Research Paper ■

Incorporating the International Classification of Functioning, Disability, and Health (ICF) into an Electronic Health Record to Create Indicators of Function: Proof of Concept Using the SF-12

NANCY E. MAYO, PHD, LISE POISSANT, PHD, SARA AHMED, MSC, LOIS FINCH, MSC, JOHANNE HIGGINS, MSC, NANCY M. SALBACH, MSC, JUDITH SOICHER, MSC, SUSAN JAGLAL, PHD

Abstract **Objective:** The purpose of this proof-of-concept study was to assess the feasibility of using a generic health measure to create coded functional status indicators and compare the characterization of a stroke population using coded functional indicators and using health-related quality-of-life summary measures alone.

Design: Multiple raters assigned International Classification of Functioning, Disability, and Health (ICF) codes to the items of the 12-Item Short Form Health Survey (SF-12). Data for comparing the information from the SF-12 and from ICF codes were derived from the Montreal Stroke Cohort Study that was set up to examine the long-term impact of stroke. Available for analysis were data from 604 persons with stroke, average age 69 years, and 488 controls, average age 62 years.

Measurement: The SF-12 provides two summary scores, one for physical health and one for mental health. Domains of the ICF are coded to three digits, before the decimal; specific categorizations of impairments, activity limitations, and participation restrictions are coded to four digits before the decimal.

Results: Persons with stroke scored, on average, approximately 10 points lower than controls on physical and mental health. The ICF coding indicated that this was attributed, not surprisingly, to greater difficulty in doing moderate activities including housework, climbing stairs, and working and was not attributed to differences in pain. Differences in mental health were attributed most strongly to greater fatigue (impairment in energy), but all areas of mental health were affected to some degree.

Conclusion: The ICF coding provided enhanced functional status information in a format compatible with the structure of administrative health databases.

■ **J Am Med Inform Assoc.** 2004;11:514–522. DOI 10.1197/jamia.M1462.

Electronic health records (EHRs) are increasingly being deployed in clinical settings. This technology provides the capacity to capture clinical information about individuals at their point of contact with the health care system and to make this information available across the continuum of care.[1,2] A very advantageous feature of EHRs is the ability to supply data for clinical, population, and health services research.[3] EHRs provide a mechanism for capturing patient-reported outcomes in a standardized format[4–7] through disease-specific or generic health measures that are increasingly important across the spectrum of health research.

The choice of health measure to include in an EHR is likely to be dictated by response burden and simplicity of data capture. Existing health-related quality-of-life (HRQL) measures are attractive for use in EHRs because they capture the key components of physical, emotional, and psychological health parsimoniously.[8,9] Such measures have undergone considerable testing and have been shown to reflect the individual's perspective on health[10] meaningfully and reliably. However, HRQL measures all have a common disadvantage: their summary scores are not readily interpretable. Accurate interpretation requires reference to normative tables and complex weighting algorithms that trade off health advantages in some domains against liabilities in other domains. One advantage of integrating an HRQL measure into an EHR is that the summary scores may be calculated automatically and are available to enrich research programs. The specific problems faced by the individuals, however, are not discernible in the summary score and, thus, not easily flagged for intervention or special services. If coded in a standard way, item-level HRQL information would be even more valuable than the

Affiliations of the authors: Department of Clinical Epidemiology (NEM, SA, JH, NMS), School of Physical and Occupational Therapy (NEM, LF, JH), Department of Clinical and Health Informatics (LP), Respiratory Epidemiology and Clinical Research Unit (JS), McGill University Health Center, Montreal, Quebec, Canada; and Department of Physical Therapy, University of Toronto, Toronto, Ontario, Canada (SJ).

Supported by the funding from Réseau de recherches en santé de population du Québec.

Correspondence and reprints: Nancy E. Mayo, PhD, Division of Clinical Epidemiology, Ross Pavilion R4.29, Royal Victoria Hospital Site, 687 Pine Avenue West, Montreal, QC, Canada, H3A 1A1; e-mail: <nancy.mayo@mcgill.ca>.

Received for publication: 09/22/03; accepted for publication: 06/21/04.

summary scores as indicators of the specific problems faced by the individual.[11] Standard coding of functional status and quality-of-life data in EHRs could also address gaps in the data available in administrative databases and large surveys for use in population and health services research.[3,12,13]

The World Health Organization's (WHO) International Classification of Diseases (ICD-9 or -10)[14] or one of its modifications[15] is the coding system used almost universally in administrative databases to code diseases, disorders, and injuries.

The International Classification of Functioning, Disability, and Health (ICF),[16] developed by the WHO over the past two decades to classify the consequences of disease, complements the ICD. The current ICF lists 1,424 categories referring to body functions (b) and structure (s), activities and participation (d), and environmental factors (e). In previous editions, the terms impairment, disability, and handicap were used to classify the consequences of disease. More recently, this negative terminology has been replaced with the more positive terms of body structure and function, activity and participation, with the corresponding negative terms being impairment, activity limitation, and participation restriction. The expression "consequences of disease" has been replaced with "components of health." All negative components of health are collectively termed disability, and all positive components are collectively termed functioning. In many scientific contexts, these terms are used interchangeably without consideration that they define different concepts. A body structure/function refers to the anatomical/physiological component of a body (e.g., oral pharynx/voice), an activity is the execution of a task (e.g., walking), and participation refers to the involvement of an individual in life situations (e.g., working). ICF also permits the discrimination between the capacity for an activity and the actual performance of it in daily life. In other words, an individual may have the capacity to walk a short distance but does not actually perform walking because it is easier to use a wheelchair. Additionally, impairment, capacity, or performance can be further qualified according to the degree of severity.

