J Tokyo Wom Med Univ
75 (12) 489～504 (2005)

Model Specification and Testing of Outcome Indicators Used for Assessment of Healthcare Service for Home-Care Neurology Patients

Keiichi ITO[1)2)], **Hiromi WATANABE**[1)3)], **Makoto IWATA**[1)],
Shoichi SASAKI[1)] and **Shinichiro UCHIYAMA**[1)]

[1)]Department of Neurology (Director: Prof. Makoto IWATA), Tokyo Women's Medical University, School of Medicine
[2)]Division of Community Health Nursing and [3)]Division of Internal Medicine,
Tokyo Women's Medical University, School of Nursing

(Accepted Sept. 30, 2005)

In order to assess the effectiveness of healthcare services for home-care neurology patients and families, an outcome indicator was developed based on the degree of difficulty in performing daily life activities. Furthermore, using structural equation modeling (SEM) and general linear model, we examined the indicators' construct validity and predictive validity. To test the construct validity of the outcome indicators, we examined whether the second-order factor model was established or not using SEM. The outcome indicators consisted of five sub-indicators: 1) Anxiety about disease and disability indicator, 2) Family care burden and strain indicator, 3) Motor dysfunction indicator, 4) Appearance of symptom indicator, and 5) Interference in social network utilization indicator, resulting in high construct validity. The result of multiple indicators model indicated that all of the indicators influenced the HRQOL (SF-36) two years after baseline survey. The aspect of SF-36 on which outcome indicators have their influences was different for each indicator. Based on whether the scores of outcome indicators were improved or not in two years, the subjects whose degree of difficulty in performing daily living activity increased in the two years showed a remarkable decrease of SF-36 domains relating to role and physical functioning.

Key words: outcome indicator, validity testing, health-related quality of life, quality of home care, neurology patients

Introduction

The growth in home health care services in Japan since the 1990s and the enactment of long-term care insurance (April 2000) have resulted in a large increase in the number of home-care beneficiaries, primarily diagnosed with neurological and cerebrovascular diseases. With the shift of care from public facilities to the home milieu, the critical need to assess the effectiveness of home care services in our society has been underscored. In particular, because neurology patients and their family caregivers have a variety of difficulties performing daily living activities, and suffer from long-term diseases, it is important to assess the effectiveness of the home care services provided to these patients and their families over a long period of time.

In recent years, the quality of services has been an increasing focus in the US health care system. This emphasis has included efforts to quantify and analyze the outcome of care[1)]. For home care, the focus on outcome has resulted in new federal requirements that home health agencies participating in Medicare collect and report patient data using a single core set of measures specified in the outcome assessment and information set (OASIS; Health Care Financing Administration 1999)[2)]. In contrast, only a few studies in Japan have reported on the outcome of home care for the client, and explored an evaluation method for the effectiveness of home care in practice[3)4)].

Most home care services under Medicare in the US, however, are subject to providing short-term care, because of the strictness of Medicare's eligibility criteria for coverage[5)]; OASIS is, therefore, designed to assess the outcome for the client at 60-day intervals. Although most home care in our society is long-term care and most home-care patients have chronic conditions such as neurological diseases, it is necessary to develop a new instrument for assessing the outcome of care in the long term.

Thus, in order to assess the effectiveness of health care services for home-care neurology patients who are receiving long-term care, and their families, we developed a multi-dimensional outcome indicator based on the degree of difficulty performing daily living activities. Furthermore, using psychometric methods and structural equation modeling (SEM), we examined the indicators' reliability, construct validity and predictive validity.

Subjects and Methods

Procedure of data collection

In the current study, a two-fold postal survey for home-care neurology patients and their families, consisting of a baseline survey and a follow-up survey 2 years later, was carried out. For each survey, we mailed a cover letter and a set of questionnaires, with a postage-paid reply envelope. The cover letter explained the purpose and procedures of the current study and the option to refuse to participate. A document of informed consent was also sent to the individuals, and written consent was obtained. Return of the questionnaire was also considered as consent to participate. In addition, telephone contact was used to instruct individuals who required assistance in responding to the questionnaires. The questionnaires returned were linked with medical records regarding diagnosis and treatment status. Before the actual research was conducted, this survey project was approved by the ethics committee board of Tokyo Women's Medical University.

Sample characteristics

We conducted the baseline survey by mail for patients over 20 years of age, who were discharged from the Neurology Word at a university hospital between April 1995 and March 2000, and their families. As shown in the flowchart (Fig. 1), 504 responses were received, resulting in a participant rate of 49% and the measurement of 463 respondents, excluding the cases of patients whose deaths were confirmed during the survey period and of insufficient response content, were used for analysis. Of the 463 respondents, 54.4% were men with a mean age of 63.4 (range = 20 to 94, SD = 15.3) and 45.6% were women with a mean age of 57.7 (range = 20 to 92, SD = 16.1). When comparing the profiles of the respondent group and non-respondent group from the baseline survey, the ages of the non-respondent group with a mean age of 56.3 (SD = 19.0) at the time of survey were significantly younger than those of the respondent group ($p < 0.001$), whereas there was no statistical significance for number of days in the hospital, gender, and the proportion of profiles of each disease group between the respondent and non-respondent groups. The breakdown of diseases for the respondent group was: cerebrovascular diseases, 28.4%; degeneration and demyelinating diseases such as amyotrophic lateral sclerosis, Parkinson's disease, spinocerebellar degeneration and multiple sclerosis, 21.9%; peripheral nerve diseases and myopathy such as polyneuropathy, myositis and severe myasthenia gravis, 19.9%, and others, 29.8%.

Two years after the baseline survey, we conducted a follow-up survey (response rate: 51.6%) on the same subjects who agreed to participate in the survey, and used the data from 201 responses for analysis. Of the 201 subjects, 52.7% were men with a mean age at the baseline survey of 63.2 (range = 21 to 83, SD = 13.8) and 47.3% were women with a mean age of 59.6 (range = 20 to 84, SD = 14.1). There were no statistically significant differences between the 201 respondents and non-respondents on the variables of gender, age, length of hospitalization, and the proportion of profiles of each disease group, whereas the level of physical disability measured by the activities of daily living (ADL) scale was lower in non-respondents compared with respon-

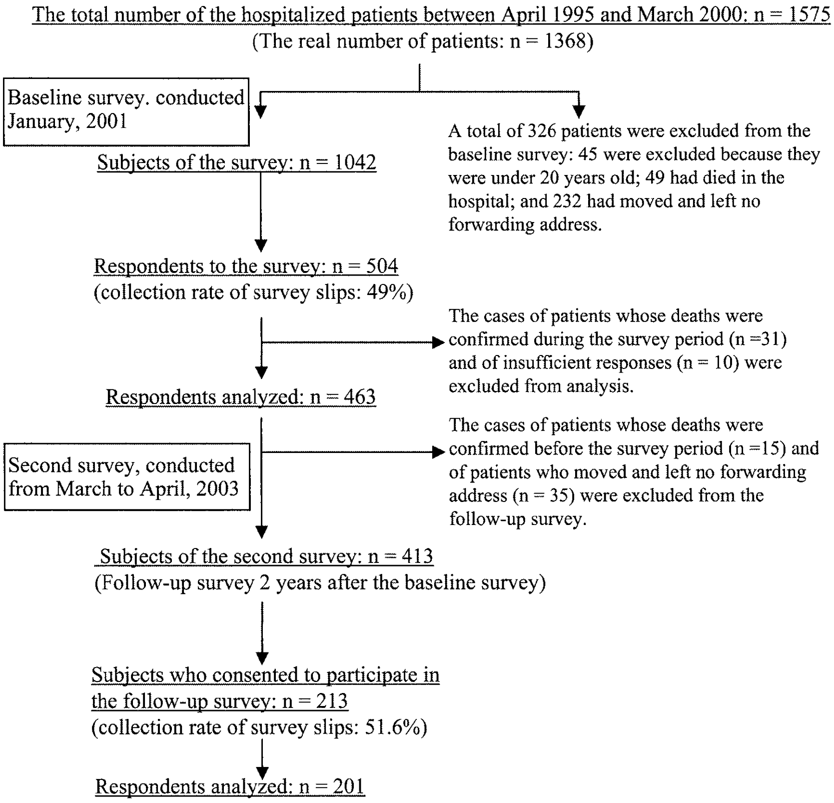

Fig. 1 Flowchart of the baseline and the follow-up surveys

dents (p = 0.038).

