures allows straightforward comparisons with general population norms and with other diseases. Moreover, comparisons of different types of injury can be made and some of the measures generate a summary score that can be converted into a composite health outcome measure, such as the Disability Adjusted Life Year (DALY).[22] Composite health outcome measures combine fatal and nonfatal consequences of diseases and injury.

We used queries of the type Injury and SF-36, SF-12, EuroQol (EQ-5D), Health Utilities Index (HUI 2 and HUI3), Quality of Well Being Scale (QWB), Nottingham Health Profile (NHP), Sickness Impact Profile (SIP), Functional Independence Measure (FIM), and WHO Disability Assessment Schedule (WHODAS II). These are generic instruments, including health domains of potential relevance for large groups of injury patients. In some studies, injury-specific and/or domain-specific measures were added. We selected studies that looked into the consequences of the following injury categories: all injury combined, major trauma, traumatic brain injury, spinal cord injury, hip fractures, other fractures, superficial injuries and wounds, sprains/strains/dislocations, amputations, and poisoning. For the studies retrieved, we summarized the characteristics of the study population, the study design, the timing of the assessment(s), the instruments used to measure functional outcome and their discriminative power and responsiveness to change, and the main findings. The results of the literature review were used as input for the consensus procedure described below.

Consensus Procedure

The ECOSA working group on postinjury levels of functioning and disability was established during a worldwide expert meeting on Measuring the Burden of Injury.[23] The group is composed of experts from the fields of traumatology, injury epidemiology, health status measurement, and health economics. It has participants from the Netherlands, Denmark, and the United Kingdom. In 2002, a draft report containing the conceptual framework and working methods used was discussed during a first expert meeting and subsequently revised and released on the ECOSA Website (www.ecosa.org). By means of a newsletter, interested researchers and practitioners were invited to participate in the discussions. In 2003 to 2004, the literature review was conducted and, based on its results, draft guidelines for empirical researches were constructed. The guidelines aim to give practical advice on defining the patient population, selecting the measurement instrument(s), timing of follow-up, and selecting the method(s) of data acquisition. The draft guidelines were extensively debated during an electronic discussion among the group members. This led to revised guidelines, which were released on the ECOSA Website and provoked further comments by several experts. These were presented and discussed during a workshop on methods of injury research, which was organized within the framework of the 7th World Conference on Injury Prevention and Control in 2004. In this article, we

present the literature findings and guidelines for the following categories of patients: all injuries combined, major trauma, and hip fractures. Traumatic brain injury and spinal cord injury are not dealt with in this article because they have been studied rather frequently and other parties have already published guidelines on these patient groups.[24,25] Superficial injuries and wounds, sprains/strains/dislocations, amputations, and poisoning are not dealt with because of an almost complete lack of empirical data for these injury types. The heterogeneous group of "other fractures" has been looked at in several studies, but for most types of fractures the number of studies is still too small to be used as input for guideline development.

RESULTS

Literature Review

Population-based studies on injury-related disability are scarce (Table 1). We identified eight studies (being reported in 10 articles),[13,17,18,26–32] which all used different inclusion criteria for their study population (e.g., different age ranges, trauma center patients versus emergency department [ED]-treated patients), different generic measures (EQ5D, SF-36, QWB, FIM, SIP) for health status measurement, and different timings of assessment. Six of the studies used a longitudinal design with different periods and timings of follow up.[13,18,26–31] All studies looked at heterogeneous patient populations and included injuries of different levels of severity. High prevalences of health problems within and after the first year of injury were a common finding of the studies. Four studies made a comparison with general population norms[13,17,18,28] showing that, in the medium and long term, injury patients as a group were worse off than age- and sex-matched reference groups. One study used an injury-specific measure (FCI) in addition to generic measures and appeared more sensitive to the loss of cognitive functioning and hand-arm movement.[32] Longitudinal studies from the United States[27,28] and the United Kingdom[29,30] showed no further improvements after 12 months in populations of admitted adult trauma patients. We observed that in populations including moderate to low severity injury (ED-treated patients), different generic measures (EQ-5D, SF-36) were able to discriminate between the health status of injured patients and noninjured persons and between patients with different types of injuries. A problem related to moderate- to low-severity injury populations,[18,26] however, was the difficulty in acquiring acceptable response rates.

Studies on disability in the most severely injured patients are increasingly conducted. This development started in the 1990s, when several studies using self-designed questionnaires were executed in several countries.[33–37] However, before the new millennium, only one study using a generic instrument among a selected sample of severely injured patients could be identified.[38] But since then, this approach has increasingly been applied, as can be derived from Table 2. We identified nine studies[38–47] that all aimed at including "major trauma patients". Most of these studies, however, did not clearly describe their inclusion criteria in terms of Injury Severity Score (ISS) cutoff points.

Of the available generic instruments, SF-36/SF-12 has most often been applied among major trauma patients.[38,40–45] These applications have shown that this instrument enables discrimination between different subgroups and is able to pick up changes in health status between discharge and 12-month follow up.

Two studies have used EQ-5D, also showing discriminative power and the ability to identify changes in health status.[39,47] Five of the studies used a longitudinal design, but all with only one measurement after discharge.[38,39,42,43,46] Two studies made a comparison with general population norms[43,47] showing that, in the long term, major trauma patients were worse off than age- and sex-matched reference groups. A common finding of all studies was the high prevalence of disabilities found at follow-up, mostly at 12 or 24 months after discharge. One study used an injury-specific instrument (HASPOC) including a generic instrument (SF-12) and assessments by clinicians. It showed that HASPOC—contrary to SF-12—discriminates between polytrauma and single-injury patients.[45] Indications were found that, in major trauma, no further improvements occur after 24 months.[47]

We identified eight studies looking at the functional consequences of hip fractures with generic measures (Table 3).[15,16,48–53] The majority of these studies dealt with community-dwelling patients (65+ years) who were eligible for operation, and excluded persons with cognitive impairments. Only two studies collected follow-up information from a comprehensive sample of hip-fracture patients without prior exclusions.[16,53] Data from patients with cognitive impairments were successfully acquired with the help of proxy respondents. Three generic measures have been applied so far among hip-fracture patients: EQ-5D,[15,48] SF-36,[49–52] and NHP.[16,53] In addition, several disease-specific measures have been used. The Cummings hip scale is a functional status scale developed for patients with hip problems, addressing basic and instrumental activities of daily living.[54] The Lower Extremity Measure (LEM) is a modification for patients with hip fractures of a measure that was originally developed for patients undergoing preservation surgery for a tumor of an extremity.[51] The Osteoporosis Assessment Questionnaire (OPAQ) is an instrument designed to assess the health-related quality of life in all types of osteoporotic patients, including patients with hip fractures.[55] The literature also shows an interesting application of the Rehabilitation Activities Profile (RAP).[53] This is a 15-item instrument for assessing recovery in rehabilitation medicine, providing more detailed data on communication, mobility, personal care, and household and leisure activities than the aforementioned generic measures.