As with the ICD system, the ICF system could conceivably be incorporated into administrative databases providing key information for policy makers and researchers about the activity limitations and participation restrictions that translate directly into need for services. If provided in a standard way, information on these functional limitations across populations would be very helpful for planning rehabilitation, home care, and long-term care resources as well as explaining outcomes associated with health care services or interventions.[17]

The ICF has the potential to provide standard coding for the various items on the different measures commonly used in the health field that would permit the characterization of the specific functional problems of individuals.[17–19] A coded problem list could be produced independent of which of the many possible measures were used. The codes could then be transmitted to the agencies responsible for maintaining the health databases and incorporated into these population-based resources. Efforts to code existing measures of functional limitation with the ICF are just beginning. Codes have been reported for the Western Ontario and McMaster Universities osteoarthritis measure,[20] the New York Heart Association Functional Class,[17] and the Karnofsky Performance Status.[17] Cieza et al.[21] have provided ICF codes for six HRQL measures, including the 36-Item Short Form Health Survey (SF-36). In spite of growing interest in pursuing coding activities, the practical advantage of moving from a standardized outcome measure to a standard coding system has not yet been demonstrated.

This study is a proof of concept that used the ICF to encode answers to the individual questions in the 12-Item Short Form Health Survey (SF-12) HRQL measure. The purpose of this study was to assess the feasibility of using the individual items in a generic health measure to create coded functional status indicators and to compare the characterization of the disability of a stroke population using these coded functional indicators and using HRQL summary measures alone.

Methods

Data were obtained from patients enrolled in the Montreal Stroke Cohort Study that was set up to examine the long-term impact of stroke.[22] Interviews were conducted with community-dwelling persons with stroke (n = 604) and age- and gender-matched individuals without stroke (n = 488). Persons with a first ever, ischemic, or hemorrhagic stroke were identified from ten acute care hospitals in the Montreal area using an inception cohort approach. Those who were discharged back to the community, either directly or via rehabilitation, were interviewed over the telephone with a battery of tests covering activity participation and HRQL. For persons who could not respond for themselves, caregivers provided information about functional status, but proxies were not used for the HRQL measures. There were no exclusions based on age or language. Randomly generated samples of computerized telephone listings were used to assemble the peer control group. The same battery of tests was administered to the control cohort. Basic demographic information was also obtained. The results at six months post-stroke have been described in a previous report.[22] This paper presents information from the first interview for persons with stroke conducted around the time of reentry into the community and from the controls. The time frame for the first interview varied from 2 weeks to 3 months post-stroke. Because this is a proof-of-concept study, the timing of the interview is not relevant.

The participants in this study were interviewed using the Measuring Outcomes Study SF-36.[23] Because the response burden of this measure might be beyond the threshold for inclusion on a generic EHR, we used only the answers to the SF-12[24] subset of the SF-36 in this study. The SF-12 provides two summary scores, one for physical health, called the Physical Component Summary, and one for mental health, called the Mental Component Summary.

The SF-12[24] is a valid and reliable measure of health status reflecting the values and preferences for health from the person's perspective. It correlates very highly with the SF-36 (>0.95), the well-known longer version,[23] and the validity of the SF-36 has been extensively established against other measures of health status.[23,25] The test–retest reliability of the two summary measures, Physical and Mental Health, has been reported to range from 0.76 to 0.89.[24] Reliability coefficients indicate how much variability in a score is due to true variation in the trait and how

much is due to chance or random error. For valid group comparisons, reliability needs to exceed 0.70.[26] For this study, the measurement issue is the extent to which self-report reflects the individual's actual capacity for those traits that can be directly observed, like climbing stairs or doing housework. It is not possible to measure pain or psychological distress by methods other than self-report. There is evidence of good correlation between patient reports of functional limitations and ratings based on direct observation by a rehabilitation professional (90% agreement)[27–29]; however, when there were disagreements, it was because the individuals reported better function than what was actually observed. This overestimation of ability was greater among persons with moderate to severe functional limitations.[27] So, if anything, any reported limitation is more likely to be more rather than less severe when directly measured.

Eight rehabilitation professionals independently assigned the 12 items of the SF-12 to alphanumeric ICF codes. Codes are preceded by the letters b signifying impairments of body function, s signifying impairments of body structure, and d signifying activities and participation. Each letter is followed by either a three- or four-digit code representing the level of granularity captured. For example, walking is a three-digit code (d450) and walking short distances is a four-digit code (d4500). As many as two decimal places are used for severity qualifiers that include none, mild, moderate, severe, complete, or not specified. The first decimal place of a "b" code qualifies the extent of the impairment and in the case of a "d" code, the level of performance.

Raters were instructed to identify all codes that would fit the SF-12 item and indicate the severity qualifier that best fit the response options of the SF-12. Subsequently, the two senior raters reviewed the pool of selected codes and came to a consensus as to the code that best reflected the meaning of the SF-12 item. Wherever possible, a four-digit code was chosen over a three-digit code to be as specific as possible about the functional limitation. If more than one four-digit code applied, then the code having a 9 ("unspecified") as the fourth digit before the decimal was chosen.

Statistical Methods

Means for the Physical Component Summary and Mental Component Summary were calculated and contrasts across gender, age, and income were made using multiple linear regression. The distribution on functional limitations as coded by the ICF classification was compared separately for persons with stroke and controls across age, gender, and income using logistic regression for multiple ordinal categories.[30] This analysis provides a proportional odds ratio and a test of homogeneity across cut points. Persons with stroke and controls were also compared on SF-12 subscales using multiple linear regression and on ICF coded functional limitations using categorical logistic regression. For the latter analysis, when the score test indicated significant heterogeneity, analyses focused on contrasting the most severe categorization of functional limitation and binary logistic regression was used. To facilitate logistic models, the contrast variables age and income were categorized as follows: age, younger than 55/55 to 64/65 to 80/older than 80 years; income, less than $10,000/$10,000 to $19,000/$20,000 to $29,000/$30,000 to $39,000/$40,000 to $49,000/more than $50,000. An α level of 0.05 indicated statistical significance.