Outcome indicators for assessing the effectiveness of home care service for neurology patients

1) Definition of outcome indicator

The outcome indicator was defined by three types of outcome, according to Shaughnessy et al[6]: end-result outcome, intermediate-result outcome, and utilization outcome. End-result outcomes refer to changes in functional ability, physiologic condition, symptom distress, and emotional condition. Intermediate-result outcomes reflect a quantified non-functional outcome of care and can be pivotal in attaining certain end-result outcomes (e.g., a dichotomy reflecting change in the extent of family caregiver strain is an intermediate-result outcome). Utilization outcomes are a quantification of the health services that are potentially attributable to home care under consideration.

2) Scale development

We started the formulation of a questionnaire, including items indicating multi-dimensional outcomes, from a hypothetical measurement scale. As the first procedure, we started with the content analysis of qualitative data obtained from our previous study[7], which was related to the problematic events which neurology patients and family members experienced while giving home care. Based on the degree of difficulty in performing daily living activities obtained from these analytical results, and the relevant literature by Kramer[8] which presented a conceptual framework of Medicare quality indicators of home health care for neurological conditions, we created a measurement scale for multidimensional outcome indicators consisting of 30 items in total which were designed tentatively to cover five dimensions. The five dimensions for assessing the effectiveness of home care services for neurological patients and their family caregivers are changes in: 1) functional status measured by ADL and instrumental activities of daily living (IADL), 2) health status signs and symptoms, 3) family/caregiver strain, 4) unmet needs, and 5) utilization. The category of satisfaction was not included in the initial items, since global satisfaction measures tend to be influenced by many factors unrelated to the quality of home health care services[9].

The measurement scale, consisting of 30 items, used a four-point Likert-type response format, scored as: 1 = no problem, 2 = small problems, 3 = considerable problems and 4 = big problems for each item, for the clarity and ease of administration of both patients and their families. It was designed to show a higher degree of difficulty performing daily living activities with higher scores.

Measures

The following measurement scale for ADL was included to assess the functional status of neurology patients. Health-related quality of life (HRQOL) measurement was also included to assess the predictive validity of the newly developed outcome indicators.

1) Measurement of activities of daily living

Independence in ADL was determined by a partially modified Katz's index of independence in ADL developed by Katz[10] to measure the physical ability to function of the study participants. Katz's index is the most widely used of all functional assessment indices in studies determining the condition of ambulatory patients[11]. For the current study, independence was determined in six activities: bathing, dressing, waking, communication, continence, and feeding. Although through a series of questions from the original scale, participants were rated on a 3-point scale of independence for each activity, we revised the scoring system to count the number of activities in which the individual was dependent, measured on a scale of 0 to 6 (0 = independent in all six functions, 6 = dependent in all six functions).

2) Health-related quality of life measurement

HRQOL was assessed using the Medical Outcome Study 36-Item Short Form Health Survey, version 1.2 (SF-36)[12]. The SF-36 is a generic, self-administered survey, which has been widely used for varying chronic conditions, such as neurological diseases. This questionnaire consists of 36 questions, from which eight different domains can be calculated: physical functioning (PF), role-physical (RP), bodily pain (BP), general health (GH), vitality (VT), social functioning (SF), role-emotional (RE), and mental health (MH). The time frame is given as 'during the past 4 weeks'. Scores of the SF-36 range from 0 to 100, with the maximum score of 100 indicating the best possible health state and 0 indicating the worst health. The reliability and validity of the Japanese version of this scale have been supported by several studies[13][14]. Published sex and age-specific norms of the SF-36 are available from the general population in Japan.

Statistical analysis

All statistical analyses were performed with SAS statistical software package version 8.2™[15] and SPSS version 11.5[16] for Windows. In descriptive statistics, frequencies, means, and standard deviations were computed to obtain a sample profile on demographic and disease-related variables. Continuous variables were compared using student's t-test. To compare the categorical data, chi square test with Yates' correction was used. Pearson product-moment correlation coefficient was also used for this analysis. For each statistical analysis, cases with missing values were deleted from relevant analysis. A significance level of 0.05 was used for statistical tests, unless otherwise stated.

1) Internal consistency reliability

The internal consistency reliability of the outcome indicators was determined by calculating Cronbach's alpha coefficient, which provides an indication of the degree of convergence among different items hypothesized to represent a unified construct. Scales with reliabilities of more than 0.70 are recommended for the purpose of comparing groups of patients, whereas a higher reliability criterion of 0.90 is recommended for greater precision in analyzing individual patient scale scores[17].

2) Factorial validity

To analyze the scale structure of the outcome indicators, we used the psychometric method and explanatory factor analysis (EFA) with oblique promax rotation accompanied by the maximum likelihood method. Before EFA, in order to enhance internal consistency, unnecessary items were excluded; items with a manifestation frequency of less than 5% or more than 95% were eliminated from this analysis. An item of either which showed a correlation coefficient of more than 0.70 between two items was also eliminated. The number of factors in our procedure using EFA was fixed at five according to the initial hypothesis.

3) Construct validity

SEM and maximum likelihood estimation procedures using the CALIS procedure of SAS were employed to examine the construct validity of the outcome indicators. In order to test the construct validity of the outcome indicators, we examined whether the second-order factor model was established or not using SEM. In addition, to test homogeneity of the second-order factor model constituted above, we prepared the six datasets that divided each into two groups according to the three kind of patient's profiles: gender, age-class (under 65 years old/65 years old and over), and disease group (neurological/cerebrovascular). Parameter estimates and values of fit indices for each of the six datasets were computed based on the second-order factor model constituted from SEM. The reason SEM was used is that it is the best model for implementing factor analysis and regression analysis simultaneously, and that it allows investigators to test a prespecified a priori relationship and to determine if a reasonable fit exists between the five factors model and the raw data. SEM also has the ability to incorporate latent variables, which are hypothesized and unobserved concepts that can only be approximated by observed or measured variables, into the analysis[18].

A standardized coefficient was used to estimate causal effects. Standardized coefficients allow comparison of variables with different units of measurement. In the present study, five different fit indices are reported; chi-square likelihood ratio statistic (Chi-square/d. f.), goodness of fit index (GFI), adjusted goodness of fit index (AGFI), root mean square error of approximation (RMSEA) estimate, and the comparative fit index (CFI) were used to estimate overall model fit.

Chi-square likelihood ratio statistic[19] assesses the magnitude of the discrepancy between the sample and fitted covariance matrices. The chi-square statistically tests the lack of fit. A non-significant chi-square is desired because it indicates agreement between the proposed model and the data. That is, there is no significant discrepancy between the sample and the fitted model[20].