The generic and disease-specific measures used so far showed similar differences between subgroups. All these instruments were responsive to changes within the first 3 to 6 months after the injury but the disease-specific measures (and the RAP) showed somewhat greater effect sizes than the

Table 1 Available Empirical Studies (1995–2005) With Generic Measures Into Disability Among Comprehensive Injury Populations (All Injuries Combined)

Author, Year, Country (Reference Number)	Study Population	Measure	Design; Timing Follow-up (% Response)	Discriminative Power	Responsiveness to Change	Limitations	Main Findings
Kopjar 1996: Norway[26]	N = 775; treated at ED and/or admitted; age 16–78	SF-36	Longitudinal; 2 and 6 months; (61%, 38%)	Patients with activity restrictions have lower scores on all SF-36 dimensions than patients without such restrictions	Significant improvement between 2 and 6 months at most SF-36 dimensions	Health domains missing; no long-term follow up; low response rate at 6 months	At 2 months: 17% report activity restrictions
Holbrook et al. 1998, 1999; United States[27,28]	N = 1,048; admitted to trauma center; age 18+	QWB	Longitudinal; 6, 12, 18 months; (79%, 79%, 74%)	Effect on outcome depends on injury location; serious extremity injury is independent predictor	Significant improvement between 6 and 12 months; no further improvement between 12 and 18 months	Health domains missing; no short-term follow up; exclusion of neurotrauma	Summary score at 18 months (0.68) is below general population norms
Baldry Currens 1999, 2000; United Kingdom[29,30]	N = 330; admitted >3 day; age 5+	FIM	Longitudinal; 3, 6, 12, 24 months (80% at all moments)	Differences in disability prevalences between minor and major trauma and between different injured body regions are identified	Significant improvement between 3 and 6 and 6 and 12 months; no further improvement beyond 12 months	Health domains missing	At 12-month follow up, 16% problems at FIM motor score and 12.% problems at FIM cognition score
Michaels et al. 2002; United States[31]	N = 247; admitted to trauma center; age 18+; without severe neurotrauma	SF-36	Longitudinal; during admission, 6, 12 months (100%, 75%, 51%)	Patients with extremity fractures were more impaired in physical function, role-physical and bodily pain at 6 and 12 months than nonorthopedic injuries	Significant improvements between baseline and 6 months and 6 and 12 months	Health domains missing; no short-term follow up; exclusion of neurotrauma	All subscales of SF-36 remain below general population norms at 12 months
MacKenzie et al. 2002; United States[32]	N = 1,587; admitted to trauma center; age 18–59; blunt motor vehicle injury	SF-36, SIP, FCI	Longitudinal; 12 months (78%)	FCI is more sensitive than SIP and SF-36 to loss of cognitive functioning and hand/arm movement		Only one follow-up assessment; restriction to blunt motor vehicle injury	SF-36 physical health summary ranges from 36.2–50.2; SF-36 mental health summary ranges from 45.1–51.3

Table 1 Available Empirical Studies (1995–2005) With Generic Measures Into Disability Among Comprehensive Injury Populations (All Injuries Combined) (continued)

Author, Year, Country (Reference Number)	Study Population	Measure	Design; Timing Follow-up (% Response)	Discriminative Power	Responsiveness to Change	Limitations	Main Findings
Aitken et al. 2002; United States[13]	N = 310; admitted to academic children's hospital; age 3–18; AIS >1	CHQ,pediatric FIM	Longitudinal; discharge, 1, 6 months; (63%, 56%, 45%)	Differences in CHQ scores by level of injury severity are identified	Significant improvement between 1 and 6 months	Health domains missing; no long-term follow up; restriction to children with AIS >1	Mean physical and psychosocial summary scores remain significantly below general population norms
Inaba et al. 2003; Canada[17]	N = 171; admitted to trauma center; age 65+	SF-36	Cross-sectional; 1, 5–4, 5 years; (75%)			Health domains missing; cross-sectional design; one assessment per patient varying in time after injury; restriction to 65+	On 7 dimensions of the SF-36 the trauma patients remain below general population norms
Meerding et al 2004; Netherlands[18]	N = 4,639; treated at ED and/or admitted; age 15+	EQ-5D	Longitudinal; 2, 5, 9 months (39%, 24%, 12 %)	Differences in disability prevalence between different injured body regions are identified	Significant improvement between 2 and 5 months; no further improvement between 5 and 9 months	Health domains missing; no long-term follow up; low response rates	Summary score at 9 months (0.74) is below general population norms

Table 2 Available Empirical Studies (1995–2005) With Generic Measures into Disability Among Major Trauma Patients

Author, Year, Country (Reference Number)	Study Population	Measure	Design; Timing Follow-up (% Response)	Discriminative Power	Responsiveness to Change	Limitations	Main Findings
Brenneman et al. 1997; Canada[38]	N = 195; admitted and ISS >9	SF-36	Longitudinal discharge; 12 months (44%)	Better scores on 7 dimensions of the SF-36 for patients who returned to work	Significant improvement between discharge and 12 months; more improvement among those who returned to work	Health domains missing; limited to one follow-up assessment; ISS cutoff point 9; exclusion neurotrauma	Functional status at discharge predicts future employment status
Badia et al. 2001; Spain[39]	N = 69; admitted to the ICU	EQ-5D	Longitudinal; prior health status, 12 months (89%)	Differences in disability (changes) between diagnostic groups are identified	Major trauma patients have a significantly lower health status than prior to admission (estimated by proxies and patients)	Health domains missing; limited to one follow-up assessment; no ISS cutoff point; small sample size	EQ-5D summary score drops from 0.90 prior to admission to 0.75 at 12-month follow up
Pirente et al. 2001, 2002; Germany[40,41]	N = 56; admitted and "severely injured"	SF-36, GOS, EQ-5D, POLO-Chart	Case control; 12 months (77%)	Higher proportion of problems on all SF-36 dimensions among trauma patients than among controls		Limited to one follow-up assessment; no ISS cutoff point; small sample size	High prevalence of problems (30% to 51%) on all SF-36 dimensions at 12-month follow up
Stalp et al. 2001; Germany[42]	N = 150; admitted and ISS >9	SF-12, FIM, GOS, MFA	Longitudinal; 24 months	Differences in disability prevalences by injured body region are found		Limited to one follow-up assessment; ISS cutoff point 9; no response rates provided	SF-12 indicated an outcome of more than satisfactory in 63% of cases at 24-month follow up
MacKenzie et al. 2002; United States[43]	N = 1,587; admitted >72 hours or to ICU; age 18–59; blunt motor vehicle injury	SF-36 + cognitive function scale	Longitudinal; 12 months (78%)	Cognitive function scale discriminates well between patients with and without brain injury		Health domains missing; limited to one follow-up assessment; no ISS cutoff point; restriction to blunt motor vehicle injury	Mean SF-36 score in all dimensions except vitality and mental health are below general population norms
Stalp et al. 2002; Germany[44]	N = 312; admitted and ISS >15	SF-12, FIM, GOS, MFA, HASPOC	Longitudinal; 24 months (81%)	Differences in disability prevalences by injured body region are identified; highest prevalences found for lower extremity injury		Limited to one follow-up assessment	General outcome (SF-12): grade I/II (very good/good) 34%, grade III/IV (satisfactory/sufficient) 54%, grade V/VI (deficient/bad) 12%

Table 2 Available Empirical Studies (1995–2005) With Generic Measures into Disability Among Major Trauma Patients (continued)

Author, Year, Country (Reference Number)	Study Population	Measure	Design; Timing Follow-up (% Response)	Discriminative Power	Responsiveness to Change	Limitations	Main Findings
Zelle et al. 2003; Germany[45]	N = 170; admitted and "polytrauma"	SF-12, HASPOC	Case control; 24 months	HASPOC; contrary to SF-12-discriminates between polytrauma and single injuries		Limited to one follow-up assessment; no ISS cutoff point; no response rates provided	Hannover Score for Polytrauma Outcome (HASPOC) outperforms generic instrument (SF-12)
Dimopolou et al. 2004[46]	N = 117; multiple trauma admitted to ICU	NHP, GOS, Rosser Disability Scale	Longitudinal; 12 months (74%)	Severe head injury has a significantly worse outcome than other injury		Health domains missing; limited to one follow-up assessment; no ISS cutoff point	90% of patients has a problem in at least one of the six domains related to subjective health status
Vles et al. 2005; the Netherlands[47]	N = 196; admitted and ISS >15	EQ-5D + cognitive function scale, GOS	Cross-sectional; 1–6 years (85%)	Differences in disability prevalences by injured body region and injury severity are identified		Health domains missing; cross-sectional design; one assessment per patient varying in time after injury	Summary score at long-term follow up (0.76) is below general population norms; high prevalence of disabilities in all EQ-5D domains

generic measures. The vast majority of studies used a longitudinal design, but all with different assessment moments. Prefracture levels of functioning have been assessed in five studies,[15,49,51–53] based on recall (patients or proxies) within the first week of the hospital admission. The results obtained to date seem to indicate that the recovery of hip fracture patients takes place within the first 4 months after the event. Afterward, however, their health status seems to remain far below prefracture levels and general population norms. One study found that prefracture levels were already below levels of age- and sex-matched controls.[52]

Guidelines

Based on the findings from the literature, we developed guidelines for conducting follow-up studies among trauma patients. First of all, a set of general guidelines was agreed upon that address the definition of the patient population, choice of the measures, timing of the assessments, and data collection procedures. Application of the general guidelines is advised for all types of injury. In extension to the general guidelines, additional measures and/or moments of follow up may be used to capture the consequences of specific types of injury in more detail. To illustrate this strategy (i.e., the use of a common core supplied with additional measures/moments), some specifications will be presented for patients with major trauma and hip fractures. The general guidelines are described below and summarized in Table 4.