Results

Table 1 shows the ICF codes identified for each SF-12 item by the consensus raters, by all raters in this study, and by another ICF team in Germany.[21]

The only item of the SF-12 that was not definable using the ICF framework was item 1 (evaluating own health). Item 2 had two possible codes, one for activities related to recreation and leisure and one for housework. For the purposes of illustrating the additional information gained by coding the specific items, we chose to use the code d640, for housework, because this is an important limitation for service planning in the community. It is also likely that people who cannot

Table 1 ■ The ICF Codes Identified for the SF-12 by Consensus of the McGill Raters and the ICF Research Branch of the German WHO Collaborating Center for the Family of International Classifications

	ICF Code*	
SF-12 Item	McGill Consensus (No. of Raters)	German ICF Research Branch
1. Evaluating own health	No match	nd (not definable)
2. Doing moderate activities such as moving a table, pushing a vacuum cleaner, bowling, playing golf	d920 (2) d640 (2)	d920 Recreation and leisure d640 Doing housework
3. Climbing several flights of stairs	d4551 (8)	d4551 Climbing, includes steps
4–7. Interference in work† or regular activities due to physical and/or emotional health	d859 (3)	d859 Work and employment otherwise unspecified d850 Remunerative employment b230 Carrying out daily routine
8. Pain interfering with work or housework	b280 (5)	b280 Sensation of pain
9. Feeling calm and peaceful	b1263 (6) Psychic stability	b152 or b1529 Emotional functions or unspecified
10. Having a lot of energy	b1300 (8)	b1300 Energy level
11. Feeling downhearted and blue	b1265 (5) Optimism	b152 or b1529 Emotional functions or unspecified
12. Interference with social activities	d9205 (7)	d9205 Socializing, unspecified

*Rating could be at three- or four-digit level.
†Work items of the SF-12: 4, accomplished less work or regular activities due to physical health; 5, limited in kind of work or regular activities due to physical health; 6, accomplished less work or regular activities due to emotional health; 7, did not work or do regular activities as carefully due to emotional health.

do housework cannot do moderately demanding recreational activities either.

The four items referring to work were the most challenging to code (items 4, 5, 6, 7). The SF-12 items qualify the limitation in working according to amount of time, quality, cause, and impact on physical or emotional health. The ICF is more concerned with the type of limitation that a person would have in carrying out work and thus the codes are highly specific. The code that fit the best and where there was the greatest degree of agreement for each of the four items was d859, "work and employment, other specified and unspecified." This code was also one chosen by the German ICF group.[21]

Table 2 gives the ICF codes along with the severity qualifiers chosen by the eight raters. Note that work is dichotomized; therefore, no severity code can be applied. There was very little disagreement on severity ratings; seven or eight of eight raters agreed on each rating.

Available for analysis were data from 604 persons with stroke and 488 controls. As illustrated in Table 3, the average age of the persons with stroke was 69 years and of the controls 62 years.

The following tables and figures contrast the information obtained from analyses of the two SF-12 summary scores, Physical Health (PCS) and Mental Health (MCS), with that obtained when analyses were carried out according to specific functional limitation, with the associated severity rating coded using the ICF. Because ICF is concerned with disability in terms of impairments, limitations, and restrictions, here globally termed functional limitations, it is more desirable to have a lower level of functional limitation. Higher numbers are preferable for the SF-12 summary scores.

Analyses on SF-12 Summary Scores

Table 4 shows the mean PCS and MCS summary scores across gender, age, and income categories for persons with stroke and controls. Also given are the regression coefficients that provide an estimate of the difference between two means adjusted for the other variables. For persons with stroke, there was a statistically significant difference between men and women on mental health with women scoring on average 2.72 points below men. For controls, women scored lower than men on both physical and mental health. For controls, there was also a statistically significant difference in physical health for persons older than the age of 80 years compared with persons younger than 80. For persons with stroke, age did not have an effect on physical or mental health. Other significant differences were observed for income, with more wealthy individuals reporting better health status.

Analyses by ICF: Gender

Table 5 contrasts men and women on functional limitations as coded through the ICF. The table presents the proportion of men and women reporting a functional limitation along with the odds ratio (OR) and 95% confidence interval (CI). An OR greater than 1 indicates that women were more likely than men to report a functional limitation; a 95% CI that excludes 1 indicates statistical significance. All ORs are adjusted for age and income. There were common gender differences between stroke and controls for climbing stairs, psychic stability, and optimism, with women consistently reporting significantly higher rates of functional limitations than men in both stroke and control groups. For persons with stroke, there were additional gender differences in housework and energy limitations. In neither the stroke group nor the controls did gender have an impact on pain or socializing.

Analyses by ICF: Age

Analyses revealed that only persons older than 80 years of age experienced a greater degree of functional limitation than younger persons, and these contrasts are presented in Table 6. The proportion of persons with stroke between 55 and 80 years of age with any functional limitation in the area of housework was 55%; the proportion with a severe limitation was 28%. For older persons with stroke, these proportions (any and severe) were 43% and 30%, respectively. For controls, the proportions reporting housework limitations were 16% for younger persons and 48% for persons older than 80 years of age; severe limitations were rarer: 5% and 15%, respectively, for these two age groupings. The impact of age on housework limitations and on climbing stairs was very strong among controls (OR 4.16; 95% CI = 2.03–8.51 and OR 4.11; 95% CI = 2.05–8.24, respectively). Among persons with stroke, the only limitation for which age was a factor was climbing stairs.