GFI indexes the relative amount of observed variance and covariance accounted for by the model[21]. AGFI is goodness of fit index adjusted for degrees of freedom. Both GFI and AGFI range from 0 to 1.00, with the former indicating the absence of model fit and the latter indicating perfect model fit; Values of 0.90 or above usually indicate good model fit.

RMSEA is a measure of the discrepancy between the observed and model implied covariance matrices per degree of freedom, suggesting that values of RMSEA of 0.05 or less indicate good fit, and less than 0.08 indicate an adequate fit[22][23].

CFI ranges from 0 to 1.00, with 0 indicating a poor fit, 1.00 indicates a perfect fit, and is derived from the comparison of a restricted model with a null model (one in which each observed variable represents a factor). CFI also provides a measure of complete covariation in the data, and a value larger than 0.90 indicates a psychometrically acceptable fit to the data[24].

In addition to assessing overall model fit, SEM also permits investigators to assess the degree of variance accounted for in each dependent variable and to determine whether individual path loadings are significantly different from 0. If calculated t values exceed the value of 1.96, then the parameter estimate is statistically significant[25].

4) Predictive validity

The predictive validity of the outcome indicators was tested using the multiple indicators model which is a sub-model of SEM, and the general linear model (GLM)[15]. The multiple indicators model is a regression analysis model explaining the relationship among latent variables as constructs. This model allows the investigators to deal with distinguishing three kinds of measurement errors: those of independent variables, those of dependent variables and the error for describing the causal relation among constructs, whereas the generally used regression analysis assumes an error term only in the dependent variable. In the present study, two latent variables, those of the outcome indicator and the HRQOL, were used for examining the causal relation between these constructs. The multiple indicators model was constructed, supposing that each indicator of the baseline survey would have its influence on each subject's SF-36 score after two years.

Analysis of covariance (ANCOVA), which is the main technique used in the GLM procedure, was also used to assess the predictive validity; the influences on SF-36 extended by the change in outcome indicators in two years were analyzed. In addition, least squares means (LSM) of SF-36 scores adjusted values for each outcome indicator of the baseline survey were also computed, to remove the influence of confounding variables.

5) Clinical relevancy

To assess the clinical relevancy of the outcome indicators, we examined the relationship between each outcome indicator and the subjects' profiles: gender, age, and ADL. Furthermore, we calculated the rate of patient improvement and stabilization for each outcome indicator for three disease categories, cerebrovascular diseases, demyelinating and degenerative diseases, peripheral nerve disorders and myopathy, based on scores of the changes in outcome indicators in two years. The rates of improvement and stabilization were calculated according to the methodology of the Outcome-Based Quality Improvement System in Medicare[6].

Results

1. The scaling structure in the outcome indicators

Before EFA, five unnecessary items were excluded. Five factors were extracted from the 25 remaining items by EFA using an oblique promax rotation accompanied by the maximum likelihood method (Table 1).

According to our results, the first factor consisted of seven items concerning needs, problems with the progression of the disease/disorder and anxiety, which the present health care system has difficulty covering, and which were interpreted as "anxiety about disease and disability indicator". The second factor consisted of four items concerning the caretaker, and was interpreted as "family care burden and strain indicator". The third factor consisted of five items showing IADL problems centering on movement ability, and was interpreted as "motor dysfunction indicator". The fourth factor consisted of five items concerning ADL problems and complaints of disturbed comfort, and it was interpreted as "appearance of symptoms indicator". The fifth factor consisted of four items concerning problems receiving consultation in a specialized hospital, necessitated by worsening of the disease, and specialist advice for treatment, and was interpreted as "interference in social network utilization indicator". The results of EFA indicated that the outcome indicators could be divided into five factors with comparatively simple structures and with high factor loadings held in items included in each sub-indicator.

The results of EFA, using the oblique promax rotation method, also show an inter-factor correlation coefficient. The maximum value of the inter-factor correlation coefficient was between the first and second factors ($r = 0.784$), and the minimum value was between the fourth and the fifth factors ($r = 0.417$). Regarding the reliability coefficient, Cronbach's alpha of five factors was between 0.78 and 0.92 and fulfilled the acceptance criteria, whereas the reliability coefficient of the fourth outcome indicator was slightly low.

2. Second-order factor model structure of outcome indicators by structural equation modeling

Figure 2 shows the second-order factor model of the outcome indicators. As the observed variables structuring the hypothetical model, three observed variables with high factor loading and comparatively simple structures of relevant factors based on the results of EFA were allocated to each factor. The reason is that one of the limitations of this model involves the maximum number of indicator variables that can be effectively studied[26]. For this measurement, we decided to acknowledge the residual covariance between error variables within a logically valid range, and examined the second-order factor model's goodness of fit.

Table 1 Factor pattern coefficients† of outcome indicators consisting of 25 items

Item	Factor 1	2	3	4	5	Communality
1) Feels anxious because medicines always need to be kept on hand.	**.364**	−.216	.150	.343	.160	.473
2) Feels that there is no way to cope with disease because sees no marked improvement.	**.868**	−.040	.103	.058	−.143	.754
3) Feels anxious about how to cope with a disease which is difficult to cured.	**.855**	.044	.121	.026	−.148	.812
4) Has difficulty working because of occasional symptoms.	**.642**	−.062	.032	.056	.069	.486
5) Feels anxious because of a lack of knowledge about the disease.	**.449**	.151	−.012	−.017	.239	.508
6) Feels anxious and conflicted after being discharged from the hospital.	**.656**	.218	.037	−.034	.069	.751
7) Has to live concealing disease.	**.493**	−.082	−.184	.016	.236	.258
8) Family caregiver cannot get enough sleep because of the need for constant care.	−.129	**.746**	−.146	.401	.002	.701
9) Family caregiver feels uneasy about caregiving.	.098	**.875**	.059	−.067	−.025	.876
10) Family caregiver hardly has any spare time to go out.	−.157	**.774**	.164	−.002	.051	.710
11) Family caregiver feels strong anxiety and conflict toward caregiving.	.180	**.703**	.087	−.080	.077	.820
12) Takes time to go to hospital and tires easily.	.075	−.070	**.598**	.147	.066	.540
13) Has difficulty with daily living activities such as preparing food and eating because of insufficient recovery.	.290	.115	**.533**	−.023	−.109	.605
14) Has a problem being accompanied by a family member when going to the hospital and for rehabilitation.	−.078	.199	**.621**	.076	.058	.665
15) Has difficulty in daily living activities because of walking difficulty.	.054	.093	**.844**	−.021	−.054	.836
16) Has difficulty going to the hospital because of problems of waking up and down stairs and riding in trains or other vehicles.	−.028	.003	**.907**	−.061	.044	.777
17) Has problems excreting and often has to a use laxative to fight constipation.	−.108	.002	.309	**.415**	.085	.396
18) Has difficulty urinating and a frequent desire to urinate.	−.062	.139	.231	**.407**	.129	.513
19) Has difficulty swallowing food because of dry mouth and problems swallowing.	.005	.335	.240	**.341**	−.166	.505
20) Cannot get enough sleep.	.160	−.033	−.123	**.710**	−.012	.494
21) Cannot have a long conversation because of breathing difficulties.	.127	.324	−.087	**.355**	−.017	.367
22) There is no special hospital for emergency consultation when feeling sick.	.003	−.020	.146	.072	**.584**	.495
23) Has no one to consult about how to cope with problems.	.383	−.018	−.140	.009	**.633**	.668
24) Has difficulty keeping in touch with the hospital from which he/she was discharged.	−.088	.059	.014	.015	**.755**	.568
25) Needs and desires for rehabilitation are not fulfilled.	.099	.241	.313	−.166	**.373**	.634
Eigen value for each factor	12.179	1.705	1.233	1.171	0.947	—

†: explanatory factor analysis using an oblique promax rotation with maximum likelihood method.

