Outcome Measures for Use with Home Care Clients

John N. Morris, *Hebrew Rehabilitation Center for Aged*, Iain Carpenter, *The University of Kent*, Katherine Berg, *Brown University*, and Richard N. Jones, *Hebrew Rehabilitation Center for Aged**

RÉSUMÉ

Avant les années 1990, la plupart des recherches portant sur l'évaluation des résultats des programmes de soins à domicile, notamment sur les efforts pour retarder le placement, maintenir le fonctionnement et assurer l'autonomie des aîné(e)s, ne reposaient pas sur des fondements théoriques. Les résultats attendus des programmes étaient rarement mentionnés. L'élaboration de nouvelles politiques exige maintenant une évaluation complète des besoins et une agrégation des données recueillies. Depuis qu'un nombre croissant de patients quittent l'hôpital avec des problèmes cliniques complexes et que les objectifs de réhabilitation deviennent plus étendus, on assiste à une explosion des services de soins à domicile. Les modèles de soins sociaux, bien que toujours en cours, ne constituent désormais qu'une mince part du marché des soins offerts à domicile. Dans cet environnement en mutation, la pertinence des programmes de soins dans les établissements ou des autres soins dispensés suite à une maladie grave soulève de nouvelles interrogations. On s'interroge également sur les déplacements des clients entre les diverses installations de soins suite à des maladies graves. Cet article décrit un ensemble de mesures fonctionnelles, comportementales et sociales proposées en lien avec l'évaluation des effets des programmes de soins faisant suite à une longue maladie. L'instrument normalisé de collecte de données utilisé est le Resident Assessment Instrument for Home Care (RAI-HC). L'article présente un résumé des mesures proposées et témoigne de la validité de groupes identifiés à partir d'un échantillon national des clients de soins à domicile. Les données soulignent les différentes caractéristiques des clients entre les organismes et attestent que le RAI-HC permet de recueillir des données fiables et valides pour décrire les populations et évaluer l'efficacité des programmes.

ABSTRACT

Prior to the past decade, much research examining outcomes of home care programs, including efforts at delaying institutional placement, maintaining function, and supporting independence, was atheoretical in character. Outcomes hoped for were often unobserved. New policy developments require comprehensive assessment of need and aggregation of this assessment information. As more and more patients leave hospitals with complex clinical problems and extensive rehabilitative goals there has been a corresponding explosion of home care services. Social care models, while they still exist, are becoming a smaller component of the overall home care market. In this changing environment, questions are now being asked concerning the appropriateness of the care programs in home care and other post-acute care settings. There are also concerns that need to be addressed about movement of clients between post-acute settings. In this paper, we describe a set of proposed functional, behavioural, and social outcome measures that are germane to evaluating the efficacy of programmatic efforts within the post-acute continuum. Data were collected with a standardized data collection instrument, the Resident Assessment Instrument for Home Care (RAI-HC). We provide data summarizing these proposed outcomes and evidence of known groups validity in a cross-national sample of home care clients. Data highlight the differing characteristics of clients across these agencies and provide evidence that this standardized data collection instrument can capture data that is reliable and valid for describing populations and evaluating program effectiveness.

Introduction

During the 1970s and 1980s, there was an extensive history of research that looked at outcomes of home care programs, focussing primarily on delaying institutional placement and supporting functional independence (Kemper, 1988). Much of this research was atheoretical in character, and the hoped for outcomes were often unobserved (Weissert, Cready, & Pawelak, 1988). In recent years in the United Kingdom and United States and elsewhere, there has been a shift in the paradigm underlying the provision of services in the home, with an increasing focus on technical nursing, rehabilitation and medical care. The drive behind this change in emphasis rests on new assumptions about post-hospital care. The average length of stay has declined and patient acuity at discharge has become a more pervasive concern.

As demands on home care have increased and home care advanced technically, the demand for valid indicators of effectiveness has increased. The answer to the question "who is getting what and with what effect" has become more urgent. Without answers to such questions, the emerging pressures to shrink the home care industry (as is the case in the United States) could have detrimental effects on segments of the client population. In the United States there has been an explosion of home care services, with an accompanying shift of responsibility from acute to post-acute care. Only recently has the negative backlash to this uncontrolled explosion of

care been felt in the community. Patients are leaving hospitals with complex clinical problems, extensive rehabilitative goals, and have been going into many post-acute settings. The appropriateness of these pathways is largely unknown.

In the United Kingdom, assessment of need and aggregation of assessment information have been seen as fundamental requirements for the success of new policy developments for home-based care. However, the field lacks clear guidance on how to achieve these goals and no systematic process for evaluation has been developed (Leicester & Pollock, 1996). In particular, community based alternatives to hospital have been given scant attention in medical research. There are many critical issues that have yet to be adequately addressed. Some of these include:

- information on the safety and effectiveness of new technologies in the home,
- the social, financial and psychological effects of providing acute care at home,
- monitoring the less strictly controlled home environment which lies outside institutional safeguards, and
- medical audit and quality assurance (Lafferty, 1996).

Increasingly, government funded home care requires satisfaction of precise eligibility criteria for access (Department of Health, 1995). In order to evaluate eligibility criteria, targeting goals must be specified and measures for determining the success of meeting goals developed for different segments of the population (Spector & Kemper, 1994). Tailoring services to the needs of the population and targeting delivery to those most in need are primary components of effective care delivery.