Definition of the Patient Population

In general, a clear definition of the patient population should be the starting point of any future epidemiologic study into injury-related disability. This seems obvious, but in the available literature patient populations were often ill defined and/or not clearly reported. This hampers the linkage of disability information from specific studies to population-based injury surveillance data. The primary criteria for including patients are injury type and injury severity. Internationally accepted definitions of injury, specific injury types, and levels of injury severity should be used. Prior exclusions based on characteristics not related to the injury will produce biased results. In the literature, very often mental (e.g., patients with cognitive impairments) and/or social (e.g., patients living in institutions) exclusion criteria were used. Although understandable from the point of view of study logistics, complexity, and resources, this leads to selective patient samples not representative for the impact of injury at the population level. To include patients with mental and/or social problems, we recommend the use of standardized proxy assessment. The exclusion of patients based on their age should in principle be avoided too, unless a study specifically looks at the functional consequences of injury among a predefined age group. For major trauma, the patient selection should be based on the internationally accepted definition of major trauma patients. Currently, the most widely used definition of major trauma includes multiple or single (e.g., isolated head injury with

Table 3 Available Empirical Studies (1985–1995) With Generic Measures Into Disability Among Hip Fracture Patients

Author, Year, Country (Reference Number)	Study Population	Measure	Design; Timing Follow-up (% Response)	Discriminative Power	Responsiveness to Change	Limitations	Main Findings
Tidermark et al. 2002; Sweden[15]	N = 72; 65+; community dwelling; "internal fixation"	EQ-5D + clinical follow up	Longitudinal; 12–48 hours, 1 week, 4 months, 17 months (93%)	Large differences in EQ-5D summary score at different moments of follow-up between patients with and without complications	Significant increase in EQ-5D summary score between 0 and 4 months, followed by no further improvement afterwards (and significant decline in patients with complications)	Exclusions on mental and social criteria; small sample size	EQ-5D summary score decreased from 0.78 before fracture (based on recall) to 0.59 at 4 months and 0.51 at 17 months (remaining far below general population norms of 0.80)
Tidermark et al. 2003; Sweden[48]	N = 102; 65 + community dwelling; operated	EQ-5D + clinical follow-up	RCT; internal fixation versus total hip replacement; 4, 12, 24 months	Reduction in EQ-5D summary score significantly lower in group with total hip replacement		Exclusions on mental and social criteria; small sample size: no response rates provided	Internal fixation leads to more complications (36% versus 4%) than total hip replacement and worse functional outcome
Peterson et al. 2002; United States[49]	N = 38; 65+; community dwelling; operated <48 hours	SF-36 + Cummings hip scale + clinical follow-up	Control patients within an RCT; 2–5 days, 2, 6, 7, and 12 months		Most SF-36 dimensions and the Cummings scale showed significant improvement between 0 and 6, months, but no further improvements afterwards	Exclusions on mental and social criteria; small sample size: no response rates provided	Most SF-36 dimensions and the Cummings scale had decreased at 12 months in comparison to the prefracture status (based on recall) and remained below general population norms
Tosteson et al. 2001; United States[50]	N = 67	SF-36 + time tradeoff	Cross-sectional; 1–5 years postinjury; comparison between groups (no fracture, vertebral, hip)	Significant differences in SF physical component score between groups; lowest score for patients with hip and vertebral fracture		Exclusions on mental criteria; small sample size; cross-sectional design; one assessment per patient varying in time after injury; no response rates provided	Hip fracture patients had significantly lower QALYs (0.63) compared with vertebral fractures (0.82) and no fractures (0.91)

Table 3 Available Empirical Studies (1985–1995) With Generic Measures Into Disability Among Hip Fracture Patients (continued)

Author, Year, Country (Reference Number)	Study Population	Measure	Design; Timing Follow-up (% Response)	Discriminative Power	Responsiveness to Change	Limitations	Main Findings
Jaglal et al. 2000; Canada[51]	N = 43, community-dwelling	SF-36 + LEM	Longitudinal, 0,6 weeks, 6 months		Both the SF-36 physical component score and the LEM significantly improved between all the time moments; effect sizes of LEM were greater than of SF-36	Exclusions on mental and social criteria; small sample size; no response rates provided; no long-term follow up	Most SF-36 dimensions decreased at 6 months in comparison to the prefracture status (based on recall) and remained below general population norms
Randell et al. 2000; Australia[52]	N = 32; 65+, low trauma fracture; patients	SF-36 + OPAQ-2	Longitudinal, <1 week, 12 weeks	At baseline, mean scores reported by fracture patients were lower than controls for all domains in both the SF-36 and OPAQ-2	Among hip fracture patients there was a significant reduction in both physical and social component of SF-36 between 0 and 12 weeks; among controls there was not	Exclusions on mental criteria; small sample size; no response rates provided; no long-term follow up	Hip fracture patients had lower baseline scores and experienced a significant deterioration on both SF-36 and OPAQ-2
Van Balen et al. 2001; the Netherlands[16]	N = 102, 65+, low trauma fracture patients; operated	NHP	Longitudinal, 1 week, 1 month, 4 months (72%)		Significant improvement of all NHP dimensions between 1 week and 4 months	Small sample size; no long-term follow-up	Most NHP dimensions (physical mobility, pain, social isolation, and emotional problems) remained below general population norms at 4 months
Van Balen et al. 2003; the Netherlands[53]	N = 254, 65+, low trauma fracture patients; operated	NHP + RAP + BI	Longitudinal, 1 week, 1 month, 4 months (72%)		Significant improvement of all NHP dimensions and RAP and BI between 1 week and 4 months: greater effect sizes of RAP and BI	No long-term follow up	RAP and BI scores at 4 months remained below prefracture levels; most NHP dimensions at 4 months remained below general population norms

Table 4 Guidelines for the Conduction of Empirical Follow-up Studies Measuring Injury-Related Disability

Definition of patient population
Use international definitions of injury (severity) to include patients
Make no prior exclusions based on mental or social criteria
Use standardized proxy assessment to include patients with mental and/or social problems
Choice of measures
Use combination of EQ-5D and HUI3 as common core of measures (all studies)
Extend the common core of measures with an injury-specific instrument (specific subgroups)
Timing of assessments
Use a longitudinal design with multiple assessments over time
Use assessments at 1, 2, 4, and 12 months postinjury as common core of timings (all studies)
Extend the common core of timings with (an) extra assessment moment(s) (specific subgroups)
Make a retrospective assessment of the preinjury health state
Data collection
Collect data on determinants, using international definitions and classifications
Develop and use protocols for the collection of data among patients not being able to give self-reports
Install response raising measures

Abbreviated Injury Score [AIS] of 4 or 5) trauma with an Injury Severity Score >15. ISS is preferred above AIS because it provides opportunities to make comparisons within groups containing patients with multiple injuries as well as patients with single injuries. If AIS/ISS is not directly available, the International Classification of Diseases TO Abbreviated Injury Scale (ICDTOAIS) conversion program to derive severity scores may be used.[56] A possible alternative to classify and select patients by level of severity is the New Injury Severity Score (NISS).[57] Recent studies have found that NISS outperforms ISS in predicting mortality[58–61] and disability[62] of specific patient groups. NISS and ISS show similar distributions among patients.[63] In the future, NISS could possibly replace ISS as the preferred injury severity scoring system, but the scientific community has not yet reached a conclusion on this issue. Our group therefore advises to follow current standard practice, such as using ISS cutoff points for patient selection purposes until the scientific debate might decide otherwise. Until then, in conducting outcome analyses, both ISS and NISS should be used and compared. In conducting outcome analyses, physiologic scoring systems (e.g., RTS) should be included as well.[64] These systems, however, are less suited for the selection of major trauma patients because they cannot be assessed in intubated patients and are not universally used and recorded. If direct or indirect severity scoring systems cannot be applied, the group of trauma patients admitted to the intensive care unit could be followed up as an alternative, but variations in health care (e.g., bed availability) will limit comparability. For patients with hip fractures, it is advised to select all proximal femoral fractures. If specific selections of injuries are studied, these should be reported clearly.