Analysis by SEM showed that the fit of the model was adequate at chi-square/d.f. = 0.621 ($p > 0.05$), GFI = 0.936, AFGI = 0.906, CFI = 1.000 and RSMEA estimate = 0.001. The standardized path coefficient was positive for the five latent variables (first-order factors) from the observed variables (0.520–0.926). The standardized coefficient was also positive for the second-order factor (overall outcome indicator) from the five first-order factors (0.716–0.929). The coefficients of determination (R-Square) of the five latent variables ranged from 0.512 to 0.864. The 15 observed variables ranged in R-Square from 0.428 to 0.858, except item No. 11 which had a coefficient of determination of 0.270. As shown in Table 2, the observation of homogeneity of the second-order factor models which were constituted from six deferent datasets divided by the patient's profiles, indicated that the fit of the model was also adequate, judging from the values of each parameter estimates and fit indices, whereas the value of fit indices for the dataset of cerebrovascular group showed insufficient model fit. It was suggested that the five indicators showing each aspect of the degree of difficulty in performing daily living activities could possibly be integrated in one factor as an overall outcome indicator.

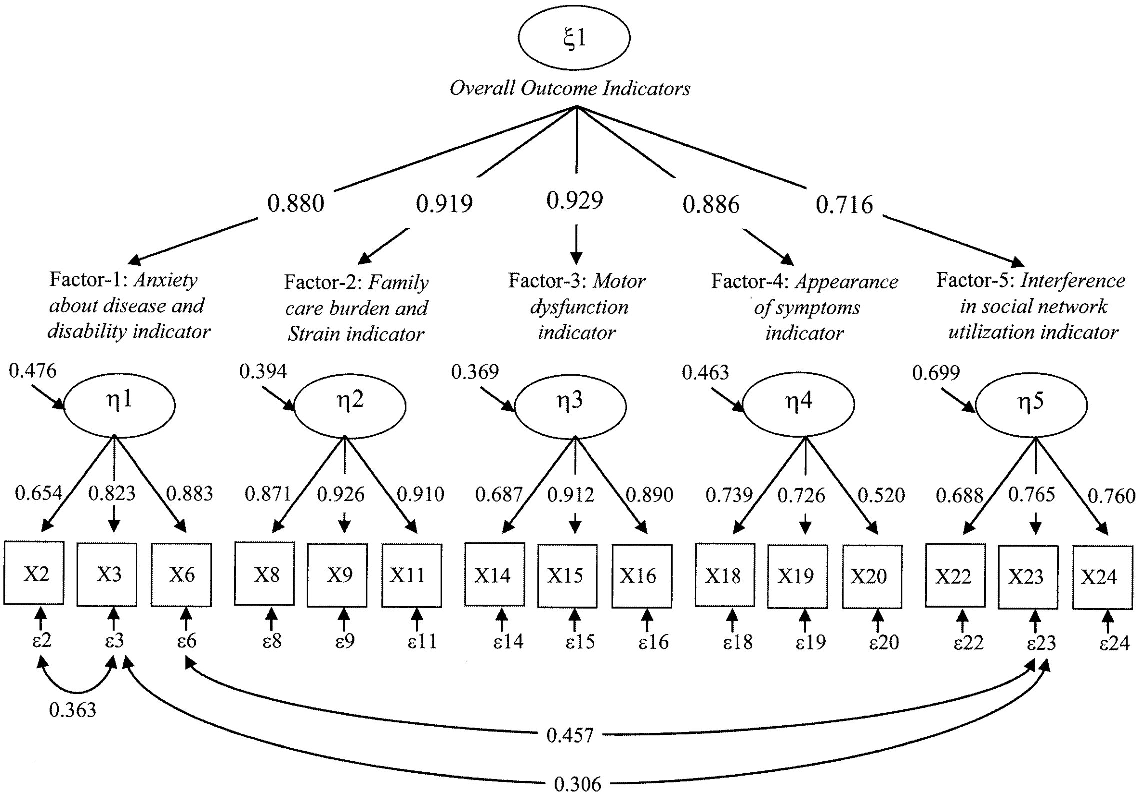

Fig. 2 Second-order factor structural model of the outcome indicators using structural equation modeling of PROC CALIS
All loadings are significant at the 0.05 level. Overall model fit: chi-squares/d. f. = 0.621, goodness of fit index (GFI) = 0.936, adjusted goodness of fit index (AGFI) = 0.906, root mean square error of approximation (RMSEA) estimate = 0.001, comparative fit index (CFI) = 1.000. An elliptical shape: latent variables, A box: observed variables (see Table 1 for a variable number.), ξ : exogenous variables, η : endogenous variables, ε : the residual for observed variables.

3. Relationship between the scores of outcome indicators and patient's profiles

With regard to the difference of gender, the degree of difficulty for females in the third outcome indicator for "motor dysfunction" was significantly high ($p<0.001$). The correlation coefficient between age and each indicator is low, ranging from −0.064 to 0.171. Accordingly, these indicators were considered to be applicable regardless of age. The correlation coefficient between ADL and each indicator was moderate, ranging from 0.431 to 0.720.

4. Patient improvement and stabilization of the outcome indicators

As shown in Table 3, when calculating the rates of improvement and stabilization of the outcome indicators over two years among patients, the rate of improvement ranged from 0.258 to 0.398 and that of stabilization ranged from 0.561 to 0.700 in all diseases. The rate of improvement in the group of cerebrovascular diseases was lower, compared with that in the demyelinating and degenerative diseases, peripheral nerve disorders and myopathy groups. In particular, low improvement of the second indicator for "family care burden and strain" was observed in the cerebrovascular diseases group.

5. Assessment of the predictive validity of the outcome indicators

1) Assessment using Multiple Indicators Model

As noted, we conducted the follow-up survey for the same subjects as the first one, and analyzed the responses from 201 participants 2 years after the baseline survey. For the first, we constructed a multiple indicators model, supposing that each indicator

Table 2 Parameter estimates based on second-order factor model using SEM for each of patient's profiles