To have maximum utility, outcome measures should be embedded in the information that home care professionals routinely collect or use. This would have the benefit of minimizing the bureaucratic burden on care staff and care agencies as primary clinical data can be used for the secondary purposes of performance measurement. These data can form the basis of a database for informing future services and research (Carpenter & Bernabei, 1995).

The Resident Assessment Instrument for Home Care (RAI-HC) was designed to provide the basis of a comprehensive assessment of the needs of individuals cared for in their homes (Morris et al., 1997a). It was developed on the same principles as the MDS/RAI (Morris et al., 1990) which has been shown to be useful for monitoring indicators of outcome and quality of care in nursing home populations (Ljunggren, Phillips, & Sgadari, 1997; Phillips, Hawes, Mor, Fries, & Morris, 1996). The individual items in the RAI-HC cover a broad spectrum of functional domains; where possible, the items are compatible with measures in the RAI for nursing homes.

The RAI-HC is a standardized assessment tool for clinical use. Items in the instrument describe client performance and capacity in a variety of areas. It is designed for use by clinical professionals (nurses, social workers, or physicians). Assessments require direct questioning of the client and primary family caregiver (if available) and observation of the client in the home environment. Assessment may also require review of secondary documents. Ultimately, the assessor weighs available information and clinical judgement in recording ratings for the client. Assessment data are linked to care planning guidelines for problems, risk factors, and areas of potential benefit triggered by the assessment. These assessment protocols guide the assessor through best practice in developing a care plan for the identified problem, risk or potential. In this way the assessor derives direct benefit from use of the assessment while service providers, purchasers, and policy makers can aggregate the assessment data for the purposes of evaluating outcomes and population based service planning (Morris et al., 1997a).

In this paper we have selected a broad spectrum of RAI-HC items for use as outcome measures. Fifteen measures are proposed. In some instances, items stand alone as indicators (6 of the 15 outcome measures) or are collapsed into hierarchical indices (3 of the 15 outcome measures) or combined as summative scales (6 of the 15 outcome measures). The interassessor reliability for single item indicators and items comprising hierarchical indices have been published previously (Morris et al., 1997b) and are summarized below. In this report, we compare internal consistency reliability for measures composed as summative rating scales.

Method

During the course of usual home health care activities, home health care assessors from agencies in Ontario, Canada ($n = 410$), two United States sites (Michigan and Texas, $n = 1,070$ and $n = 410$, respectively), several areas within Japan ($n = 1,004$) and Italy ($n = 290$) collected RAI-HC assessments on home health care clients. Subjects in each cross-national site were unselected and represent a cross-section of home health care clients typical for each of the agencies adopting the RAI-HC. Some subjects were long-time recipients of home health care, others were new to the home health care agency at the time of assessment. These cross-sectional assessments were combined into a single data set. There is some variation in the content of different releases of the RAI-HC. Data presented in this analysis consider only those elements comparable across all sites (with the exception of services utilization and reason for assessment).

The inter-rater reliability of the RAI-HC has been reported earlier in a cross-national sample of different composition Morris et al., 1997b). Assessment areas relative to IADL self performance demonstrated excellent mean interassessor reliability (weighted kappa, $\kappa_w = 0.79$), as did IADL

self performance (κ_w = 0.77) IADL difficulty (κ_w = 0.75), and stamina (κ_w = 0.74). Areas relevant to memory and decision making also demonstrated excellent reliability (κ_w = 0.79), as did items relating to communication (κ_w = 0.84). Indicators of depression and anxiety demonstrated fair to good interassessor reliability (κ_w = 0.61). History of falls had excellent reliability (κ_w = 0.81) as did bladder continence (κ_w = 0.81). Indicators of elder abuse, including fear of family member, demonstrated excellent interassessor reliability (κ_w = 0.79). Health status indicators, including self-reported health, demonstrated good reliability (κ_w = 0.58) and the reliability of service utilization assessment was excellent (κ_w = 0.75). Overall, these previously published results suggest most RAI-HC domains have interassessor reliability coefficients suitable for comparison of groups. In the current analysis, we extend these results by considering the internal consistency reliability for proposed outcome measures constructed as summated rating scales.

Comparisons were performed by country and by known groups defined by a shared characteristic, problem or service. We considered five overlapping non-exclusive known groups. The data set is comprised of home care clients from several nations. In each nation, the selection pressures towards home care vary. Therefore, it is illustrative to see how the clinical outcome measures behave in known groups likely to be encountered by home health caregivers in any country. Means and variability about those means displaying patterns across groups that match expectations for those groups adds to the face validity of the outcome measures, and serves as evidence of known groups validity for these measures (c.f., (Bellelli, Frisoni, Bianchetti, & Trabucchi, 1997; Rossi, Wright, & Anderson, 1983; Spector, 1992)).

