Geriatrics and Gerontology International 2004; 4: 146–150

ORIGINAL ARTICLE

Improvement of inappropriate prescribing and adverse drug withdrawal events after admission to long-term care facilities

Yumiko Mita,[1] Masahiro Akishita,[1] Katsuaki Tanaka,[1] Shizuru Yamada,[1] Ryuhei Nakai,[2] Eigo Tanaka,[3] Tetsuro Nakamura[4] and Kenji Toba[1]

[1]Department of Geriatric Medicine, Kyorin University School of Medicine, [3]Mizukusaki-En and [4]Research Institute of Aging Science, Tokyo, and [2]Department of Neurology, Fukuoka University School of Medicine, Fukuoka, Japan

Background: The objectives of this study were to determine whether medications, particularly inappropriate prescribing, would be reduced after admission to long-term care facilities, and whether adverse drug withdrawal events (ADWEs) would occur in relation to discontinuation of medications.

Methods: The study consists of a retrospective survey using medical chart review in five health service facilities for the elderly in Japan. All the patients who were admitted to the facilities between January 2001 and December 2002 ($N = 627$) were participants in the study. Medications taken on admission, at 1 month and 3 months after admission, and events (significant worsening of the disease status, accidents, new symptoms and signs, and other acute events) during a 3-month period were recorded. Inappropriate prescribing was determined using Beers' criteria with some modification. ADWEs were determined using the Naranjo causality algorithm.

Results: On admission, the patients were taking 3.5 ± 2.5 (mean $\pm$ SD) drugs. One month later, the number of prescribed drugs was decreased by 17% ($P < 0.01$ vs on admission), but did not show an additional reduction 3 months later. Inappropriate prescribing was found in 10% of the patients taking drugs on admission, but the number of inappropriately prescribed medications was reduced by 33% after 1 month. Of 105 events recorded, only five (2% of the patients with drug reduction) were considered ADWEs; three cases of confusion, a case of depression, and a case of hyperglycemia, following discontinuation of psychotropic drugs, antidepressants and a sulfonylurea, respectively.

Conclusion: Adverse drug withdrawal events were not frequent despite the significant reduction of medications after admission to long-term care facilities. This might be because the rate of reduction was relatively high for inappropriately prescribed medications.

Keywords: adverse drug reaction, long-term care, medical expense, medical injury, pharmacotherapy.

Introduction

Adverse drug reactions in elderly people increase with age,[1-3] with most being attributable to medication errors that are preventable.[3,4] Age-dependent changes in pharmacokinetics and pharmacodynamics, polypharmacy and non-compliance related to patients' functional

Accepted for publication 12 February 2004.

Correspondence, Masahiro Akishita, MD, PhD, Department of Geriatric Medicine, Kyorin University School of Medicine, 6-20-2 Shinkawa, Mitaka, Tokyo 181–8611, Japan. Email: akishita-tky@umin.ac.jp

decline may play a role.[1,3] In particular, polypharmacy resulting from multiple pathology in elderly people is a critical problem leading to adverse drug reactions.[1–3] To prevent polypharmacy, review of prescriptions is essential according to evidence-based medicine and criteria for inappropriate prescribing.[5,6] In fact, inappropriate use of medication in elderly people has been reported to be as frequent as 16% to 25%.[7–9]

Conversely, discontinuation of medications to improve polypharmacy or inappropriate prescribing may induce adverse drug withdrawal events (ADWEs),[10] although the net effect on adverse drug reactions can be favorable in elderly outpatients.[11] Fixed payment insurance systems restrict medication use, possibly leading to a reduction of inappropriate prescribing and/or an increase of ADWEs. In health service facilities for the elderly in Japan, where functional training and nursing/personal care are provided under long-term care insurance,[12] a fixed payment system including prescribing of medication is applied. Accordingly, it is hypothesized that prescribed drugs, particularly inappropriate prescribing, would be reduced after admission to the facilities, and that ADWEs would occur in relation to discontinuation of medications. To test this hypothesis, we performed a retrospective chart review of a total of 627 patients in five health service facilities for the elderly, and found that prescribed drugs can be reduced with few ADWEs in such a frail elderly population with chronic diseases.

Methods

Sample and data collection

The data were derived from five health service facilities for the elderly (Mahoroba-no-Sato, Nagano; Moeu-no-Sato, Nagano; Himawari-En, Fukuoka; Milleniu-Sakuradai, Tokyo; Mizukusaki-En, Tokyo) in Japan. Institutional medical charts were reviewed for all the patients admitted between January 2001 and December 2002. Diagnoses of each patient were not recorded because they were unclear from the institutional charts, but Alzheimer's disease, cerebrovascular disease and osteoporosis were the main causes of disability in each institution. The average basic activities of daily living, as measured by the Barthel index, were 70–80 points out of 100 points according to the institutions. Medications that the patients were taking on admission and prescribed drugs 1 month and 3 months after admission were recorded. Similarly, all the events (significant worsening of the disease status, accidents, new symptoms and signs, and other acute events) during a 3-month period were recorded. The institutions that managed the patients before admission were categorized as acute care hospitals, outpatient clinics (home), sanitarium-type wards, special nursing homes for the elderly and health service facilities for the elderly. Patients with voluntary discharge within 3 months excluding cases of death or transfer to another hospital were excluded, and a total of 627 patients were analyzed. The director of each institution gave written approval to the participation in this study. The study protocol was approved by the committee on ethics and the institutional review board of Kyorin University School of Medicine.