Parameters	Gender		Age-class		Disease category	
	Man ($n^{\dagger}$ = 252)	Woman (n = 211)	under 65 (n = 235)	65 & over (n = 228)	Neurological (n = 329)	Cerebrovascular (n = 128)
Factor loadings						
ξ_1						
$\gamma_{1,1}$	0.885	0.876	0.873	0.909	0.863	0.937
$\gamma_{2,1}$	0.912	0.931	0.888	0.963	0.914	0.936
$\gamma_{3,1}$	0.937	0.925	0.946	0.887	0.917	0.988
$\gamma_{4,1}$	0.909	0.878	0.872	0.897	0.903	0.931
$\gamma_{5,1}$	0.682	0.762	0.717	0.728	0.674	0.837
η_1						
$\kappa_{2,1}$	0.824	0.822	0.845	0.790	0.820	0.812
$\kappa_{3,1}$	0.674	0.642	0.630	0.688	0.631	0.687
$\kappa_{6,1}$	0.866	0.904	0.921	0.844	0.887	0.866
η_2						
$\kappa_{8,2}$	0.922	0.923	0.937	0.914	0.918	0.952
$\kappa_{9,2}$	0.882	0.864	0.872	0.880	0.857	0.916
$\kappa_{11,2}$	0.933	0.885	0.914	0.903	0.911	0.929
η_3						
$\kappa_{14,3}$	0.902	0.880	0.917	0.877	0.893	0.875
$\kappa_{15,3}$	0.819	0.590	0.882	0.529	0.828	0.452
$\kappa_{16,3}$	0.923	0.895	0.898	0.925	0.908	0.903
η_4						
$\kappa_{18,4}$	0.730	0.752	0.732	0.747	0.714	0.777
$\kappa_{19,4}$	0.777	0.646	0.683	0.769	0.756	0.661
$\kappa_{20,4}$	0.545	0.479	0.461	0.571	0.447	0.667
η_5						
$\kappa_{22,5}$	0.615	0.758	0.740	0.613	0.662	0.756
$\kappa_{23,5}$	0.776	0.750	0.738	0.843	0.764	0.789
$\kappa_{24,5}$	0.838	0.697	0.760	0.758	0.780	0.724
Errors in equations						
ζ_1	0.466	0.482	0.487	0.416	0.506	0.348
ζ_2	0.411	0.364	0.459	0.271	0.406	0.353
ζ_3	0.349	0.379	0.325	0.461	0.398	0.155
ζ_4	0.416	0.478	0.490	0.442	0.430	0.365
ζ_5	0.731	0.648	0.697	0.685	0.739	0.547
Correlated measurement errors						
ε_1-ε_2	0.530	0.223	0.254	0.535	0.294	0.587
ε_2-ε_{14}	0.276	0.350	0.405	0.154	0.332	0.237
ε_3-ε_{14}	0.449	0.493	0.485	0.354	0.453	0.411
Values of fit indices						
GFI	0.903	0.894	0.919	0.889	0.919	0.839
AGFI	0.859	0.844	0.881	0.838	0.881	0.765
Chi-Square/d.f.	1.049	1.112	0.825	1.131	0.804	1.952
RMSEA	0.025	0.034	0.001	0.036	0.001	0.098
CFI	0.995	0.990	1.000	0.989	1.000	0.937

ξ: the second-order factor (exogenous variable), η: the first-order factor (endogenous variables), γ: the structural coefficient linking the latent variables (ξ, η), κ: the coefficients from the endogenous variables to the observed variables, ζ : the residual for endogenous variables, ε: the measured error for observed variables. $\dagger$: cases with missing values were deleted from this analysis.

of the baseline survey would have an influence on each subject's HRQOL (SF-36) after two years. The multiple indicators model was constructed assuming that the outcome indicators with structural concepts consisting of three variables would have an influence on HRQOL with structural concepts consisting of eight domains of SF-36. Figure 3 shows the strength of the first indicator on HRQOL. The path coefficient from the first indicator was -0.640; the increase in the degree of difficulty in performing

Table 3 Change in outcome indicators by neurological disease categories two years after the baseline survey

Disease categories	Outcome indicator§				
	1	2	3	4	5
Rate of improvement†					
Overall diseases	0.398	0.316	0.358	0.258	0.325
Cerebrovascular diseases	0.333	0.125	0.286	0.294	0.200
Demyelinating and degenerative diseases	0.444	0.375	0.441	0.265	0.343
Peripheral nerve disorders and Myopathy	0.462	0.400	0.320	0.136	0.350
Rate of stabilization‡					
Overall diseases	0.561	0.700	0.560	0.536	0.593
Cerebrovascular diseases	0.500	0.706	0.542	0.604	0.608
Demyelinating and degenerative diseases	0.643	0.583	0.600	0.432	0.514
Peripheral nerve disorders and myopathy	0.586	0.824	0.546	0.533	0.630

†: If the patient's outcome improves between the baseline and follow-up point, this outcome measurement takes on a value of 1; otherwise it is 0. Patients who cannot improve (do not have problems relevant to the outcome indicator between the two points) are excluded from the computation of this measurement. ‡: If the patient's outcome does not worsen between the baseline and follow-up point, this outcome measurement takes on a value of 1; otherwise it is 0. Patients who cannot worsen (are already the most severe level of the relevant outcome indicator between the two points) are excluded from the computation of this measurement. For all the analyses, missing values were excluded from the computation of these measurements. §: Outcome indicator, 1; anxiety about disease and disability indicator, 2; family care burden and strain indicator, 3; motor dysfunction indicator, 4; appearance of symptom indicator, 5; interference in social network utilization indicator.

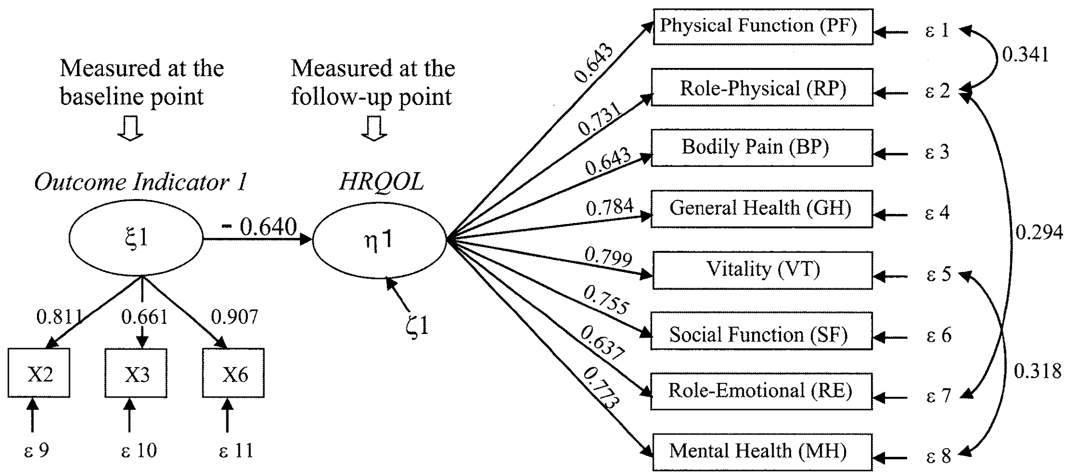

Fig. 3 The causal relationship from indicator 1 at the baseline survey point to HRQOL using the multiple indicators model constructed from structural equation modeling
All loadings are significant at the 0.05 level. Overall model fit: chi-square/d.f. = 1.141, goodness of fit index (GFI) = 0.925, adjusted goodness of fit index (AGFI) = 0.878, root mean square error of approximation (RMSEA) estimate = 0.038, comparative fit index (CFI) = 0.991. An elliptical shape: latent variables, A box: observed variables (see Table 1 for a variable number), ξ: exogenous variables, η: endogenous variables, ζ: the residual for endogenous variables, ε: the residual for observed variables.

daily living activities tended toward lowering HRQOL. The remaining four path coefficients from the second indicator, the third indicator, the fourth indicator and the fifth indicator were −0.542, −0.576, −0.492 and −0.498, respectively. Thus all indicators had influenced the HRQOL conditions