Choice of the Measures

The working group has made an attempt to define a "common core of measures" to be used in all studies into injury-related disability. For this purpose, seven generic measures (EQ-5D, HUI3, SF-36, SIP, QWB, HUI3, WHODAS II) and one injury specific measure (FCI) were judged according to a list of assessment criteria. The results of this assessment are summarized in Table 5.

As a first criterion, we defined that all body functions, and activities and participation (A&P) domains of the ICF, that are relevant for a substantial part of injury patients should be included in the common core. According to the working group the common core of measures should at least include the following ICF domains: Cognition, Emotion, Pain, Problem solving, Ambulation, Use of hand/arm/fingers, Self care, Household activities, Interpersonal interactions (including sexual activities), School and/or work, and Recreation. We found that none of the measures studied contains all these relevant health domains. For example, the only measure containing both cognition and emotion is the HUI3. EQ-5D, SF-36, SIP, and QWB include emotion, but miss the cognitive dimension. FCI, FIM, and WHODASII on the other hand include cognition, but miss emotion. The use of hand/arm/fingers, which seems very relevant for a substantial part of injury patients, is only included in the FCI and the HUI3. But these two instruments provide no information at all on the social consequences of injury (e.g., usual activities and social interaction) that is well covered in all other measures. This shows that to cover all relevant health domains of the ICF at least two measures should be combined: a measure focusing on functional capacities of the patient on the one hand (HUI or FCI) and a measure including social participation on the other hand (all other measures).

In addition to the relevance of health domains, the working group used several other assessment criteria. The common core measures should be applicable to different injury types and severity levels, which should be documented in the literature. They should be applicable to the widest range of age groups, and to other health problems as well. They should provide a link to utility measures to calculate summary measures of population health. They should be suitable for self-assessment by questionnaire, be available in different languages, and, last but not least, meet the criterion of brevity and simplicity.

After judging all measures according to these criteria, the working group recommends the use of the EQ-5D in combination with HUI3 as the preferred common core of measures in all studies.

EQ-5D and HUI3 are complementary with respect to the ICF health domains included. A questionnaire combining the two measures will need only 10 to 15 minutes completion time.

HUI3 is the preferred measure to assess functional capacities after injury and should be used in all studies. The

Table 5 Assessment of Appropriateness Health Status Measures for Application to Injury Patients

Assessment Criteria	EQ-5D	FCI	FIM	HUI3	SF-36	SIP	QWB	WHODASII
1. Inclusion of relevant health domains	Anxiety/depression, pain/discomfort, ambulation, self care, usual activities (household, school, work)	Cognition, vision, hearing, speech, bowel control, bladder control, sexual function, problem solving, communication, ambulation, bending/lifting, hand-arm movement	Cognition, bowel control, bladder control, problem solving, communication, ambulation, eating, grooming, bathing, dressing, toileting, social interaction	Cognition, emotion, vision, hearing, speech, pain, problem solving, communication, ambulation, dexterity (use of hands and fingers)	Emotion, pain, ambulation, bending/lifting, bathing, dressing, regular daily activities (household, school, work), social activities	Alertness, emotion, communication, ambulation, eating, self care, household management, work, social interaction	Emotion, vision, hearing, speech, pain, bowel problems, bladder control, urinary problems, paralysis/stiffness extremities, ambulation, bending/lifting, self care, usual activities (household, school, work, sports)	Cognition, problem solving, communication, ambulation, eating, bathing, dressing, household management, social interaction, sexual activities, school, work
2. Documented applicability to injuries of different types and severity levels	Documented applicability to comprehensive injury populations, including minor injury, major trauma and hip fractures	Documented applicability to hospital-admitted patients with blunt motor vehicle injury and major trauma applicability to low energy trauma has yet to be tested	Documented applicability to hospital admitted patients; applicability to low-energy trauma has yet to be tested	Documented applicability to osteoporotic fractures applicability to other trauma populations has yet to be tested	Documented applicability to comprehensive injury populations, including minor injury, hospital-admitted patients with blunt motor vehicle injury, major trauma and hip fractures	Documented applicability to hospital-admitted patients with blunt motor vehicle injury and major trauma; applicability to low-energy trauma has yet to be tested	Documented applicability to hospital-admitted patients applicability to low-energy trauma has yet to be tested	Applicability to injury populations has yet to be tested
3. Applicability to other health problems	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
4. Applicability to the widest possible age ranges	Applicable to children (5+), adolescents, adults and the elderly; for children 0–4 some adaptations needed	Designed for adult populations; pediatric version for children	Applicable to children (5+), adolescents, adults and the elderly; for children 0–4 pediatric version (WEEFIM)	Applicable to children (5+), adolescents, adults, and the elderly	Applicable to adolescents, adults, and the elderly	Applicable to adolescents, adults, and the elderly	Applicable to adolescents, adults, and the elderly	Applicable to adolescents, adults, and the elderly
5. Link to utility scores	Yes; based on health state valuations by the general UK population	Yes; based on ratings from an US expert panel (n = 114); validity has to be tested	No	Yes; based on valuations by a representative sample of the general US population	Yes; recently developed and labeled SF-6D; validity not yet tested on a large scale	No	Yes; based on valuations by a representative sample of the general US population	No
6. Suitability for self-assessment	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes

Table 5 Assessment of Appropriateness Health Status Measures for Application to Injury Patients (continued)

Assessment Criteria	EQ-5D	FCI	FIM	HUI3	SF-36	SIP	QWB	WHODASII
7. Availability in different languages*	Catalan, Croatian, Czech, Danish, Dutch, Finnish, French, German, Hungarian, Italian, Norwegian, Polish, Portuguese, Spanish, Swedish, Turkish, UK English	English	English	Dutch, English, French, German, Italian, Portuguese, Spanish, Swedish	Bulgarian, Croatian, Czech, Danish, Dutch, Finnish, French, German, Greek–Hungarian, Iceland, Italian, Latvian, Lithuanian, Norwegian, Polish, Portuguese, Romanian, Russian, Spanish, Swedish, Turkish, UK English, UK Welsh	English	Danish, Dutch, Finnish, French, German, Italian, Norwegian, Portuguese, Russian, Spanish, Swedish, UK English	Dutch, English, French, German, Greek, Italian, Romanian, Russian, Spanish, Turkish
8. Brevity	5 items + VAS; 2–4 minutes completion time	10 items; 8–10 minutes completion time	18 items; 10–15 minutes completion time	8 items; 8–10 minutes completion time	36 items; 5–10 minutes completion time; 12 items; 2–4 minutes (SF-12)	136 items; 30 minutes completion time	69 items; 15–20 minutes completion time	36 items; 5–10 completion time; 12 items; 2–4 minutes completion time

*Inventary in 2005.

applicability of HUI3 has been shown already in patients with low-energy trauma.[65] This measure can be used in a wide age range including children from 5 years old onwards[66] and the (very) elderly.[65] It has a direct link to utility scores based on valuations by a general population sample in North America.[67] The measure is suitable for self-assessment, available in eight different languages, and needs only 8 to 10 minutes completion time.

EQ-5D is the preferred measure to assess complementary health domains not well captured by HUI3 (including social participation). For several reasons, we have rated EQ-5D higher than SF-36, QWB, SIP, FIM, and WHODASII. Recent studies have documented that EQ-5D is applicable to comprehensive injury populations, including injury of low to moderate severity,[14,18] major trauma,[39,47] and hip fractures.[15,48] The measure is applicable to a wide age range, including children from at least 5 years old onwards[14,68] and the (very) elderly.[15,48,69] It has a direct link to utility scores, based on health state valuations by the general UK population.[70] The measure is suitable for self-assessment, available in 17 different languages, and needs only 2 to 4 minutes completion time.