Analyses by ICF: Income

Figure 1 presents the distribution of the functional limitations by three categories of income: <$19,000, $20,000 to $39,000, and more than $40,000. Table 7 presents the results of the ordinal regression models. For persons with stroke, there were many areas in which persons with higher incomes were advantaged (lower risk of reporting a functional limitation; OR 1.0). Only in the area of housework was there an advantage even with moderate income. For controls, higher income was also an advantage in many areas of functioning.

Table 2 ■ ICF Severity Rating of the Response Categories of the SF-12 Items

(SF-12 Question) Impairment or Disability	Complete (.4)	Severe (.3)	Moderate (.2)	Mild (.1)
(2) d6409 Housework	Quite limited			Limited a bit
(3) d4551 Climbing stairs	Quite limited			Limited a bit
(8) b2800 Pain	Extremely	Quite a bit	Moderately	A little bit
(9) b1263 Psychic stability*	All the time	Most of the time/a good bit of the time	Some of the time	A little of the time
(10) b1300 Energy*	All the time	Most of the time/a good bit of the time	Some of the time	A little of the time
(11) d1265 Optimism	All the time	Most of the time/a good bit of the time	Some of the time	A little of the time
(12) d9205 Socializing	All the time	Most of the time	Some of the time	A little of the time

*Items 9 (feeling calm and peaceful) and 10 (having a lot of energy) are worded positively, whereas the ICF coding refers to disability, so all the time would be a very severe impairment.

Table 3 ■ Description of the Study Samples

	Stroke	Controls
No.	604	488
No. of women (%)	259 (43)	330 (68)
No. of men (%)	345 (57)	158 (32)
Mean age (SD)	69.3 (12.6)	62.2 (12.2)

Analyses by Stroke or Control

Table 8 provides estimates of the extent to which persons with stroke reported greater functional limitations or lower health status than controls after adjustment for age, gender, and income. Pain was least affected by stroke and work was most affected. There was also an important effect of stroke on energy level, stair climbing, and housework. Using the SF-12 component scores, persons with stroke scored an estimated 10 points lower for physical health and 8.6 points lower on mental health than controls.

Discussion

The SF-12 component summary scores indicated that persons with stroke scored, on average, approximately 10 points lower than controls on physical and mental health. The ICF coding indicated that this was attributed, not surprisingly, to greater difficulty in doing moderate activities like housework, climbing stairs, and working and less to differences in pain. Differences in mental health were attributed most strongly to greater fatigue (impairment in energy), but all areas of mental health were affected to some degree. At this crude level, comparing persons with stroke with members of the general population, the SF-12 and the ICF coding provided largely the same information except that the ICF codes would be more suitable for inclusion in an administrative database.

In in-depth analyses, the advantage of the ICF coding emerges. Using the SF-12, there were no differences between men and women on physical health, yet the ICF coding revealed that women reported more problems with moderate activities like housework. This may be a true difference because of the greater physical problems that women may have or a difference in reporting of problems. For controls, physical health differences between women and men were not related to differences in ability to do moderate activities like housework but from differences in stair climbing and work. For mental health, the SF-12 and the ICF recoding yielded the same degree of information: women scored lower than men in all areas.

The SF-12 analyses for persons with stroke showed that age had no impact, but the ICF analysis identified that older persons with stroke had greater difficulty climbing stairs than younger persons. This could have important ramifications for service provision and for independent living. Older controls had more difficulty than younger in both physical and mental health, and the ICF analysis did not provide any additional information. With respect to income, wealthier persons with stroke reported better health than persons with lower incomes, and, according to the ICF codes, this was because of less difficulty with moderate activities like housework and stair climbing. This may reflect the ability of wealthier persons to have more conveniences or help in the home to assist in these areas. Interestingly, for persons with stroke, there was no effect of income on the SF-12 summary measure MCS, but the ICF analysis indicated that there were significant differences in all the specific mental functions. The additional information provided by the ICF coding would be helpful in flagging lower income persons for closer surveillance of mental health problems and may explain differences in outcomes not explained by stroke-related impairments.

With the advent of electronic health records, the items on the SF-12 (and on other measures) might be captured as a byproduct of individual health care and then automatically translated into ICF codes that could be exported to administrative databases. The main barrier to adding functional status indicators to administrative databases is the data collection burden. If data collection can occur during individual health care encounters, inclusion of functional status data in administrative databases may be feasible.

Table 4 ■ Physical and Mental Health Scores of the SF-12 According to Gender, Age, and Income Category

	Stroke (n = 604)		Controls (n = 488)	
Contrast	PCS Mean (SD)	MCS Mean (SD)	PCS Mean (SD)	MCS Mean (SD)
Men	41.0 (10.2)	45.8 (11.8)	51.3 (6.8)	55.1 (7.3)
Women	38.7 (10.0)	43.4 (10.9)	49.3 (8.6)	50.7 (10.2)
β [SE] M:W	−2.04 [1.1]	−2.72 [1.18]*	−1.74 [0.81]*	−4.36 [0.98]†
Age (yr)				
<55	41.7 (9.7)	43.6 (12.2)	51.5 (7.6)	51.2 (10.0)
55–64	41.5 (10.2)	43.5 (10.1)	50.5 (7.6)	52.7 (9.2)
65–80	39.4 (10.5)	45.1 (11.9)	49.6 (7.8)	52.2 (9.7)
>80	39.0 (9.4)	46.0 (11.1)	45.1 (9.3)	54.5 (8.3)
β [SE] ≥80:<80	−0.62 [1.32]	1.53 [1.49]	−4.93 [1.53]†	2.87 [1.85]
Income				
<$19,000	37.56 (11.41)	43.38 (13.35)	47.25 (9.22)	50.13 (11.13)
$20,000–$39,000	40.50 (9.48)	42.18 (10.50)	51.38 (7.12)	51.79 (9.77)
>$40,000	44.18 (10.30)	46.96 (12.11)	51.69 (6.99)	54.31 (7.56)
β [SE] >$40,000:<$19,000	6.05 [2.07]†	3.39 [2.33]	3.72 [1.22]†	3.02 [1.48]*

SE = Standard error.
β coefficients are interpreted as adjusted (for other variables in the model) mean differences; β/SE is equivalent to a t-test.
*$p < 0.05$.
†$p < 0.005$.