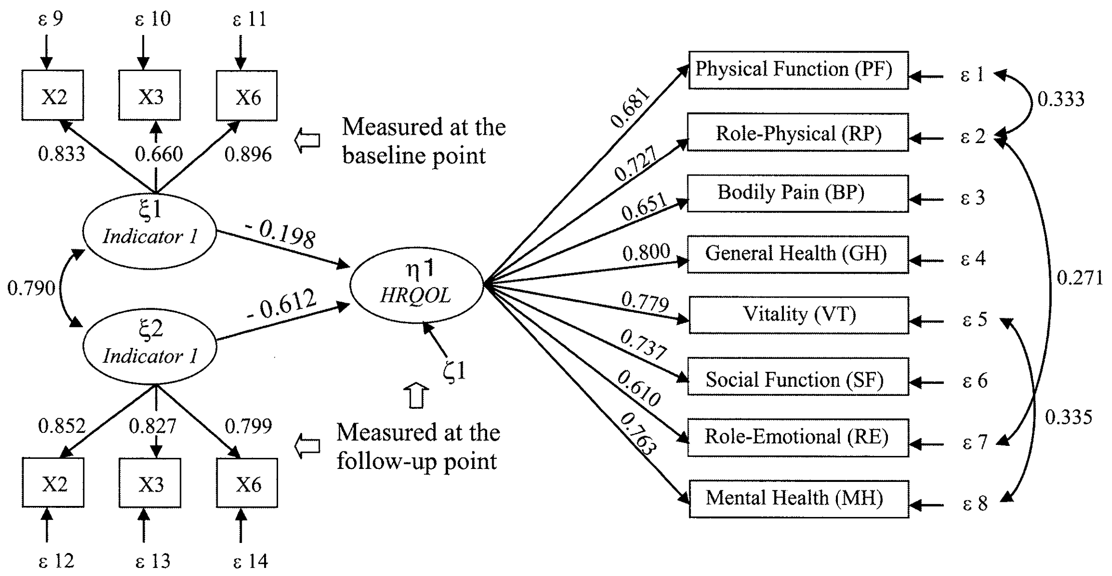

Fig. 4 The causal relationship from indicator 1 of the baseline survey and follow-up survey to HRQOL using the multiple indicators model constructed from structural equation modeling (SEM)
All loadings are significant at the 0.05 level. Overall model fit: chi-square/d.f. = 1.014, goodness of fit index (GFI) = 0.906, adjusted goodness of fit index (AGFI) = 0.868, root mean square error of approximation (RMSEA) estimate = 0.012, comparative fit index (CFI) = 0.999. An elliptical shape: latent variables, A box: observed variables (see Table 1 for a variable number), ξ : exogenous variables, η : endogenous variables, ζ : the residual for endogenous variables, ε : the residual for observed variables.

two years after the baseline survey. In particular, the first indicator for "anxiety about disease and disability" influenced HRQOL the most.

Secondly, the predictive validity was tested by a constructed model in which outcome indicators of both the baseline survey and the follow-up survey influenced HRQOL. In Fig. 4, the upper left side shows the outcome indicators determined at the baseline survey, and the lower part shows the same indicators determined by the follow-up survey. In this model, there is a correlation between structural concepts expressing indicators, and each indicator influences HRQOL. This model shows that the influence on HRQOL extended from outcome indicators determined at the baseline survey was comparatively low (path coefficient: from −0.009 to −0.098). On the other hand, the influence on HRQOL extended from the first, second, third, fourth and fifth outcome indicators determined with the second survey were −0.612, −0.611, −0.783, −0.309 and −0.740, respectively. Generally, the outcome indicators determined from the follow-up survey, excluding the fourth indicator, had large influences on HRQOL, and among them, the third indicator for "motor dysfunction" and the fifth indicator for "interference in social network utilization" were shown to have large influences on HRQOL.

2) Assessment by using general linear model

Figure 5 shows the analysis of influence on HRQOL extended by the change in outcome indicators in two years using GLM. Because the scores of outcome indicators determined at the baseline time point might possibly become confounding variables which extend their influence on both the change in the scores of outcome indicators in two years and HRQOL at the same time, the influence of the scores of outcome indicators at the baseline time point were adjusted. In this measurement, the GLM was constructed using each SF-36 domain as a dependent variable and the change in outcome indicators in two years as explanatory variables.

The results of GLM indicated that the aspect of HRQOL which outcome indicators influence was different for each indicator. Based on whether the

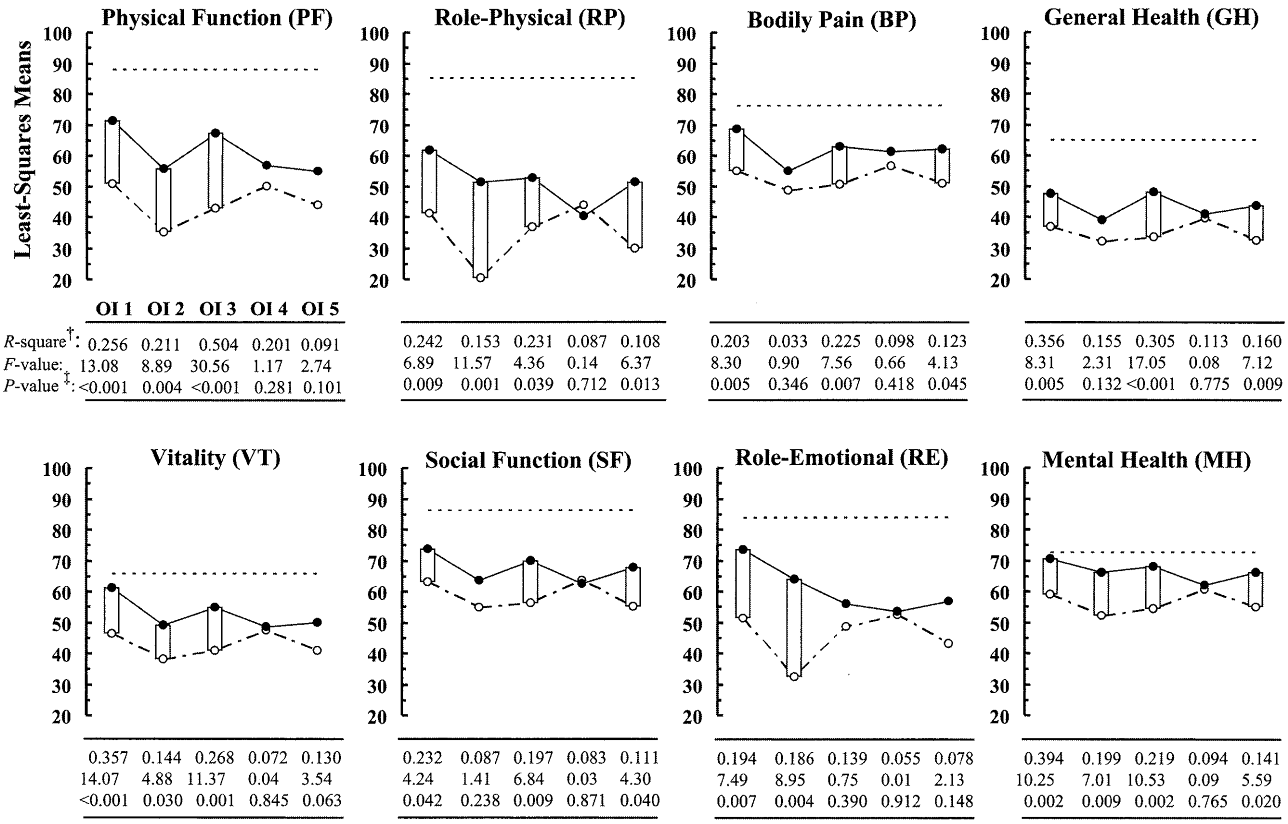

Fig. 5 The analysis of influences on HRQOL (SF-36) extended by the change in outcome indicators over two years using the General Linear Model
† : analysis of covariance (ANCOVA), adjusted the baseline score of each outcome indicator. ‡ : test; Fisher's least significant difference test. Least-Squares Means: adjusted mean score of each domain in the SF-36.
● : improvement group, ○ : no change or worsening group, □ : range between the improvement and the no change or worsening group in which statistical significance ($p < 0.05$) was observed, ------ : published sex and age-specific norms of the SF-36 available from the general population of Japan.

scores of outcome indicators improved or not in two years, the subjects whose degree of difficulty in performing daily living activities increased in the two years showed a remarkable decrease in HRQOL relating to role functioning and physical functioning. Specifically, the first outcome indicator significantly influenced all SF-36 domains, and in particular, had a large influence on PF, BP, VT, and MH. The second outcome indicator had a large influence on the domains relating to role functioning of RP and RE. The third outcome indicator had a large influence on the four domains of PF, GH, SF, and MH. The fourth outcome indicator had no significant influence on all domains, and a low correlation to PF. The fifth outcome indicator moderately influenced the five domains other than PF, VT, and RE, and in particular influenced GH, RP, and SF.