Towards this end, we have defined five non-exclusive groups based on known presenting characteristics. These groups describe with broad brushstrokes client populations similar with respect to expected care trajectory. More specifically, a *post-hospital group* is defined as subjects whose reason for referral to home health care services was post-hospital care, i.e. to assess any necessary care requirements following a stay in a hospital (Morris et al., 1997a). An *Alzheimer's disease (AD) or other dementia group* included subjects with a physician's diagnosis of probable AD or other dementia. A group *any neurological or psychiatric diagnosis* was defined, expanding the AD or other dementia group to include head trauma, multiple sclerosis, Parkinsonism, or any psychiatric diagnosis. A *high technology treatment group* was defined as including any subject receiving special treatment in the 14 days preceding assessment. These treatments included alcohol treatment, treatment for drug addiction, blood infusions, chemotherapy, cardiac rehabilitation, continuous positive airway pressure (CPAP), peritoneal dialysis (CAPD), renal dialysis, Holter monitor, IV infusion, medication by injection, ostomy care, oxygen therapy, radiation therapy, respiratory therapy or tracheotomy care. The fifth group, a

physical or occupational therapy group (OT/PT), was defined as those *subjects receiving occupational or physical therapy at any time in the 14* days preceding the RAI-HC assessment.

The RAI-HC collects more than 200 variables on each client. The current analysis considers 15 outcome measures. These measures are grouped into five domains: physical functioning, cognitive functioning, psychosocial functioning, clinical complexity, and services utilization.

The four physical functioning outcomes include a measure of IADL difficulty, IADL involvement, ADL dependence, and stamina. The IADL difficulty outcome is a hierarchical index capturing pattern of difficulty with ordinary housework, preparing meals and using the telephone. Ratings of clients are combined into a scale from 0–6, with higher values indicating greater IADL difficulty. A score of 0 indicates no difficulty in any of these three areas. Scores of 1–3 indicate some difficulty in one to three areas, respectively, and scores of 4–6 indicate great difficulty in one to three areas, respectively.

The IADL involvement outcome is a summated rating scale based on involvement with three IADL activities (ordinary housework, meal preparation, phone use). Each activity is scored 0–3, indicating independence (0), some help (1), full help (2) or performance of task by other (3). The IADL involvement scale range is 0–9 with higher values indicating greater dependence upon others for IADL tasks.

The ADL difficulty outcome is a summated scale that incorporates RAI assessments of difficulty with mobility, transfers, locomotion, dressing, eating, toileting, personal hygiene and bathing. Each area of potential difficulty is scored on a 0–4 scale, indicating independence (0), supervision (1), limited assistance (2), extensive assistance (3), and total dependence or "activity did not occur" (4). The total scale has a range of 0–32, with higher values indicating more overall ADL difficulty.

The stamina outcome is a summated scale based on four dichotomous indicators of physical capacity: out of the house infrequently (1 day a week or less), less than two hours of physical activity in the last week, unable to climb stairs on own, locomotion in home requires supervision or physical help. The scale has a 0–4 range, with higher values suggesting lower tolerance for physical activity.

The two cognitive functioning outcomes include the Cognitive Performance Scale (CPS©), a hierarchical index used to rate the cognitive status of nursing home residents (Morris et al., 1994) and a combination of two RAI-HC communication items assessing the client's ability to understand and make themselves understood. The CPS has a range of 0–6 with increasing values indicating more severe cognitive impairment. Each item in the communication outcome variable is scored 0–3, indicating the frequency (always/sometimes/rarely/never) in which the client is able to understand or make themselves understood. Combined, this summated

scale has a range of 0–6, with increasing values indicating poorer communication ability.

The three psychosocial functioning measures include mood disturbance, pattern of alcohol use, and fear of family member or other caregiver. The fear of family member or other caregiver measure is a dichotomous indicator based on a single RAI-HC item. The assessor is to interview the client in a non-threatening manner away from family, friends and caregivers and/or make observations based on interactions and quality of the interaction between the client and caregivers. The assessor attempts to determine if the client is afraid of the caregiver or withdraws (emotionally or physically) from the caregiver.

The alcohol use measure is a hierarchical index. It combines frequency of alcohol use with two items capturing the character of alcohol use. The measure is scored on a 0–3 range. Scores of 0 indicate no use in the previous 90 days, 1 indicates some use less than daily in the previous 90 days, 2 indicates daily use in the previous 90 days. Any alcohol use in concert with an admission on the part of the client or caregiver that the client had been told to cut down or required a drink early in the day results in a score of 3.

The mood disturbance measure is a summated rating scale based on RAI indicators of depression, anxiety and sad mood. Symptom clusters used for this measure include sad/depressed feelings, withdrawal from usual activities, tearfulness, anxious complaints, and persistent anger. Each item in this section is scored 0–2 indicating the frequency in which the indicator occurred in the past 30 days. A score of 0 indicates the assessor made a determination that no indicator was expressed in the past 30 days. A score of 1 indicates that some symptom was exhibited up to five days per week over the past 30 days, and a 2 indicates that some symptom was exhibited six or seven days per week over the past 30 days. The summary mood disturbance measure has a range of 0–10, with higher scores indicating greater symptom burden over the preceding 30 days.

Four clinical complexity measures are proposed. Two are dichotomous indicators: any history of falls in the previous 180 days, and verbal expression of feeling in poor health on the part of the client. Included in this domain is an ordinal index of continence, and a summative rating scale capturing pain. The assessor rates the client's continence over the preceding 14 days as continent (0), usually continent (1), occasionally continent (2), frequently incontinent (3) or incontinent (4). The pain measure is a summative rating scale based on three dichotomous indicators: frequent complaints of pain, unusually intense pain, and pain disruptive of usual activities.