Table 5 ■ Extent to Which Women Differ from Men on Functional Limitations (% Any/% Severe or Complete) as Illustrated by OR and 95% CI

	Stroke (% Any/% >Severe)			Controls (% Any/% >Severe)		
ICF Code	Men (n = 345)	Women (n = 259)	OR (95% CI)	Men (n = 158)	Women (n = 330)	OR (95% CI)
d6409 Housework	43/25	53/33	1.4 (1.0–2.0)	16/3	20/7	1.3 (0.7–2.2)
d4551 Climbing stairs	53/37	69/53	1.7 (1.1–2.4)	21/7	34/15	1.7 (1.1–2.8)
d8599 Work	71	74	1.2 (0.8–1.8)	16	26	1.9 (1.2–3.2)
b2800 Pain	39/20	43/22	1.0 (0.7–1.5)	31/7	38/11	1.3 (0.9–1.9)
b1263 Psychic stability	55/18	65/19	1.4 (1.0–2.0)	27/3	48/15	2.6 (1.7–4.0)
b1300 Energy	75/37	85/48	1.6 (1.1–2.2)	40/8	46/15	1.3 (0.9–1.9)
d1265 Optimism	57/17	70/28	1.8 (1.2–2.5)	38/8	57/11	1.9 (1.3–2.8)
d9205 Socializing	59/20	65/23	1.0 (0.7–1.5)	30/6	42/8	1.8 (1.2–2.5)

Except for work, all functional limitations are classified polychotomous ordinal and, therefore, ORs are cumulative across all possible cut points. ORs are adjusted for age (younger than 55/55 to 64/65 to 80/older than 80 years) and income categories (less than $10,000/$10,000 to $19,000/$20,000 to $29,000/$30,000 to $39,000/$40,000 to $49,000/more than $50,000); OR greater than 1.0 indicates that women are more likely than men to report functional limitations; 95% CIs that exclude 1.0 indicate statistical significance. CIs with lower limits of 1.0 have been rounded up from 0.95 or greater.

CI = confidence interval; ICF = International Classification of Functioning, Disability, and Health; OR = odds ratio.

Table 6 ■ Extent to Which Older Persons Differ from Younger Persons on Functional Limitations (% Any/% Severe or Complete) as Illustrated by OR and 95% CI

	Stroke (% Any/% >Severe)			Controls (% Any/% >Severe)		
ICF Code	55–80 yr	>80 yr	OR (95% CI)	55–80 yr	>80 yr	OR (95% CI)
d6409 Housework	55/28	43/30	1.16 (0.74–1.81)	16/5	48/15	4.16 (2.03–8.51)
d4551 Climbing stairs	57/40	71/57	1.62 (1.04–2.55)	26/11	64/24	4.11 (2.05–8.24)
d8599 Work	73	70	0.84 (0.50–1.39)	22	33	1.53 (0.70–3.34)
b2800 Pain	41/21	42/18	0.97 (0.62–1.53)	36/9	30/12	0.94 (0.45–1.98)
b1263 Psychic stability	61/18	54/15	0.67 (0.44–1.03)	42/11	33/6	0.68 (0.32–1.44)
b1300 Energy	78/42	83/43	0.96 (0.64–1.45)	44/12	48/15	1.19 (0.60–2.33)
d1265 Optimism	62/22	62/20	0.88 (0.58–1.34)	51/10	48/6	0.72 (0.36–1.43)
d9205 Socializing	62/21	62/21	0.99 (0.65–1.51)	38/7	33/3	0.66 (0.31–1.42)

Except for work, all functional limitations are classified polychotomous ordinal, and, therefore, ORs are cumulative across all possible cutpoints. ORs are adjusted for gender and income categories (less than $10,000/$10,000 to $19,000/$20,000 to $29,000/$30,000 to $39,000/$40,000 to $49,000/more than $50,000). Functional limitations are reported for persons between 55 and 80 years of age and for persons older than 80. ORs greater than 1.0 indicate that persons older than 80 years report more functional limitations than younger persons; 95% CIs that exclude 1.0 indicate statistical significance.

CI = confidence interval; OR = odds ratio.

As part of the process to test the ICF, the Canadian Institute for Health Information[31] consulted with physiotherapists. Most did not perceive a great need for the ICF in physiotherapy and noted a medium to high level of difficulty in using it. The combination of lack of perceived utility and relative difficulty of use is likely to discourage use of the ICF as part of daily charting practices in physiotherapy, particularly in addition to the standardized outcome measures that are already used.

Routine use of outcome measures is the cornerstone of evidence-based care. Many of these outcome measures have had decades of research, both conceptual and empirical, and millions of dollars have been invested in their development, validation, and refinement. These outcome measures have been developed for evaluative purposes[32] and have been shown to be responsive and sensitive to change. The deployment of an EHR could simultaneously shift and reduce the burden of coding items from existing measures.

The work presented here illustrates how a coding system can be used to translate information from a standard outcome measure into functional indicators. The observation that there was some variability in the codes chosen by the different raters does not invalidate the findings. It is often the case in examining administrative databases for diagnostic codes that several are listed. The researcher or administrative organization using the codes invokes their own rules for identifying which codes are used to select people with various diagnoses, etc. Thus, having more codes would be preferable to having fewer.