SF-36 is the second best alternative to be combined with HUI3. SF-36 is also applicable to a wide range of injury populations, including injury of low to moderate severity,[26] blunt motor vehicle injury,[43] major trauma,[38,40–42] and hip fractures.[49–52] The measure is suitable for self-assessment, available in 23 different languages and needs only 5 to 10 minutes completion time (with 2 to 4 minutes completion time for the SF-12, which also has been applied successfully in injury populations).[44,45] However, there are some disadvantages of SF-36/SF-12 in comparison to the EQ-5D. Although a link of SF-36 to utility scores has been developed and labeled SF-6D,[71] the validity of this approach has not yet been tested on a large scale. There is a problem with valuations of patients in severe health conditions, which are systematically underestimated.[72] This would lead to underestimation of health problems of injury patients with higher severity levels.

QWB is a third possible alternative to be combined with HUI3, but is rated lower than EQ-5D and SF-36 because it has yet to be tested in low-energy trauma and because of the length of the questionnaire.

In principle, because the ICF health domains are complementary, EQ-5D could be combined with FCI. However, combination with HUI is preferred for several reasons. Contrary to HUI, FCI is not applicable to other health problems and the utility score of this measure, based on a relatively small US expert panel (n = 114), has yet to be validated. Moreover, FCI has not yet been translated from English into other languages.

The other measures that were assessed have several shortcomings. FIM and SIP have yet to be tested in low-energy trauma, have no link to utility scores, and have several practical disadvantages, such as the need of specific interviewer skills (FIM) or length of the questionnaires (SIP). WHODASII holds promising features, but is not recom-

mended yet because too little is still known about its validity. Moreover, a link to utility scores is missing.

In studies focusing on specific groups of injury patients, the recommended common core of measures may be extended with an injury-specific (or disease-specific) measure, additionally addressing problems that are frequently occurring in that specific injury population. The common core of measures (EQ-5D and HUI3) is appropriate for studies focusing on major trauma. However, because many patients with major trauma suffer from head injuries, additional use of the Glasgow Outcome Scale (GOS) is advised. This will allow comparisons of studies on major trauma with specific studies on traumatic brain injuries. For patients with hip fractures, the common core of measures may be extended with an appropriate specific measure, such as the Cummings Hip Scale,[54] the Lower Extremity Measure,[51] and the Osteoporosis Assessment Questionnaire.[55]

Timing of the Assessments

The working group recommends conducting longitudinal studies with multiple assessments over time. In this way, both recovery patterns and permanent consequences of injury can be assessed.

Each measurement moment should be representative for one of the four phases of trauma recovery: acute treatment phase, rehabilitation phase, adaptation phase, and stable end situation. Standardization of the timings of the assessments is of equal importance to standardization of the measures used. Therefore, the working group has developed a "common core" of timings as well. The working group is aware that, because of practical and/or financial restrictions, it will often be impossible to do a large series of measurements. All studies, however, should measure at one or more of the assessment moments advised below. If possible, a first follow-up measurement is advised at 1 month (acute treatment phase, range 0–8 weeks). The first 2 months after injury reflect the period of major health effects for injury patients of all severity levels, as shown by several studies.[13–15,18,26,49] Second and third follow-up measurements are advised at 2 months (rehabilitation phase, range 1–3 months) and 4 months (adaptation phase, range 3–6 months) respectively. The literature shows that for the majority of patients, recovery is mainly occurring within the first half year after injury.[13–15,18,26–30,49] With measurements at 2 and 4 months, this period of recovery is well covered. Finally, it is advised to do a measurement of the stable end situation at 12 months (range 6–24 months). Studies in the United States and United Kingdom have shown that comprehensive injury populations show no further recovery after 12 months.[28,30] The working group further recommends the retrospective assessment of preinjury health status within the first week after the injury, as implemented in several follow-up studies of patients with hip fractures.[15,49,51–53]

Longitudinal studies with multiple measurement moments to study recovery patterns of major trauma patients are a priority issue because most outcome studies on major trauma patients have to date been restricted to a single moment of follow-up. The common practice in this field to define moments of follow-up based on the time since hospital discharge should be abandoned. Major trauma patients may show further improvement (in particular in social participation) after 12 months.[47] Therefore, in studies focusing on major trauma, it is recommended to add an extra measurement at 24 months to the common core of timings.

Data Collection

We recommend collecting data on the largest sample of patients feasible. Comprehensive injury populations need sample sizes of at least 1,000 patients and are preferably stratified with over-representation of the more severe injuries. For injury-specific studies, samples of at least 100 patients are recommended. We recommend collecting additional data on possible determinants of disabilities. The following variables are important: age, sex, socioeconomic status, comorbidity (physical and psychological), complications, and social support. For this purpose, internationally accepted definitions, classifications, and measurements should be used. Comorbidity, for example, can be derived from existing classifications.[73] The partitioning of disability to different competing conditions is very difficult because of the variable severity of comorbidities. One approach is to measure the presence of significant preinjury comorbidities by including a question in the baseline data collection such as "before your injury did you suffer from a disability or long-term health problem that limited your normal activities?" Analysis of changes in status for those with and without preexisting disability will help measurement of the injury-related component. Questions on the use of health services for reasons other than the injury on a before and after basis are also helpful and have been included in the UK Burden of Injuries study that was recently initiated. Protocols should be developed for the collection of data among specific patient groups. Several subgroups of injury patients will not be able to give self-reports on their health condition. Major examples are patients with severe head injuries, patients on mechanical ventilation, and patients with cognitive impairments, young children, and psychiatric patients. For these patient groups, the protocols should allow the collection of data with the help of proxy respondents (parents, partners, or other caregivers of the patient). We recommend installing response-raising measures when conducting a follow-up study into injury-related disability.[74] Without specific measures the risk of low response rates is high, in particular when injuries of low severity are included.[18,26] We recommend sending reminders to all nonrespondents at all moments of follow up. It is further advised to make one person responsible for the collection of follow-up data, such as a medical resident with interest in scientific research. A prerequisite to be met is that ethical rules about informed consent should be applied.

DISCUSSION

Based on a review of the literature and a consensus procedure, we have developed guidelines for the conduction of empirical follow-up studies into injury-related disability. We recommend use of EQ-5D and HUI3 as the common core of measures in all studies. We advise to conduct longitudinal studies with multiple assessments over time. In this way, both recovery patterns and permanent consequences of injury can be assessed. As the common core of timings, a series of measurements at 1, 2, 4, and 12 months is recommended. In extension to the general guidelines, additional measures and/or moments of follow up may be used to capture the consequences of specific types of injury. For studies focusing on major trauma, for example, we recommend an additional follow-up measurement at 24 months postinjury.

Our group is the first to develop guidelines for measuring injury-related disability among the complete and heterogeneous scope of injury patients with all levels of severity. Previously, guidelines have been developed by an international group of experts (the Cologne group 2002) for specific patient groups, including patients with major trauma.[75,76] Similar to our group, EQ-5D was one of the advised measures for major trauma patients by the Cologne group as well to permit economic analyses. In addition, however, they have recommended using SF-36 as a validated global quality-of-life measure. We have proposed HUI3 as additional measure instead because of the inclusion of some important health domains that are missing in both EQ-5D and SF-36: cognition and hand-arm movement. EQ-5D is preferred above SF-36 because the latter measure has been shown to underestimate the consequences of severe health conditions,[72] which makes the measure less suitable for injuries of higher severity levels. For similar reasons, the Prevention of Falls Network Europe has recommended to include EQ-5D instead of SF-36 in a common outcome data set for fall injury prevention trials.[77] Moreover, EQ-5D can be extended with a question on cognitive functioning,[78] as applied already in injury patients.[14,18,47]

The literature has shown variation in utility scores by health status measure,[79] including variation between EQ-5D and HUI3.[80] Comparative analyses on EQ-5D versus HUI3 among injury patients have not yet been conducted. Calculation of both utility scores is therefore advised, providing a range of uncertainty when quantifying the impact of injury on population health.