The final outcome domain is services utilization. This is operationalized as hours of informal and formal services received over the past week. Although the RAI assesses formal services utilization over the past 14 days, the values are expressed as hours over one week for the purposes of this analysis. Formal services include home health aides, visiting nurses,

homemaking services, meals, volunteer services, physical therapy, occupational therapy, speech therapy, day care or day hospital, or social worker in the home. Similarly, informal IADL and ADL support provided by family or friends over the previous week (in hours, weekdays and weekends) is estimated by questioning the client and caregiver as available. Data describing the hours of formal service utilization are not available for clients from Japan and Italy. In addition, informal service data are unavailable for clients from Japan and Canada.

Analytic Method

Cross-national sites were compared in terms of subject characteristics, the internal consistency reliability of measures (for those measures composed as summative scales) and in terms of the distribution of the proposed outcome measures. Site means and proportions for subject characteristics and proposed outcome measures were compared to summary statistics reflecting all sites weighted equally. Comparisons of subject characteristics are summarized in Table 1. Proposed outcome measure means and proportions are summarized in Table 2. Table 3 presents standardized effect scores ((overall mean – national mean)/overall standard deviation) for each measure by site and known group for only those standardized differences that are greater than 0.20 standard deviations. This table provides a summary of differences relative to the overall sample for all measures on an equivalent scale. Standardized differences indicating a lower mean in the nation or client group relative to the overall total are shaded.

Results

Subject Characteristics
Subject characteristics of cross-national clients are summarized in Table 1. The overall sample is predominantly older persons (mean age 74.9 years over all sites); clients range in age from less than 24 years to more than 100 years of age. There are cross-national differences in the age distribution. The Japanese sample contained no subjects younger than age 40, and had a mean age of 77.4 years. The Texas site had a much larger proportion of clients less than 65 years of age (25.4%), relative to Italy (16.2%), Michigan (13.8%), Japan (11.9%) and Ontario (7.0%).

Overall, about one-third (31.3%) of the sample was male. The Japanese sample had significantly more men (40%), while Ontario and Michigan were represented by fewer men (24.1% and 26.4%, respectively).

The five known clinical groups were unevenly distributed across national sites. For example, post-hospital clients were not found in Texas and Japan. Thus, the post-hospital group contains subjects from Michigan, Ontario and Italy only. The four remaining clinical groups were represented by subjects from each cross-national site. The prevalence of AD and other dementias was greater among subjects from Italy (27.9%) relative to

Table 1

Subject characteristics by national site. Relative frequency distribution (per 100 subjects). Inter-RAI cross-national study of home health care clients (n = 3,633)

	Total (n = 3,633)	U.S.A. Michigan (n = 1,070)	U.S.A. Texas (n = 410)	Canada Ontario (n = 859)	Japan (n = 1,004)	Italy (n = 290)
Age (years)						
< 24	1.0	0.1	2.6	0.0	0.0	0.9
25–29	0.5	0.4	1.0	0.3	0.0	0.5
30–34	0.6	0.1	1.0	0.0	0.0	1.4
35–39	0.7	0.7	2.2	0.1	0.0	0.0
40–44	0.7	1.5	0.7	0.3	0.2	1.7
45–49	2.1	1.5	3.5	0.6	1.5	2.6
50–54	2.2	1.4	3.9	1.3	1.5	1.7
55–59	3.1	1.4	4.1	1.5	2.3	4.3
60–64	4.6	6.7	6.4	2.9	6.4	3.1
65–69	8.2	10.1	12.4	6.7	8.7	4.3
70–74	13.5	14.2	13.4	12.9	16.3	12.2
75–79	15.9	21.2	10.2	20.8	16.6	16.5
80–84	18.6	17.7	15.5	21.3	20.2	17.9
85–89	16.9	15.4	13.4	20.3	14.6	19.1
90–94	8.6	6.0	6.9	9.0	8.7	10.1
95–99	2.5	1.1	2.5	2.0	2.8	3.1
100+	0.3	0.1	0.3	0.1	0.2	0.5
mean age (years)	74.9	75.6	71.5***	79.0***	77.4*	76.8
Sex						
men	31.3	30.2	24.1**	26.4**	40.0***	35.5
women	68.7	69.8	75.9	73.6	60.0	64.5

Table 1 contd.

	Total (n = 3,633)	U.S.A. Michigan (n = 1,070)	U.S.A. Texas (n = 410)	Canada Ontario (n = 859)	Japan (n = 1,004)	Italy (n = 290)
Education[1]						
less than high school	68.0	54.0***	71.5***	57.2*		88.5***
high school	17.5	27.0	14.6	18.7		10.1
more than high school	14.5	19.0	13.9	24.1		1.4
Clinical Group[2]						
Post hospital care	7.2	9.0	0.0	16.6	0.0	11.0
AD or other dementia	15.7	17.6	11.7	11.4	15.7	27.9
Neurological/Psychiatric	27.1	34.8	28.0	23.6	16.6	43.4
High tech treatments	22.5	31.1	23.4	11.8	18.9	28.0
OT or PT	12.8	21.9	5.6	10.0	10.4	2.8

[1] – p-values relate to test of proportion with "less than high school" equal to overall sample.
[2] – statistical testing not reported for these descriptors.
* $p < .05$ difference in means (proportions) relative to total.
** $p < .01$ difference in means (proportions) relative to total.
*** $p < .001$ difference in means (proportions) relative to total.

subjects from the U.S. (Michigan 17.6%, Texas 11.7%), Ontario (11.4%) and Japan (15.7%). The same pattern was revealed for the any neurologic or psychiatric diagnosis group. The Michigan sample had the greatest proportion of subjects receiving high technology treatments (31.1%), followed closely by Italy (28.0%). The proportion of subjects receiving occupational or physical therapy varied considerably across site, ranging from a low of 2.5 per cent among those subjects from Italy to 21.9 per cent among those subjects from the Michigan site.