This study used the SF-12 to illustrate the feasibility of using a standard outcome measure to create a coded list of functional status indicators. In reality, it is more likely that development of EHRs will proceed with disease-specific add-ons, in which case, outcomes specific for each condition are likely to be included. This would avoid some of the problems with having one SF-12 item referring to several activities (e.g., doing moderate activities such as moving a table, pushing a vacuum cleaner, bowling, playing golf). There is no way of knowing, except through individual cognitive debriefing, to which activity the person was referring when responding, although these activities are all quite similar with respect to the body structures and functions required (i.e., all require upper and lower limb strength, mobility, and endurance).

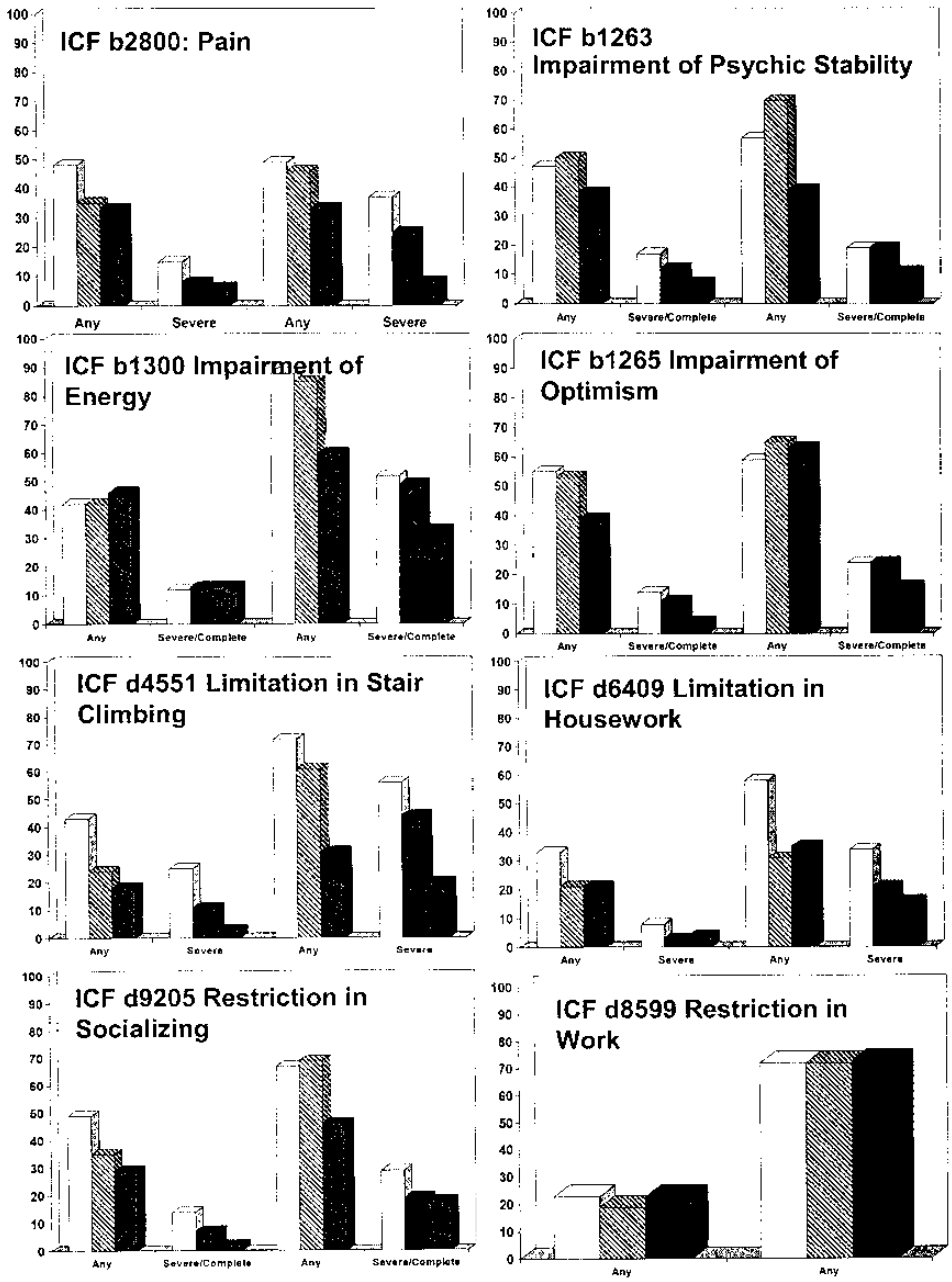

Figure 1. Proportion of persons with stroke reporting any or any severe functional limitation according to three income categories. Less than $19,000 (*white columns*), $20,000 to $39,000 (*hatched columns*), and more than $40,000 (*dark gray columns*).

Cognitive Aspects of Survey Methodology (CASM) is a framework used for understanding the cognitive processes that respondents deploy in reading, comprehending, and interpreting questions and in formulating and providing answers. The CASM model suggests that there are four stages in the response process: comprehension, retrieval, judgment, and response.[33,34] For subjects who have difficulty with any of these stages, there is likely to be greater discordance between verbal reports and performance-based assessments. This is likely the explanation for the finding of Korner-Bitensly and Wood-Dauphinne[27] that 10% of persons, on follow-up after discharge from inpatient rehabilitation, overestimated their functional capacity. For persons with stroke, the SF-12 would not be the first choice for a measure of function. To be comprehensive, measures of basic and instrumental activities of daily living, emotion, pain, and participation would need to be included. However, the advantage of the SF-12 is that it captures indicators of health applicable to all.