Discussion

The following four contributions were presented in this study: First, the outcome indicators that were perceived by patients and their families in order to evaluate the effectiveness of healthcare service for the home-care neurology patients discharged from the university hospital have been developed. Second, a part of the validity of the constructive concept of the outcome indicators was verified by the structural equation modeling. Third, the improvement rates and the stabilization rates of the outcome indicators of each neurological disease were obtained based on the difficulties in performing daily living activities for two years. Fourth, as a result of the general linear model on the values of the effects for HRQOL caused by the variance of the outcome indicators in two years, the aspects of HRQOL caused by the variance of outcome indicators differed by each indicator.

The outcome indicators consist of the sub-indicators for five dimensions: 1) "Anxiety about disease and disability indicator" consists of items relating to needs difficult to meet by current healthcare services and those relating to the difficulties for the treatment and anxiety both accompanied by the progress of the disease. These items are shown to be unique problems for neurological disease patients including intractable cases and their families[27]. 2) "Family care burden and strain indicator" includes family care burdens and relates to a scale which has been developed in order to measure the feeling of care burden of families. 3) "Motor dysfunction indicator" includes IADL impairment centered on moving ability. Neurological disease patients often have a motor nerve dysfunction in their upper and lower limbs, which as a result causes the problems in performing daily living activities. They also have difficulties in going to a hospital when they feel ill since they cannot freely use public transportation. 4) "Appearance of symptom indicator" consists of the contents relating to the problems of physical ADL and symptoms interfering with comfort, that is, the items of this indicator are related to the appearance of physical symptoms, which cause difficulties in maintaining basic living activities. 5) "Interference in social network utilization indicator" relates to the difficulties in receiving healthcare service for home care. As the number of patients with a neurological intractable disease is small, the number of easily accessible specialized hospitals is also small. Accordingly, the patients have difficulty in receiving professional service when they want to go to a local specialized hospital due to worsening symptoms or when they want to receive professional advice as to the care of the symptom. The outcome indicators developed as above include three types of outcome measures, that is, end-result outcome, intermediate-result outcome, and utilization outcome. Furthermore, it is considered that these outcome indicators meet the content validity since they cover the difficult problems that need to be solved, which neurological patients and their families face in daily living.

In this study, the structural equation modeling was used to verify the construct validity of the outcome indicators due to the reasons described below. An explanatory factor analysis has often been used for the assessment of the validity of the constructive concept. However, it has been shown that there are many limits on the validity verification of the constructive concept using explanatory factor analysis. Explanatory factor analysis is a multivariate analysis method, in which data are condensed by explaining the correlations between many variables with the correlations between a few common factors and the observed variables[28]. Characteristically, for explanatory factor analysis, a model assumes that there are correlations between all factors but no correlation between common factors and unique factors or between unique factors. Accordingly, although the explanatory factors are effective in explaining the structure establishing the correlations between observed variables, it is desirable to apply covariance structural equation modeling for hypothesis verification.

On the other hand, the structural equation modeling has a large amount of freedom in constructing the model. Although it is possible to construct any model, however inappropriate, it can be rejected by fit indices. As it is impossible to calculate all possible models, it can never be made certain that the final model of the outcome indicators applied in this study is the most optimal. However, although the final model applied to this study shows the possibility of being an overall outcome indicator consisting of five factors, it shows a good fit with the model. In the final model applied, by a calculated residual matrix being taken into consideration, three covariances were specified between error variables. A possible explanation for the fact is that an error covariance between the items (items No. 2 and 3) included in "anxiety for diseases and disability indicator" shows the possibility of a similar meaning of wording, and that the existence of an error covariance between the items (items No. 23, 3 and 6) consisting of "anxiety about disease and disability indicator" and "interference in social network utilization indicator" shows the possibility that there exists a correlation between anxiety and social network relating to the cure of neurological disease patients which is not mentioned in the final model applied. The observation of homogeneity in the final model indicated the insufficient model fit for only the dataset of the cerebrovascular group. One possible explanation for the fact is that the number of subjects included in this dataset was less than what was required empirically for the computation of SEM. Although it is furthermore necessary to sift and refine the items of outcome indicators, and to test the cross validity in a different population in the future, judging from the high values of the fit indices and the determination coefficients, we consider that the validity of the developed outcome indicators is not impaired.

The improvement rates of each outcome indicator of home-care neurological patients for two years were in the range of 26 to 40%, and the stabilization rates were in the range of 54 to 70%. When considered according to disease group, it is clear that the improvement rates of the "family care burden and strain indicator" and the "interference in social network utilization indicator" of the cerebrovascular diseases group were lower than those of the demyelinating and degenerative diseases group and the peripheral nerve disorders and myopathy group. One possible explanation of the above fact in the case of cerebrovascular diseases with serious ADL impairment is that the care burden weighs heavily on the physical function and overall daily living activities of the patients, and it reflects the actual situation that they cannot effectively utilize the service of long-term care insurance. Accordingly, the importance of mental health care, not only for the patients but also for family caregivers, should be taken into consideration. It is shown that the aspects of HRQOL influenced by whether there were improvements in outcome indicators over the course of two years or not, differ by each indicator. In particular, the "motor dysfunction indicator" and the "anxiety about disease and disability indicator" contributed to the lowering of many aspects of HRQOL for home-care patients. Furthermore, the patients' own sense of inadequacy toward their social role was also a large factor in lowering HRQOL. In order to improve patients' HRQOL, health professionals continuously involved in medical intervention and care of the patients after they are discharged from the hospital should not only pay attention to their clinical condition, but also to discover and solve the diverse psychological and social problems which patients are burdened with.

Compared to the progress of outcome assessment for home care in the US[1)5)], only a few outcome studies on the home care of home-care neurology patients in Japan have been conducted. However, as the number of patients receiving home care is further increasing, the progress of outcome assessment studies is an imperative task from the viewpoint of quality of care[29)]. As the limits of this study, it is necessary to further sift and refine the items of the developed outcome indicators. Because these indicators are assessment scales which use patient perceptions, and proxy scales are used for the measurements, it cannot be denied that measurement by these outcome indicators lacks credibility compared to the actual objectively measurable condition of patients. However, patient perceptions and perception of the proxies of patients with severe chronic illness such neurological diseases have been utilized more than ever to measure the quality of healthcare[30)]. From the above, the developed outcome indicators are useful as an instrument to assess the effectiveness of healthcare service for the home-care neurology patients. For future study, outcome indicators unique to each neurological disease should be developed and their availability verified.