Cross-National Internal Consistency Reliability of Outcome Measures
Internal consistency reliability (Cronbach's alpha, α) was estimated for each site and for all sites combined (each site weighted equally) for measures composed as summated rating scales. Internal consistency reliability was not computed for proposed outcome measures that are single RAI items, and those composed as hierarchical indices. Previously published interassessor reliability estimates for such measures is summarized above. The IADL involvement scale had α estimate of 0.78 (each cross-national site weighted equally) and was 0.74, 0.70, 0.72, 0.82, and 0.77 in the Michigan, Texas, Ontario, Japan and Italy samples, respectively.

The ADL dependence outcome indicator also demonstrated very high internal consistency reliability cross-nationally ($\alpha = 0.95$) and within each site (0.93, 0.95, 0.88, 0.96, 0.95 in the Michigan, Texas, Ontario, Japan and Italy samples, respectively).

The stamina outcome scale demonstrated good internal consistency reliability across all sites ($\alpha = 0.70$) and within each site (0.61, 0.60, 0.61, 0.75, 0.80 for the Michigan, Texas, Canada, Japan and Italy samples, respectively). The IADL difficulty indicator is not constructed as a summated rating scale and therefore internal consistency reliability is not computed.

Among the cognitive indicators, the CPS is not constructed as a summated rating scale so the internal consistency reliability is not computed. The communication outcome indicator is constructed as a summated rating scale and the items constituting this outcome demonstrated an overall internal consistency reliability estimate of 0.89, and 0.77, 0.90, 0.78, 0.89, and 0.94 in Michigan, Texas, Ontario, Japan and Italy.

The mood disturbance measure is a summated rating scale, and demonstrates good internal consistency reliability ($\alpha = 0.68$ overall, and 0.69 0.78, 0.73, 0.72, 0.53 in Michigan, Texas, Ontario, Japan, and Italy, respectively). The pain measure achieved good internal consistency reliability in the overall sample ($\alpha = 0.75$) and cross-nationally (0.73, 0.69, 0.70, 0.51, 0.83, in Michigan, Texas, Ontario, Japan and Italy).

The comparison of means and proportions on the 15 outcome measures described above are presented in Table 2, below. Panel "a" contains cross-national comparisons; panel "b" contains comparisons across known clinical group.

Table 2a
Means (and standard deviation) for key outcomes by nation. Inter-RAI cross-national study of home health care clients ($n = 3,633$)

	Total[1]	U.S.A. Michigan	U.S.A. Texas	Canada Ontario	Japan	Italy
Physical Functioning						
IADL difficulty	3.32 (1.29)	3.22 (1.31)*	3.39 (1.25)	3.53 (1.51)***	3.41 (1.14)*	3.06 (1.23)***
IADL involvement	5.48 (2.94)	4.83 (2.81)***	5.64 (2.49)	3.66 (2.54)***	7.19 (2.7)***	6.10 (2.87)***
ADL dependence	8.19 (9.7)	5.35 (7.48)***	8.19 (9.24)	2.23 (4.88)***	12.49 (10.53)***	12.04 (10.7)***
stamina	2.33 (1.38)	2.33 (1.26)	2.41 (1.22)	1.62 (1.24)***	2.61 (1.42)***	2.70 (1.46)***
Cognitive Functioning						
CPS	1.60 (1.9)	1.36 (1.67)***	1.55 (1.92)	0.91 (1.36)***	2.08 (2.13)***	2.10 (2.06)***
communication	0.91 (1.6)	0.66 (1.21)***	0.87 (1.54)	0.41 (0.98)***	1.5 (1.87)***	1.10 (1.93)**
Psychosocial Functioning						
mood disturbance	1.36 (1.9)	1.64 (2.01)***	1.18 (2.01)*	1.06 (1.74)***	1.16 (1.79)**	1.76 (1.81)***
alcohol use	0.23 (0.62)	0.10 (0.48)***	0.10 (0.37)***	0.20 (0.54)***	0.20 (0.54)***	0.70 (0.99)***
fear of relative	0.02 (0.13)	0.02 (0.12)	0.02 (0.15)	0.01 (0.08)***	0.02 (0.15)	0.02 (0.15)
Clinical Complexity						
any fall history	0.40 (0.49)	0.51 (0.5)***	0.34 (0.47)***	0.36 (0.48)*	0.28 (0.45)***	0.52 (0.5)***
pain	1.19 (1.13)	1.45 (1.22)***	1.45 (1.27)***	1.32 (1.14)**	0.63 (0.79)***	1.12 (0.97)*
incontinence	1.15 (1.49)	1.13 (1.37)	1.26 (1.57)	0.68 (1.15)***	1.45 (1.62)***	1.22 (1.58)
felt in poor health	0.41 (0.49)	0.48 (0.5)**	0.39 (0.49)	0.37 (0.48)	0.36 (0.48)**	0.48 (0.5)**
Services Utilization						
formal services	7.5 (11.9)	5.0 (10.7)***	12.5 (14.4)***	8.6 (13.7)	.	.
informal services	36.8 (51.5)	48.6 (59.1)***	19.5 (52.6)***			47.5 (47.2)***