Similar to the Cologne group, we recommend measuring the preinjury health status retrospectively and measuring the long-term consequences at 12 and 24 months postinjury. A difference between our guidelines and the Cologne group in the recommended moments of follow-up concerns the first months after injury, where we have advised a larger series of assessments (at 1, 2, and 4 months contrary to one measurement at 3 months by the Cologne group). We recognize that assessments of major trauma patients during this phase will often be very difficult. Nevertheless, if feasible, these measurements should be conducted to have data for fully quantifying injury-related disability and complete estimations of the health benefits of prevention and trauma care.

A common core of measures and assessment moments is highly needed to obtain improved and more consistent scientific knowledge on injury-related disability, and is therefore recommended. However, it is recognized that using a common core of measures and assessment moments may have some disadvantages for specific groups within the widely varying population of injury patients. Owing to a ceiling effect, EQ-5D and—to a lesser extent—HUI3 are not the most appropriate choices if the main goal of a study is to distinguish between minor levels of impairment.[79] But the relevance of this issue in quantifying injury consequences can be questioned. International research on the most appropriate injury indicators has already advised restrictions to (various types of) injury at the medium to higher end of the severity spectrum.[81–83]

In studies focusing on specific types of injury, as a general rule, additional measures may be selected in addition to the common core. All health domains that are relevant for a specific patient population should in principle be included in the set of measures that is used. If not, incomplete disability information underestimating the impact of injury will be collected. Relevant health domains may be identified by relating the ICF to specific expert knowledge from the injury field. This procedure has, for example, been applied for the selection of measures to assess functional outcome after burns.[84] In this field, the common core of measures may be extended with burn-specific measures to capture specific consequences, such as esthetic aspects and interpersonal relationships.[85] Moreover, in this field, measures have been developed and validated that are well tailored to the specific developmental stages of children of different ages.[86–88]

The common core of measures seems appropriate for children ages 5 years and over. EQ-5D has previously been tested among children 5 years and older,[65] and a recent study has shown good discriminative power and responsiveness to change among injured children in this age group.[14] The Health Utilities Index has a pediatric version for children aged 5 and older, which has been tested and validated.[69] However, several aspects related to measuring injury-related disability among children were not extensively dealt with by our working group and are still open to debate. First of all, other measures that could be superior to the EQ-5D and HUI3, such as the Child Health Questionnaire (CHQ),[89] are available for children aged 5 and older. But the advantages and disadvantages of the different measures for injured children are not well established because a comparative study within the injury field has still to be conducted. Moreover, the question remains how to assess injury-related disability among very young children (0–4 years). For infants aged 1 through 4 years, the Infant Toddler Quality of Life Questionnaire (CHQ-IT) has been developed.[90] This measure was

recently tested in a small sample (n = 31 at 1 month, n = 15 at 6 months after injury) of injured toddlers within the framework of a study on functional outcome after pediatric trauma of moderate to high severity (ISS >8).[91] This type of study, combined with the use of other measures, such an adapted version of EQ-5D, should be repeated in larger samples of injured preschool children. Another issue often discussed in the health outcomes literature concerns the most appropriate respondent when assessing children's health.[92] Functional outcomes may differ after completion of a health status measure by child or parent.[93] Studies among injured children have shown that parental reports may overestimate the child's functioning, especially when assessing the physical functioning. The parents may realize that the injury could have resulted in a worse outcome and that, with respect to the initial injury, their child is doing quite well.[14,94] Although the exact values between child and adult might be different, previous research suggests that at least the ranking order will be largely the same.[65] It has been argued that it is at least possible to obtain child self reports in a school-aged population (12–18 years).[95]

At all ages, the set of measures used should in principle be extended with clinical data (assessments by health care workers). In the literature, disability information from injury patients is often fully based on patient self reports with the help of written questionnaires or personal interviews. The clinical relevance of the self-reported information can better be assessed in a setting where clinical follow ups also take place. Studies combining the collection of patient self reports with clinical follow-up data should become the standard. In Germany, comprehensive measures consisting of both self reports and predefined clinical assessments and procedures have been developed and applied among patients with major trauma.[44,45] Their list of predefined clinical assessments and procedures seems promising in addition to the common core of measures (i.e., EQ-5D and HUI3) proposed by our working group.

The proposed common core of timings will provide data on recovery patterns of trauma patients and on the temporary and permanent consequences of injury. We provided a framework, distinguishing four phases after injury: acute treatment, rehabilitation, adaptation, and stable end situation. It is possible, however, that injury may increase the risk of specific other diseases after one or more decades. In this situation, there will be no stable end situation but accelerated degeneration instead. Professional football players, for example, have increased risks of developing osteoarthrosis many years after their career,[96] which could be a result of injury-related disability. Empirical follow-up studies as proposed by the working group will end at 12 to 24 months and therefore not provide data on the "very long-term consequences" of injury. To capture this dimension of the burden of injury, other study designs (e.g., case-control studies and longitudinal linkage of databases) are recommended.

The working group has developed guidelines to improve empirical data collections on injury-related disability in terms of consistency, completeness, and comparability. Consistent international empirical data will allow valid burden of injury calculations, valid cost-effectiveness analyses of injury prevention and trauma care, and valid assessments of quality of care with composite health outcome measures. A necessary next step will be to test the guidelines in follow-up studies of large cohorts of patients with trauma of different severity levels in different countries. In addition, international consensus procedures as described in this article should be continued and extended. To date, a limited number of experts from a small number of countries have participated in the ECOSA working group, and a limited number of experts in the relevant clinical specialties in particular. For practical and financial reasons, all group members were participants of a European Union funded project to standardize calculation methods of medical costs of injuries.[97] This provided opportunities to arrange group meetings without additional costs, but prevented invitation of other international experts The discussions of our group and of related initiatives (e.g., the Cologne group 2002 and expert groups in North America) should therefore be jointly continued in the future, and we will seek funding for this aim. A further discussion on the preferred common core of measures and assessment moments in a broad international forum with input from methodological and clinical experts from all over the world is essential. We hope that our guidelines will be tested, provoke further international discussions, and will finally lead to broad international consensus on this major research topic.

REFERENCES

1. World Health Organization. International Classification of Functioning, Disabilities and Health (ICF). Geneva: World Health Organization; 2001.
2. Segui-Gomez M, MacKenzie EJ. Measuring the public health impact of injuries. *Epidemiol Rev.* 2003;25:3–19.
3. Beeck EF van, Looman CWN, Mackenbach JP. Mortality due to unintentional injuries in the Netherlands, 1950–1995. *Public Health Rep.* 1998;113:427–439.
4. Barker M, Power C, Roberts I. Injuries and the risk of disability to teenagers and young adults. *Arch Dis Child.* 1996;75:156–158.
5. Bouillon B, Neugebauer E. Outcome after polytrauma. *Langenbeck's Arch Surg.* 1998;383:228–234.
6. Masson F, Vecsey J, Salmi LR, et al. Disability and handicap 5 years after head injury: a population-based study. *J Clin Epidemiol.* 1997;50:595–601.
7. Thornhill S, Teasdale GM, Murray GD, et al. Disability in young people and adults one year after head injury: prospective cohort study. *BMJ.* 2000;320:1631–1635.
8. Goldstein FC, Levin HS. Cognitive outcome after mild and moderate traumatic brain injury in older adults. *J Clin Exp Neuropsychol.* 2001;23:739–753.
9. Hellawell DJ, Taylor RT, Pentland B. Cognitive and psychosocial outcome following moderate or severe traumatic brain injury. *Brain Inj.* 1999;13:489–504.
10. Hallin P, Sullivan M, Kreuter M. Spinal cord injury and quality of life measures: a review of instrument psychometric quality. *Spinal Cord.* 2000;38:509–523.
11. Gittler MS, McKinley WO, Stiens SA, et al. Spinal cord medicine. 3. Rehabilitation outcomes. *Arch Phys Med Rehabil.* 2002;83;S65–71.