A limitation of the SF-12 is that it does not capture the full spectrum of functional limitations that are consequences of many conditions or accidents. In designing an EHR for use with health applicable to all, it is widely known and validated and is short. Therefore, it provides sufficient detail to illustrate the principle of generating a coded list of functional status indicators from a standard outcome measure. Future work is ongoing by us and by the German ICF research group[20,21] to validate ICF core sets for various conditions,

Table 7 ■ Odds Ratios and 95% Confidence Intervals for the Effect of Income on Functional Limitations for Persons with Stroke and Controls

	Stroke		Controls	
ICF Code	$20,000–39,000 vs. <$19,000	>$40,000 vs. <$19,000	$20,000–39,000 vs. <$19,000	>$40,000 vs. <$19,000
d6409 Housework	0.38 (0.22–0.65)	0.45 (0.22–0.94)	0.46 (0.23–0.92)	0.47 (0.21–1.05)
d4551 Climbing stairs	0.67 (0.38–1.14)	0.22 (0.10–0.48)	0.43 (0.25–0.73)	0.29 (0.15–0.57)
d8599 Work	0.99 (0.53–1.87)	1.05 (0.46–2.41)	0.88 (0.47–1.64)	1.19 (0.60–2.36)
b2800 Pain	0.75 (0.44–1.27)	0.39 (0.19–0.82)	0.57 (0.35–0.93)	0.55 (0.31–0.96)
b1263 Psychic stability	1.13 (0.67–1.89)	0.45 (0.22–0.92)	1.01 (0.63–1.62)	0.46 (0.26–0.83)
b1300 Energy	0.82 (0.50–1.36)	0.43 (0.22–0.84)	1.05 (0.64–1.71)	1.29 (0.75–2.23)
d1265 Optimism	1.21 (0.73–2.03)	1.03 (0.53–2.03)	0.85 (0.53–1.34)	0.52 (0.30–0.89)
d9205 Socializing	0.87 (0.52–1.44)	0.45 (0.23–0.90)	0.52 (0.32–0.85)	0.38 (0.21–0.69)

which will contribute additional evidence of validity and reliability of functional status indicators.

The response categories of the SF-12 refer to the degree of limitation or restriction, whereas the ICF uses severity indicators to qualify function. The terms mild, moderate, severe, and complete may mean different things to different people in different environmental contexts, which is a limitation of the ICF for use in health services planning and resource allocation. Our research team is in the process of developing behavioral and verifiable indicators that represent the severity of different activity limitations and participation restrictions.

During the process of developing a preference-based stroke index (PBSI),[35] we developed statements to represent ranked categories to describe limitation or restriction in ten domains important for the stroke population: walking, stair climbing, physically demanding activities, quiet recreational activities, work or usual activities, driving a car, memory, speech, coping, and self-esteem. For example, for the walking limitation domain of the PBSI, the three response options could be modified to obtain severity ratings suitable for coding: (1) I am able to walk in the community because I need to (none or mild depending on whether there was a walking problem indicated elsewhere); (2) I am able to walk inside the house, but I have difficulty walking alone outside (moderate); (3) I am able to walk only a few steps (severe) or I use a wheelchair (complete). For the PBSI item for limitation in stair climbing, the three severity ratings are as follows: (1) I can go up and down several flights of stairs (none if no walking limitation or mild if there is); (2) I can go up and down only a few steps (moderate); and (3) I cannot go up and down stairs (complete). Thus, future work will develop coding algorithms for other functional status measures and descriptive statements to be used to assign a code for severity.

Table 8 ■ Estimated Impact of Stroke on Functional Limitations and Physical and Mental Health of the SF-12

ICF Coded Variable	OR (β)	95% CI (SE)
Housework	4.4	3.2–6.0
Climbing stairs	4.4	3.3–5.0
Work	10.2	7.3–14.1
Pain	1.6	1.2–2.1
Psychic stability	2.5	1.9–3.2
Energy	5.0	3.8–6.5
Optimism	2.1	1.7–2.8
Socializing	3.0	2.3–3.9
PCS	(−10.0)	(0.67)*
MCS	(−8.6)	(0.78)*

ORs are estimated from logistic regression for multiple ordinal categories (except for work, which is dichotomous); β and associated SE are estimated from linear regression. β/SE is equivalent to a t-test. All models include the adjustment factors of age category (younger than 55/55 to 64/65 to 80/older than 80 years), gender, and income category (less than $10,000/$10,000 to $19,000/$20,000 to $29,000/$30,000 to $39,000/$40,000 to $49,000/more than $50,000). ORs greater than 1.0 indicate that people with stroke are more likely than controls to report functional limitations (reworded for ease of interpretation because all the ORs were greater than 1.0; 95% CIs that exclude 1.0 indicate statistical significance. β coefficients are interpreted as the difference between persons with stroke and controls on physical and mental health, taking into account age, gender, and income.
*Statistically significant $p < 0.05$.

Essentially, at this level, we have illustrated that the SF-12 and ICF capture the same constructs, providing evidence of face or content validity. Future work is planned to criterion-validate the coded list of functional status indicators generated from the SF-12 and other measures and to have each person's generated list rated directly validated by the patient him- or herself and by a therapist. Severity ratings and further validation are important next steps that will follow this proof-of-concept study.

One interpretation of the results presented here is that item-level data from the SF-12 tells more about function than do the global summary scores for physical and mental health. Although this seems obvious, the items themselves provide no usable information unless there is a mechanism to communicate the ratings about the functional status of an individual to (1) people responsible for treatment and (2) organizations responsible for resource management. The mechanism is a list of functional status indicators coded to be compatible with administrative data. This study demonstrates the feasibility of creating a coded list of functional status indicators from standardized questionnaires and that the information so obtained can be more informative than the usual method of reporting a standard outcome measure.

Conclusion

It was possible to translate most of the HRQL information contained in the SF-12 into ICF codes. The ICF coding

provided enhanced information about specific functional limitations experienced by persons with stroke, information that may assist with provision of services for persons with stroke and help to explain health discrepancies. The format of the ICF is more compatible with the structure of administrative health databases and would enrich these databases by providing important information on functional indicators that are known to contribute to service use and explain health outcomes. If the routine use of standard health measure questionnaires can be integrated into EHRs, then EHR systems may become a viable source of functional status information for the administrative databases used in population and health services research.