Table 2b

Means (and standard deviation) for key outcomes by known group membership. Inter-RAI cross-national study of home health care clients (n = 3,633)

	Post-hospital	AD or Dementia	Any neuro/psychiatric Problem	High Technology Treatments	OT or PT
Physical Functioning					
IADL difficulty	3.08 (1.43)**	3.29 (0.88)	3.31 (1.13)	3.22 (1.30)*	3.42 (1.11)
IADL involvement	5.05 (2.81)*	7.51 (2.24)***	6.35 (2.83)***	5.33 (2.78)	5.68 (2.62)
ADL dependence	7.38 (7.89)	12.91 (10.32)***	10.37 (10.19)***	7.96 (9.69)	8.40 (8.79)
stamina	2.39 (1.46)	2.67 (1.42)***	2.43 (1.41)*	2.39 (1.35)	2.40 (1.29)
Cognitive Functioning					
CPS	1.47 (1.90)	3.65 (1.79)***	2.58 (2.07)***	1.39 (1.89)**	1.45 (1.74)
communication	0.70 (1.37)*	2.3 (1.98)***	1.50 (1.88)***	0.75 (1.59)**	0.81 (1.32)
Psychosocial Functioning					
mood disturbance	1.52 (1.83)	1.81 (2.09)***	1.90 (2.18)***	1.67 (2.02)***	1.53 (2.16)
alcohol use	0.26 (0.66)***	0.26 (0.73)***	0.27 (0.73)***	0.26 (0.7)***	0.14 (0.51)
fear of relative	0.05 (0.22)*	0.06 (0.23)***	0.04 (0.19)**	0.02 (0.14)	0.02 (0.13)
Clinical Complexity					
any fall history	0.46 (0.50)	0.46 (0.50)*	0.49 (0.50)***	0.42 (0.49)	0.53 (0.50)
pain	1.23 (1.07)	0.85 (1.03)***	1.08 (1.11)***	1.39 (1.18)***	1.33 (1.17)*
incontinence	1.02 (1.48)	1.87 (1.64)***	1.51 (1.59)***	1.03 (1.44)*	1.10 (1.39)
felt in poor health	0.49 (0.50)*	0.31 (0.46)***	0.40 (0.49)	0.55 (0.50)***	0.41 (0.49)
Services Utilization					
formal services	5.5 (10.1)*	8.9 (16.2)	8.1 (13.8)	14.8 (22.3)***	17.3 (21.3)***
informal service	34.3 (41.1)	64.8 (70.5)***	50.7 (64.6)***	41.6 (50.1)*	37.2 (47.1)

1 – overall Inter-RAI sample means computed with each site weighted equally.

* P < .05 difference in means (proportions) relative to total.

** P < .01 difference in means (proportions) relative to total.

*** P < .001 difference in means (proportions) relative to total.

Physical Functioning

The Michigan site had significantly less ADL impairment (mean ± standard deviation, 5.35±7.48) relative to the overall mean (8.19±9.7). The Michigan sample also had significantly less IADL difficulty (3.22±1.31) and IADL involvement (4.83±2.81) relative to the overall total (3.32±1.29 and 5.48±2.94, respectively). It is somewhat surprising that, given this lower functional impairment for the Michigan sample, the mean level of physical activity (as captured by stamina (2.33±1.26)) was not similarly lower relative to the overall mean (2.33±1.38). The Ontario sample had even lower levels of ADL (2.23±4.88) and IADL impairment (3.53±1.51 for difficulty and 3.66±2.54 for involvement), and had less stamina impairment (1.62±1.24). Relative to the overall mean, the Japanese and Italian samples demonstrated profound ADL difficulty (12.49±10.53 and 12.04±10.70, respectively), IADL difficulty (3.41±1.14 and 3.06±1.23, respectively) and IADL involvement (7.19±2.7 and 6.10±2.87, respectively).

IADL difficulty was common among all clinical groups. The clinical group with the most impaired ADL functioning, relative to the overall cross-national mean, was the AD or other dementia group with a mean impairment score of 12.91 (±10.32). The any neurological or psychiatric diagnosis group also had significantly greater ADL difficulty (10.37±10.19) relative to the overall cross-national mean. The other clinical groups did not have a level of ADL difficulty significantly different from the overall mean. The post-hospital group had significantly lower mean IADL difficulty (3.08±1.43) relative to the cross-national mean, as did the high technology treatment group (3.22±1.30).

Cognitive Functioning

Cognitive functioning, as reflected by the CPS score, was poorest in sites already described as having more impaired physical functioning clients (Italy and Japan; mean CPS 2.08±2.13 and 2.10±2.06, respectively). Level of communication difficulty was also greatest in these two countries. The Japanese sample had a mean communication score of 1.50 (±1.87) and the Italian sample a mean of 1.10 (±1.93) relative to the grand mean of 0.91 (±1.6).

The mean CPS score for those with AD or other dementia (3.65±1.79) was significantly greater than the overall sample mean. The same was true for the group with any neurological or psychiatric diagnosis (2.58±2.07) while the opposite was true for those receiving high technology treatments (1.39 ±1.89). Communication difficulties followed a similar pattern across known clinical groups.

Psychosocial Functioning

Mood disturbance varied considerably across sites. Respondents from Michigan and Italy tended to have a greater burden of symptoms relative to respondents from Texas, Ontario and Japan. Mood disturbance was high in countries where fall history and poor self-rated health were highly

prevalent. Fear of a relative is very rare in all sites. The overall prevalence is 2 per cent in this sample of home care clients.