12. Sheridan RL, Hinson MI, Liang MH, et al. Long-term outcome of children surviving massive burns. *JAMA.* 2000;283:69–74.
13. Aitken ME, Tilford JM, Barret KW, et al. Health status of children after admission for injury. *Pediatrics.* 2002;110:337–342.
14. Polinder S, Meerding WJ, Toet H, et al. Prevalence and prognostic factors of disability after childhood injury. *Pediatrics.* 2005; 116:810–817.
15. Tidermark J, Zethraeus N, Svensson O, et al. Femoral neck fractures in the elderly: functional outcome and quality of life according to EuroQol. *Qual Lif Res.* 2002;11:73–481.
16. Balen R van, Steyerberg EW, Polder JJ, et al. Hip fracture in elderly patients: outcomes for function, quality of life and type of residence. *Clin Orthop.* 2001;390:232–243.
17. Inaba K, Goecke M, Sharkey P, et al. Long-term outcomes after injury in the elderly. *J Trauma.* 2003;54:486–491.
18. Meerding WJ, Looman CWN, Essink-Bot ML, et al. Distribution and determinants of health and work status in a comprehensive population of injury patients. *J Trauma.* 2004;56:150–161.
19. Unalan H, Gencosmanoglu B, Akgun K, et al. Quality of life of primary caregivers of spinal cord injury survivors living in the community: controlled study with short form-36 questionnaire. *Spinal Cord.* 2001;39:318–322.
20. Marsh N, Kersel D, Havill J, et al. Caregiver burden during the year following severe traumatic injury. *J Clin Exp Neuropsychol.* 2002; 24:434–447.
21. Struchen MA, Atchison TB, Roebuck TM, et al. A multidimensional measure of caregiving appraisal: validation of the Caregiver Appraisal Scale in traumatic brain injury. *J Head Trauma Rehabil.* 2002;17:132–154.
22. Murray CJL, Lopez AD. Global mortality, disability and the contribution of risk factors: global burden of disease study. *Lancet.* 1997;349:1436–1442.
23. Luchter S, MacKenzie EJ, eds. Measuring the burden of injury. Conference proceedings. Baltimore, MD, May 11–13, 2000.
24. Bullinger M, Azouvi P, Brooks N, et al. Quality of life in patients with traumatic brain injury-basic issues, assessment and recommendations. *Restor Neurol Neurosci.* 2002;20:111–124.
25. Wood-Dauphinee S, Exner G, Bostanci B, et al. Quality of life in patients with spinal cord injury-basic issues, assessment, and recommendations. *Restor Neurol Neurosci.* 2002;20:135–149.
26. Kopjar B. The SF-36 health survey: a valid measure of changes in health status after injury. *Inj Prev.* 1996;2:135–139.
27. Holbrook TL, Anderson JP, Sieber WJ, et al. Outcome after major trauma: discharge and 6-month follow-up results from the Trauma Recovery Project. *J Trauma.* 1998;45:315–323.
28. Holbrook TL, Anderson JP, Sieber WJ, et al. Outcome after major trauma: 12-month and 18-month follow-up results from the Trauma Recovery Project. *J Trauma.* 1999;46:765–771.
29. Baldry Currens JA, Coats TJ. The timing of disability measurements following injury. *Injury.* 2000;31:93–98.
30. Baldry Currens JA. Evaluation of disability and handicap following injury. *Injury.* 2000;31:99–106.
31. Michaels AJ, Michaels C, Smith JS, et al. Outcome from injury: general health, work status, and satisfaction 12 months after trauma. *J Trauma.* 2000;48:841–850.
32. MacKenzie EJ, Sacco WJ, Luchter S, et al. Validating the Functional Capacity Index as a measure of outcome following blunt multiple trauma. *Qual Lif Res.* 2002;11:797–808.
33. Vazquez Mata G, Rivera Fernandez R, Perez Aragon, et al. Analysis of quality of life in polytraumatized patients two years after discharge from an intensive care unit. *J Trauma.* 1996; 41:326–332.
34. Anke AGW, Stanghelle JK, Finset A, et al. Long-term prevalence of impairments and disabilities after multiple trauma. *J Trauma.* 1997; 42:54–61.
35. Lehman U, Gobiet W, Regel H. Functional, neuropsychological and social outcome of multiple trauma patients with severe head injury (in German). *Unfallchirurg.* 1997;100:552–560.
36. Lehmann U, Pape HC, Seekamp A, et al. Long term results after multiple injuries including severe head injury. *Eur J Surg.* 1999; 165:1116–1120.
37. Van der Sluis CK, Eisma WH, Groothoff JW, et al. Long-term physical, psychological and social consequences of severe injuries. *Injury.* 1998;29:281–285.
38. Brenneman FD, Redelmeier DA, Boulanger BR, et al. Long term outcomes in blunt trauma: who goes back to work? *J Trauma.* 1997; 42:778–781.
39. Badia X, Diez-Perez A, Alvarez-Sanz C, et al. Spanish GRECO Study Group. Measuring quality of life in women with vertebral fracture due to osteoporosis: a comparison of the OQLQ and QUALEFFO. *Qual Life Res.* 2001;10:307–317.
40. Pirente N, Gregor A, Bouillon, et al. Quality of life of severely injured patients one year after trauma. A matched-pair study with a healthy control group (in German). *Unfallchirurg.* 2001;104:57–63.
41. Pirente N, Bouillon B, Schäfer B, et al. A systematically developed instrument for the assessment of health-related quality of life in multiple injured patients. The Polytrauma Outcome (POLO) Chart (in German). *Unfallchirurg.* 2002;105:413–422.
42. Stalp M, Koch C, Regel G, et al. Development of a scoring system for the standardized evaluation of the quality of rehabilitation in polytraumatized patients (HASPOC) (in German). *Chirurg.* 2001; 72:312–318.
43. MacKenzie EJ, McCarthy ML, Dittuno JF. Using the SF-36 for characterizing outcome after multiple trauma involving head injury. *J Trauma.* 2002;52:527–534.
44. Stalp M, Koch C, Ruchholtz S, et al. Standardized outcome evaluation after blunt multiple injuries by scoring systems: a clinical follow-up investigation 2 years after injury. *J Trauma.* 2002; 52:1160–1168.
45. Zelle B, Stalp M, Weihs Ch, et al. Validation of the Hannover Score for Polytrauma Outcome (HASPOC) in a sample of 170 polytrauma patients and a comparison with the 12-item Short-Form Health Survey (in German). *Chirurg.* 2003;74:361–369.
46. Dimipolou I, Anthi A, Mastora Z, et al. Health-related quality of life and disability in survivors of multiple trauma one year after intensive care unit discharge. *Am J Phys Med Rehab.* 2004;83:171–176.
47. Vles WJ, Steyerberg EW, Essink-Bot ML, et al. Prevalence and determinants of disabilities and return to work in major trauma. *J Trauma.* 2005;58:126–135.
48. Tidermark J. Quality of life and femoral neck fractures. *Acta Orthop Scand Suppl.* 2003;74:1–42.
49. Peterson MGE, Allegrante JP, Cornell CN, et al. Measuring recovery after a hip fracture using the SF-36 and Cummings Scales. *Osteoporosis Int.* 2002;13:296–302.
50. Tosteson ANA, Gabriel SE, Grove MR, et al. Impact of hip and vertebral fractures on quality-adjusted life-years. *Osteoporosis Int.* 2001;12:1042–1049.
51. Jaglal S, Lakhani Z, Schatzker J. Reliability, validity and responsiveness of the Lower Extremity Measure for patients with a hip fracture. *J Bone Joint Surg.* 2000;82:955–962.
52. Randell AG, Nguyen TV, Silverman SL, et al. Deterioration in quality of life following hip fracture: a prospective study. *Osteoporosis Int.* 2000;11:460–466.
53. Van Balen R, Essink-Bot ML, Steyerberg E, et al. Quality of life after hip fracture: a comparison of four health status measures in 208 patients. *Disabil Rehabil.* 2003;25:507–519.
54. Cummings SR, Phillips SL, Wheat ME, et al. Recovery after hip fracture: the role of social supports. *J Am Geriatr Soc.* 1988;36: 801–880.
55. Silverman SL, Mason J, Greenwald M. The Osteoporosis