References ■

1. Malamateniou F, Vassilacopoulos G, Tsanakas P. A workflow-based approach to virtual patient record security. IEEE Trans Inf Technol Biomed. 1998;2:139–45.
2. Doolan DF, Bates DW, James BC. The use of computers for clinical care: a case series of advanced U.S. sites. J Am Med Inform Assoc. 2003;10:94–107.
3. Selby JV. Linking automated databases for research in managed care settings. Ann Intern Med. 1997;127:719–24.
4. Palmer RH. Process-based measures of quality: the need for detailed clinical data in large health care databases. Ann Intern Med. 1997;127:733–8.
5. Liang MH, Shadick N. Feasibility and utility of adding disease-specific outcome measures to administrative databases to improve disease management. Ann Intern Med. 1997;127:739–42.
6. Reilly CA. Examining the symptom experience of hospitalized patients using a pen-based computer. Proc AMIA Symp. 1999:364–8.
7. Ruland CM, Bakken S. Representing patient preference-related concepts for inclusion in electronic health records. J Biomed Inform. 2001;34:415–22.
8. Patrick DL, Erickson P. Health Status and Health Policy. Allocating Resources to Health Care. New York, NY: Oxford University Press, 1993.
9. Fitzpatrick R, Fletcher A, Gore S, Jones D, Spiegelhalter D, Cox D. Quality of life measures in health care. I: Applications and issues in assessment. BMJ. 1992;305:1074–7.
10. Swank AM, Robertson RJ, Deitrich RW, Bates M. The effect of acute exercise on high density lipoprotein-cholesterol and the subfractions in females. Atherosclerosis. 1987;63:187–92.
11. McKight L, Stetson P, Bakken S, Curran C, Cimino J. Perceived information needs and communication difficulties of inpatient physicians and nurses. J Am Med Inform Assoc. 2002;9(suppl): s64–9.
12. White TM, Hauan MJ. Extending the LOINC conceptual schema to support standardized assessment instruments. J Am Med Inform Assoc. 2002;9:586–99.
13. Chute CG, Cohn S, Campbell JR. A framework for comprehensive health terminology systems in the United States: development guidelines, criteria for selection, and public policy implications. J Am Med Inform Assoc. 2004;11:503–10.
14. World Health Organization. International Classification of Diseases, Disorders and Injuries, 10th revision. Geneva: World Health Organization, 1994.
15. National Center for Health Statistics. Classifications of Diseases and Functioning and Disability. Washington, DC: Centers for Disease Control and Prevention, Department of Health and Human Services, 2004.
16. World Health Organization. International Classification of Functioning, Disability, and Health. 2nd ed. Geneva: World Health Organization, 2001.
17. Iezzoni LI, Greenberg MS. Capturing and classifying functional status information in administrative databases. Health Care Financ Rev. 2003;24:61–76.
18. Clauser SB, Bierman AS. Significance of functional status data for payment and quality. Health Care Financ Rev. 2003; 24:1–12.
19. Ustun TB, Chatterji S, Kostansjek N, Bickenbach J. WHO's ICF and functional status information in health records. Health Care Financ Rev. 2003;24:77–88.
20. Weigl M, Cieza A, Harder M, et al. Linking osteoarthritis-specific health-status measures to the International Classification of Functioning, Disability, and Health (ICF). Osteoarthritis Cartilage. 2003;11:519–23.
21. Cieza A, Brockow T, Ewert T, et al. Linking health-status measurements to the international classification of functioning, disability and health. J Rehabil Med. 2002;34:205–10.
22. Mayo NE, Wood-Dauphinee S, Cote R, Durcan L, Carlton J. Activity, participation, and quality of life six months post-stroke. Arch Phys Med Rehabil. 2001;83:1035–42.
23. Ware JE Jr, Snow KK, Kosinski M, Gandek B. Health Survey: Manual and Interpretation Guide. Boston, MA: The Health Institute, New England Medical Center, 1993.
24. Ware JE Jr, Kosinski M, Keller SD. SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales. 2nd ed. Boston, MA: The Health Institute, New England Medical Center, 1995.
25. Finch E, Brooks D, Stratford PW, Mayo NE. Physical Rehabilitation Outcome Measures, 2nd ed. Hamilton: BC Decker Inc., 2002.
26. Nunnally J. Psychometric Theory, 2nd ed. New York, NY: McGraw-Hill, 1978.
27. Korner-Bitensky N, Wood-Dauphinee S. Barthel Index information elicited over the telephone. Is it reliable? Am J Phys Med Rehabil. 1995;74:9–18.
28. Collin C, Wade DT, Davies S, Horne V. The Barthel Index: a reliability study. Int Disabil Stud. 1987;10:61–3.
29. Roy CW, Togneri J, Hay E, Pentland B. An inter-rater reliability study of the Barthel Index. Int J Rehabil Res. 1988;11:67–70.
30. Scott SC, Goldberg MS, Mayo NE. Statistical assessment of ordinal outcomes in comparative studies. J Clin Epidemiol. 1997;50: 45–55.
31. Canadian Institute of Health Information. Available at: http: //secure.cihi.ca/cihiweb/. Accessed January 2004.
32. Kirshner B, Guyatt G. A methodological framework for assessing health indices. J Chronic Dis. 1985;38:27–36.
33. McColl E, Meadows K, Barofsky I. Cognitive aspects of survey methodology and quality of life assessment. Qual Life Res. 2003;12:217–8.
34. Bjorner JB, Ware JE, Kosinski M. The potential synergy between cognitive models and modern psychometric models. Qual Life Res. 2003;12:261–74.
35. Poissant L, Mayo E, Wood-Dauphinee S, Clarke A. The development and preliminary validation of a preference-based stroke index (PBSI). Health Qual Life Outcome. 2003;1:43.