Clients from the Italian site demonstrated a particularly high mean on the alcohol use outcome variable (0.70±1.81) relative to the mean for all sites combined (0.23±0.63). Further examination of this hierarchical index revealed this effect to be caused by a markedly greater frequency of daily alcohol use in Italy (25%) relative to the other sites (4% in Ontario, 3% in Japan, and 1% in the U.S. samples).

Among clinical groups, mean levels of mood disturbance are elevated among those with AD or other dementias (1.81±2.09), any neurologic or psychiatric diagnosis (1.90±2.18) and those receiving high technology treatments (1.67±2.02) relative to the overall sample mean (1.36±1.90). Mean level of alcohol disturbance is significantly elevated among all known clinical groups relative to the total sample, with the exception of the OT/PT group.

Clinical Complexity
Despite a relatively high level of physical functioning impairment, the Japanese sample had the lowest proportion with any fall history (28%, relative to 40% over all sites). Clients with a history of falls made up the majority of the Michigan (51%) and Italian samples (52%). These two sites also had the largest proportion of subjects feeling in poor health (48% each). The sample from Japan had a very low level of pain (0.63±0.79) relative to the overall sample (1.19±1.13).

Each of the known clinical groups had a higher prevalence of fall history relative to the overall sample. These effects reached conventional levels of statistical significance for the AD or other dementia group and the any neurological or psychiatric diagnosis group. Mean pain levels were significantly lower among AD or other dementia subjects (0.85±1.03) relative to the overall sample mean. The AD and other dementia group was less likely to feel in poor health (31%) relative to the overall sample (41%).

Services Utilization
Service utilization data were not available for all sites. The Japanese site did not report utilization data, and the Canadian site did not report estimates of informal services utilization. Not surprisingly, the high technology and OT/PT groups have the highest mean level of formal service utilization among the clinical groups. Informal services utilization was very high among those subjects with AD or other dementia, with a mean of 64.8 hours over the past two weeks (±70.5) relative to an overall sample (36.8 ±51.5 hours). The magnitude of the standard deviations reflect the markedly skewed nature of these data.

Standardized Effects
The standardized effects reported in Table 3 provide a means for grading the importance of each measure across site and clinical group. The stand-

Table 3
Standardized difference relative to (weighted) cross-national mean, if greater than |.20|, for key outcomes by site and known group membership. Inter-RAI cross-national study of home health care clients (*n* = 3,633)

	U.S.A. Michigan	U.S.A. Texas	Canada Ontario	Japan	Italy	Post-hospital	AD/ dementia	Neuro/ psyc. problem	High tech. treatments	OT or PT
Physical Functioning										
IADL difficulty	-0.22***				-0.20***					
IADL involvement	-0.29***		-0.62***	0.58***	0.21***		0.69***	0.29***		
ADL dependence			-0.61***	0.44***	0.40***		0.49***	0.22***		
stamina			-0.52***	0.20***	0.26***		0.24***			
Cognitive Functioning										
CPS			-0.36***	0.25***	0.26***		1.08***	0.51***		
communication			-0.31***	0.37***			0.87***	0.37***		
Psychosocial Functioning										
mood disturbance	-0.21***				0.21***		0.24***	0.28***		
alcohol use pattern		-0.21***								
fear of relative					0.75***		-0.20***		0.28	
Clinical Complexity										
any fall history	0.21***			-0.25***	0.25***					
pain	0.23***	0.23***		-0.50***			-0.30***			0.27***
incontinence			-0.31***	0.2***			0.48***	0.24***		
felt in poor health							-0.20***		0.28***	
Services Utilization										
formal services	-0.21***	0.42***								
informal services		0.23***	-0.34***		0.21***		0.54***	0.27***	0.62***	0.82***

* *P* < .05 difference in means (proportions) relative to total.
** *P* < .01 difference in means (proportions) relative to total.
*** *P* < .001 difference in means (proportions) relative to total.

ardization process places all measures on the same metric, with a mean of zero and unit variance. Effects indicating a mean level below the overall sample mean are shaded. Table 3 provides a useful visual presentation of the major differences in the types of persons encountered by home health care agencies across sites. Japan and Italy have the most impaired populations, Ontario the least impaired.

As expected, the AD and other dementia group has a very large standardized mean for the CPS outcome variable (1.08). This standardized effect is the largest of any cross-national or clinical group and across all measures. Table 3 further illustrates the very high level of alcohol use pattern for the Italian subjects, the very high formal services utilization for OT/PT group, and the low level of pain expressed by the Japanese sample.

Discussion

The proposed outcome measures, derived from the comprehensive Resident Assessment Instrument for Home Care (RAI-HC), are equally reliable in terms of internal consistency cross-nationally. Known clinical groups demonstrate expected divergence and add to the validity of these measures. Further, observed cross-national variability of these outcome variables supports their validity. For example, finding a low prevalence of fall history in the Japanese sample is supported by previously reported cross-national work and has been attributed to cultural differences in lifestyle and bowel evacuation behaviour (Lipsitz et al., 1994). Cross-national variability in personal activities of daily living is associated with corresponding variability in difficulty with performance of instrumental activities of daily living, cognitive difficulties, and mood disturbance. The fear of family member item has a very low base rate in all sites. This finding may reflect a truly low prevalence, but more likely reflects difficulty in measuring this sensitive topic. Overall, findings support the reliability and validity of the proposed outcome measures cross-nationally.