Assessment Questionnaire (OPAQ): a reliable and valid self assessment measure of health related quality of life in osteoporosis. *J Bone Miner Res.* 1993;8:S3–43.
56. Kingma J, ten Vergert E, Werkman HA, et al. A Turbo Pascal program to convert ICD-9CM coded injury into injury severity scores: ICDTOAIS. *Percept Mot Skills.* 1994;78:915–936.
57. Brenneman FD, Boulanger BR, McLellan BA, et al. Measuring injury severity: time for a change? *J Trauma.* 1998;44:580–582.
58. Sullivan T, Haider A, DiRusso SM, et al. Prediction of mortality in pediatric trauma patients: new injury severity score outperforms injury severity score in the severely injured. *J Trauma.* 2003; 55:1083–1087.
59. Lavoie A, Moore L, LeSage N, et al. The New Injury Severity Score: a more accurate predictor of in-hospital mortality than the Injury Severity Score. *J Trauma.* 2004;56:1312–1320.
60. Frankema SP, Steyerberg EW, Edwards MJ, et al. Comparison of current injury scales for survival chance estimation: an evaluation comparing the predictive performance of the ISS, NISS, and AP scores in a Dutch local trauma registration. *J Trauma.* 2005;58: 596–604.
61. Harwood PJ, Giannoudis PV, Probst C, et al. Which AIS based scoring system is the best predictor of outcome in orthopaedic blunt trauma patients? *J Trauma.* 2006;60:334–340.
62. Sutherland AG, Johnston AT, Hutchison JD. The new injury severity score: better prediction of functional recovery after musculoskeletal injury. *Value Health.* 2006;9:24–27.
63. Stevenson M, Sequi-Gomez M, Lescohier I, et al. An overview of the injury severity score and the new injury severity score. *Injury Prevent.* 2001;7:10–13.
64. Guzzo JL, Bochicchio GV, Napolitano LM, et al. Prediction of outcome in trauma: anatomic or physiologic parameters? *J Am Coll Surg.* 2005;201:891–897.
65. Adachi JD, Ionnadis G, Pickard L, et al. The association between osteoporotic fractures and health-related quality of life as measured by the Health Utilities Index in the Canadian Multicentre Osteoporosis Study (CaMos). *Osteoporos Int.* 2003;14:895–904.
66. Raat H, Bonsel GJ, Essink-Bot ML, et al. Reliability and validity of comprehensive health status measures in children: the Child Health Questionnaire in relation to the Health Utilities Index. *J Clin Epidemiol.* 2002;55:67–76.
67. Feeny D, Furlong W, Torrance GW, et al. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. *Med Care.* 2002;40:113–128.
68. Stolk EA, Busschbach JJV. Performance of the EuroQol in children with imperforate anus. *Qual Life Res.* 2000;9:29–38.
69. Brazier JE, Walters SJ, Nicholl JP, et al. Using the SF-36 and Euroqol on an elderly population. *Qual Life Res.* 1996;5:195–204.
70. Dolan P. Modeling valuations for EuroQol health states. *Med Care.* 1997;35:1095–1108.
71. Walters SJ, Brazier JE. What is the relationship between the minimally important difference and health state utility values? The case of the SF-6D. *Health Qual Life Outcomes.* 2003;1:4.
72. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Economics.* 2002; 21:271–292.
73. Tan CP, Ng A, Civil I. Co-morbidities in trauma patients: common and significant. *N Z Med J.* 2004;117:U1044.
74. Edwards P, Roberts I, Clarke M. Increasing response rates to postal questionnaires: systematic review. *BMJ.* 2002;324:1183.
75. Bouillon B, Kreder HJ and the MI Consensus group. Quality of life in patients with multiple injuries-basic issues, assessment and recommendations. Results of a consensus meeting. *Rest Neurol Neurosci.* 2002;20:125–134.
76. Neugebauer E, Bouillon B, Bullinger M, et al. Quality of life after multiple trauma. Summary and recommendations of the consensus conference. *Rest Neurol Neurosci.* 2002;20:161–167.
77. Lamb SE, Jorstad-Stein EC, Hauer K, et al. Development of a common outcome data set for fall injury prevention trials: the Prevention of Falls Network Europe consensus. *J Am Geriatr Soc.* 2005;53:1618–1622.
78. Krabbe PFM, Stouthard ME, Essink-Bot ML, et al. The effect of adding a cognitive dimension to the EuroQol multiattribute health status classification system. *J Clin Epidemiol.* 1999;52: 293–301.
79. Kopec JA, Wilson KD. A comparative review of four preference-weighted measures of health-related quality of life. *J Clin Epidemiol.* 2003;56:317–325.
80. Luo N, Johnson JA, Shaw JW, et al. Self-reported health status of the general adult US population as assessed by the EQ-5D and Health Utilities Index. *Med Care.* 2005;43:1078–1086.
81. Cryer C, Langley JD, Jarvis SN. Injury outcome indicators: the development of a validation tool. *Inj Prev.* 2005;11:53–57.
82. McClure RJ, Peel N, Kassulke D, et al. Appropriate indicators for injury control? *Public Health.* 2002;116:252–256.
83. Lyons RA, Polinder S, Larsen CF, et al. Methodologic issues in comparing injury incidence across countries. *Inj Contr Saf Prom.* 2006;13:63–70.
84. van Baar ME, Essink-Bot ML, Oen IM, et al. Functional outcome after burns: a review. *Burns.* 2006;32:1–9.
85. Kildal M, Anderson G, Fugl-Meyer AR, et al. Development of a brief version of the Burn Specific Health Scale (BSHS-B). *J Trauma.* 2001;51:740–746.
86. Daltroy LH, Liang MH, Phillips CB, et al. American Burn Association/Shriners Hospitals for Children Burns Outcomes Questionnaire: construction and psychometric properties. *J Burn Care Rehabil.* 2000;21:29–39.
87. Kazis LE, Liang MH, Lee A, et al. The development, validation, and testing of a Health Outcomes Burn Questionnaire for Infants and Children 5 years of age and younger: American Burn Association/ Shriners Hospitals for Children. *J Burn Care Rehabil.* 2002;23:196–207.
88. van Baar ME, Essink-Bot ML, Oen IM, et al. Reliability and validity of the Health Outcomes Burn Questionnaire for infants and children in The Netherlands. *Burns.* 2006;27:790–802.
89. Landgraf JM, Ware JE. The CHQ User's Manual. Boston: The Health Institute, New England Medical Center, 1996.
90. Klassen A, Landgraf JM, Lee S et al. Health related quality of life in 3 and 4 year old children and their parents: preliminary findings about a new questionnaire. *Health Qual Life Outcomes.* 2003;1:81.
91. Winthrop AL, Brasel KJ, Stahovic L, et al. Quality of life after pediatric trauma. *J Trauma.* 2005;58:468–447.
92. Eiser C, Morse R. Can parents rate their child's health-related quality of life? Results of a systematic review. *Qual of Life Res.* 2001;10:347–357.
93. Petrou S. Methodological issues raised by preference-based approaches to measuring the health status of children. *Health Economics.* 2003;12:697–702.
94. Sturms LM, Sluis CK, Groothoff HJ, et al. Young traffic victims' long-term health-related quality of life: child self-reports and parental reports. *Arch Phys Med Rehabil.* 2003;84:431–436.
95. Hennessy S, Kind P. Measuring health status in children: developing and testing a child-friendly version of EQ-5D. York, UK: Outcomes Research Group, Centre for Health Economics, University of York; 2002.
96. Turner AP, Barlow JH, Heathcote-Elliott C. Long term health impact of playing professional football in the United Kingdom. *Br J Sports Med.* 2000;34:332–337.
97. Polinder S, Meerding WJ, van Baar ME, et al. Cost estimation of injury-related hospital admissions in 10 European countries. *J Trauma.* 2005;59:1283–1290.