References

1) **Keepnews D, Capitman JA, Rosati RJ**: Measuring patient-level clinical outcomes of home health care. J Nurs Scholarsh **36**: 79–85, 2003
2) **Department of Health and Human Services, Health Care Financing Administration**: Medicare and Medicaid program: Reporting Outcome and Assessment Information Set (OASIS) data as part of the condition of participation for home health agencies and comprehensive assessment. Fed Reg-

ist **64**: 3748–3763, 1999
3) **Shimanouchi S, Oga H, Yamguchi A et al**: Evaluation method of home care agency from care effectiveness for clients: comparative study of standard value and agency based on adjustment of client's background. J Jpn Acad Com Health Nurs **3**: 76–85, 2001 (in Japanese with English abstract)
4) **Shimanouchi S, Shimizu Y, Tomoyasu N et al**: Care items and degree of implementation as it affects client outcomes in home care. J Jpn Acad Com Health Nurs **4**: 26–33, 2002 (in Japanese with English abstract)
5) **Centers for Medicare and Medicaid Services**: Medicare and Home Health Care, pp11–16, U.S. Department of Health and Human Services, Maryland (2004)
6) **Shaughnessy PW, Crisler KS, Schlenker RE et al**: Measuring and assuring the quality of home health care. Health Care Financ Rev **16**: 35–67, 1994
7) **Ito K**: Nokekkan-shogai oyobi nanbyo no zaitaku-kea taishosha wo sukuriningu suru Case-Mix shihyo no kaihatsu, pp1–34, JSPS Grant-in-Aid for Scientific Research (C) No. 09672432 (2000) (in Japanese)
8) **Kramer AM, Shaughnessy PW, Bauman MK et al**: Assessing and assuring the quality home health care: a conceptual framework. Milbank Q **68**: 413–443, 1990
9) **Cleary PD, McNeil BJ**: Patient satisfaction as an indicator of quality care. Inquiry **25**: 25–36, 1988
10) **Katz S, Ford AB, Moskowitz RM et al**: Studies of illness in the aged: The index of ADL: A standard measure of biological and psychosocial function. JAMA **185**: 914–919, 1963
11) **McDowell I, Newell C**: Measuring Health: A guide to rating scales and questionnaires, 2nd ed., pp47–121, Oxford University Press, New York (1996)
12) **Ware JE, Sherbourne CD**: The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual framework and item selection. Med Care **30**: 473–483, 1992
13) **Fukuhara S, Bito S, Green J et al**: Translation, adaptation, and validation of the SF-36 Health Survey for use in Japan. J Clin Epidemiol **51**: 1037–1044, 1998
14) **Fukuhara S, Ware JE, Kosinski M et al**: Psychometric and clinical tests of validity of the Japanese SF-36 Health Survey. J Clin Epidemiol **51**: 1045–1053, 1998
15) **SAS Institute Inc.**: SAS/STAT® User's Guide, Version 6, 4th ed, SAS Institute Inc, Cary, NC (1990)
16) **SPSS Japan Inc.**: SPSS® Base 11.5J User's Guide, SPSS Japan Inc, Tokyo (2002)
17) **Nunnally JC, Bernstein IR**: Psychometric Theory, 2nd ed, pp229–230, McGraw-Hill, New York (1978)
18) **Musil CM, Jones SL, Warner CD**: Structural equation modeling and its relationship to multiple regression and factor analysis. Res Nurs Health **21**: 271–281, 1998
19) **Stuifbergen AK, Seraphine A, Roberts G**: An explanatory model of health promotion and quality of life in chronic disabling conditions. Nurs Res **49** (3): 122–129, 2000
20) **Hu LT, Bentler PM**: Evaluating model fit. *In* Structural Equation Modeling: Concepts, Issues, and Applications (Hoyle RH ed), pp76–99, SAGE Publication Inc, Thousand Oaks, CA (1995)
21) **Sousa KH, Chen FF**: A theoretical approach to measuring quality of life. J Nurs Meas **10**: 47–58, 2002
22) **Steiger JH**: Structural model evaluation and modification: An interval estimation approach. Multivariate Behav Res **25**: 173–180, 1990
23) **Browne MW, Cudeck R**: Alternative ways of assessing model fit. *In* Testing Structural Equation Models (Bollen KA, Long JS eds), pp136–162, SAGE Publication Inc, Newbury Park, CA (1993)
24) **Bentler PM**: Comparative fit indices in structural equation models. Psychol Bull **107**: 238–246, 1990
25) **Hatcher L**: A Step-by-Step Approach to Using the SAS System for Factor Analysis and Structural Equation Modeling, pp249–342, SAS Institute Inc, Cary, NC (1994)
26) **Bentler PM, Chou C**: Practical issues in structural modeling. Sociol Methods Res **16**: 78–117, 1987
27) **Maruyama K**: Clinical studies on obstructive factors of home care in patients with chronic neurological diseases. St Marianna Med J **20**: 497–508, 1992
28) **Toyoda H**: Sokutei hyoka to kyobunsan kozo moderu. Annu Rep Educ Psychol Jpn **36**: 119–127, 1997 (in Japanese)
29) **Ringel SP, Vickrey BG**: Measuring quality of care in Neurology. Arch Neurol **54**: 1329–1332, 1997
30) **Rosenthal GE, Shannon SE**: The use of patient perceptions in the evaluation of health-care delivery systems. Med Care **35**: NS58–NS68, 1997

在宅神経疾患患者に対するヘルスケアサービスの評価に活用する アウトカム指標の開発と妥当性の検証

[1]東京女子医科大学　医学部　神経内科学（主任：岩田　誠教授）
[2]東京女子医科大学　看護学部　地域看護学
[3]東京女子医科大学　看護学部　内科学

伊藤　景一[1)2)]・渡辺　弘美[1)3)]・岩田　誠[1)]・佐々木彰一[1)]・内山真一郎[1)]

在宅神経疾患患者とその家族に対する長期間に渡るヘルスケアサービスの効果評価に活用する，日常生活行動を遂行する上での困難度を基にしたアウトカム指標の開発を行った．さらに，共分散構造方程式モデリングと一般線形モデルの手法を用いて，アウトカム指標の構成概念妥当性と予測妥当性を検証した．大学病院神経内科を過去5年間に退院した463名の対象者に対する，指標に含まれる項目の説明的因子分析の結果，アウトカム指標は5次元から構成された．共分散構造方程式モデリングを用いて各指標を一次因子とし，その上位概念としての総合的なアウトカム指標を仮定した二次因子モデルの成立の有無を検証した．最終的に採択されたモデルにおけるパス係数は全て有意（$p<0.05$）となり，適合度指標も受入れ基準を満たした．一次因子は，第1指標：疾病障害対処困難・不安指標，第2指標：家族介護負担・Strain指標，第3指標：運動機能不全指標，第4指標：身体症状発現指標，および第5指標：地域医療・ソーシャルネットワーク利用阻害指標，と解釈された．次に，各アウトカム指標が在宅神経疾患患者における2年後の健康関連QOL（SF-36）に及ぼす影響を，201名のフォローアップ対象者に対し，共分散構造方程式モデリングの一手法である多重指標モデルを用いて検証した．5指標全てが有意に2年後の健康関連QOLに影響を与えていたが，特に第1指標と第3指標から健康関連QOLに与えるパス係数の値が大きかった．また，一般線形モデルを用いて，2年間におけるアウトカム指標の改善の有無が健康関連QOLに与える影響をみた結果からは，各アウトカム指標が影響を及ぼしている健康関連QOLの側面は異なることを見出した．特に，身体機能と役割機能のドメインにおいて，改善群と非改善群における差が大きくなっていた．開発したアウトカム指標は，その統計学的な妥当性検証結果から在宅神経疾患患者と家族に対する長期間の在宅ケアサービスの効果評価に活用する上で有効であると考えられた．