While social and political factors influence the base rates of these proposed outcome measures cross-nationally, cultural factors may also influence the meaning of some of the outcome measures. In particular, we note the high mean level on the alcohol use outcome indicator found in the Italian data. The proposed alcohol use disturbance scale may not be a valid indicator of increasing pathological use of alcohol cross-nationally. There was wide cross-national variability in the relative number of "2" responses (daily use). One-quarter of the Italian home care clients fell into this category, compared to about 4 and 3 per cent in Ontario and Japan, respectively, and less than 2 per cent for the United States home care samples. This effect could be due to cultural differences in alcohol use or willingness to admit to and record alcohol use. This discrepancy highlights difficulties in the use of this measure cross-nationally. Nevertheless, there is little indication that other outcome measures suffer the same limitation.

Because the RAI-HC is a standardized assessment instrument designed for routine use in home care, outcome data such as we have demonstrated can be gathered from routine clinical practice and used for comparing outcomes between populations. The presented RAI-HC outcome measures are comparable across cultures, and offer an opportunity to compare outcomes and services cross-nationally. This may enable comparison of the effectiveness of different service provision models in different countries. The comprehensive assessment helps identify subgroups of similar clients whose outcomes can be compared.

The current findings highlight the breadth of coverage offered by the RAI-HC. This makes the tool clinically useful for a wide variety of client populations, ranging from medically complex clients needing close attention to relatively well elders who receive and require less formal support. The data also highlight the fact that differences in case mix should be considered when making regional or international comparisons. Fortunately, the comprehensive nature of this instrument allows for adequate control of potentially confounding covariates, such as age, sex, presenting problem and medical comorbidities in planning such analyses.

References

Bellelli, G., Frisoni, G.B., Bianchetti, A., & Trabucchi, M. (1997). The Bedford Alzheimer Nursing Severity scale for the severely demented: validation study. *Alzheimer Disease and Associated Disorders, 11*(2), 71–77.

Carpenter, G., & Bernabei, R. (1995). Database needs and practical models: is a minimum dataset and common database possible and/or desirable. In L. Rubenstein, D. Wieland, & R. Bernabei (Eds.), *Geriatric assessment technology: The state of the art.* Milan: Editrice Kurtis.

Department of Health. (1995). *NHS responsibilities for meeting continuing health care needs.* London: Department of Health.

Kemper, P. (1988). The evaluation of the National Long Term Care Demonstration. 10. Overview of the findings. *Health Services Research, 2*(6), 420–427.

Lafferty, G. (1996). Community based alternatives to hospital rehabilitation services: a review of the evidence and suggestions for approaching future evaluations. *Reviews in Clinical Gerontology, 6*, 183–194.

Leicester, M.C., & Pollock, A.M. (1996). Community care in South Thames (West) Region: is needs assessment working? *Public Health, 110*(2), 109–113.

Lipsitz, L.A., Nakajima, I., Gagnon, M., Hirayama, T., Connelly, C.M., & Izumo, H. (1994). Muscle strength and fall rates among residents of Japanese and American nursing homes: an international cross-cultural study. *Journal of the American Geriatrics Society, 42*(9), 953–959.

Ljunggren, G., Phillips, C.D., & Sgadari, A. (1997). Comparisons of restraint use in nursing homes in eight countries. *Age and Ageing, 26*(suppl. 2), 43–47.

Morris, J., Bernabei, R., Ikegami, N., Gilgen, R., Fries, B., Steel, K., & Carpenter, I. (1997a). *RAI-Home Care (RAI-HC)(C) Assessment Manual.* Washington, DC: InterRAI Corporation.

Morris, J.N., Fries, B.E., Mehr, D.R., Hawes, C., Phillips, C., Mor, V., & Lipsitz, L.A. (1994). MDS Cognitive Performance Scale. *Journals of Gerontology, 49*(4), M174–M182.

Morris, J.N., Fries, B.E., Steel, K., Ikegami, N., Bernabei, R., Carpenter, G.I., Gilgen, R., Hirdes, J.P., & Topinkova, E. (1997b). Comprehensive clinical assessment in community setting: applicability of the MDS-HC. *Journal of the American Geriatrics Society, 45*(8), 1017–1024.

Morris, J.N., Hawes, C., Fries, B.E., Phillips, C. D., Mor, V., Katz, S., Murphy, K., Drugovich, M.L., & Friedlob, A.S. (1990). Designing the national resident assessment instrument for nursing homes. *Gerontologist, 30*(3), 293–307.

Phillips, C., Hawes, C., Mor, V., Fries, B., & Morris, J. (1996). *Evaluation of the Nursing Home Resident Assessment Instrument*. Washington: Health Care Financing Administration.

Rossi, P.H., Wright, J.D., & Anderson, A.B. (1983). *Handbook of survey research*. New York: Academic Press.

Spector, P.E. (1992). *Summated rating scale construction: an introduction*. Newbury Park, CA: Sage Publications.

Spector, W.D., & Kemper, P. (1994). Disability and cognitive impairment criteria: targeting those who need the most home care. *Gerontologist, 34*(5), 640–651.

Weissert, W.G., Cready, C.M., & Pawelak, J.E. (1988). The past and future of home- and community-based long-term care. *Milbank Quarterly, 66*(2), 309–388.