Analysis

Inappropriately prescribed medications were determined using an updated version of the list developed by Beers with some modification.[5] Basically, we followed the list by Sloane *et al.* in which several drugs were excluded from Beers' list in consultation with Dr Beers,[5,9] reflecting changes in pharmacotherapy, but we included digoxin at more than 0.125 mg/day and oral iron at more than 325 mg/day in the list because these dosages were recorded in the medical chart. In this study, diagnosis-related inappropriate prescribing was excluded,[3] as in the study by Sloane *et al.* because the institutional chart did not include all the diagnoses of the patients.[9]

All the events were reviewed by a consultant geriatrician, and ADWEs were determined using the Naranjo causality algorithm.[13] Because detailed information, such as the effect of re-administration was lacking in most cases, a probability scale ≥ 1 (possible, probable or definite) was considered to indicate an ADWE.

The data in the text and the tables are expressed as means ± SD unless otherwise specified. Changes in the number of prescribed drugs after admission were analyzed using paired *t*-test. Differences between the groups were analyzed using ANOVA followed by Newman-Keuls' test.

Results

Number of prescribed drugs

The patients were taking 3.5 ± 2.5 drugs when admitted to the facilities (Table 1). Forty-six patients (7.3%) were not taking any drug, while 50 patients (8.0%) were on eight or more drugs. Women were taking fewer drugs than men. This sex difference seemed independent of age, although a statistically significant difference was found only at 80–89 years of age when the patients were categorized by age groups (Table 1). Interestingly, patients of 80 years or older were taking fewer drugs than those younger than 70 years, in contrast to a previous finding that the number of prescribed drugs increased according to age.[2,14,15]

As shown in Table 2, the mean number of prescribed drugs had decreased by 0.6 (17%) 1 month

Table 1 Number of drugs taken on admission according to sex and age

	All	Men	Women	*P* for sex difference
Total	3.5 ± 2.5 (627)	4.2 ± 2.8 (177)	3.3 ± 2.4 (450)	< 0.01
≤ 69 years	4.4 ± 3.1 (36)	4.6 ± 3.5 (19)	4.2 ± 2.6 (17)	0.70
70–79 years	4.0 ± 2.6 (131)	4.6 ± 3.0 (43)	3.7 ± 2.3 (88)	0.08
80–89 years	3.3 ± 2.3* (316)	4.0 ± 2.6 (81)	3.0 ± 2.2 (235)	0.02
≥ 90 years	3.5 ± 2.7* (144)	4.2 ± 2.4 (34)	3.2 ± 2.8 (110)	0.08

*$P < 0.05$ versus ≤ 69 years by Newman-Keuls' test.
Data are expressed as mean ± SD. Number of subjects is indicated in parentheses.

Table 2 Changes in number of prescribed drugs after admission to health service facilities for the elderly

	No. of subjects	On admission	After 1 month	After 3 months
Total	627	3.5 ± 2.5	2.9 ± 2.2*	3.0 ± 2.1*
Type of institution before admission				
Acute care hospital	115	4.8 ± 3.3†	4.2 ± 2.9*†	4.1 ± 2.7†
Outpatient	200	3.6 ± 2.3	2.8 ± 1.8*	2.9 ± 2.0
Special nursing home	24	3.3 ± 2.1	2.5 ± 1.7*	2.6 ± 1.8*
Sanitarium-type ward	188	3.1 ± 2.3	2.6 ± 1.9*	2.6 ± 1.9*
Health service facility	100	2.6 ± 1.8	2.4 ± 1.6*	2.5 ± 1.7*
Facility				
A	83	4.9 ± 3.4	4.6 ± 3.0	4.6 ± 2.4
B	80	4.2 ± 2.8	3.9 ± 2.4*	4.0 ± 2.5
C	39	4.1 ± 2.7	2.4 ± 1.7*	2.2 ± 1.4*
D	172	3.2 ± 1.9	2.4 ± 1.5*	2.4 ± 1.5*
E	253	3.0 ± 2.2	2.6 ± 1.9*	2.5 ± 1.9
Event				
No	517	3.5 ± 2.5	2.8 ± 2.0*	2.8 ± 2.1*
Yes	104	3.6 ± 2.7	3.4 ± 2.5****	3.7 ± 2.2***‡

*$P < 0.01$ versus on admission by paired *t*-test; **$P < 0.01$ versus after 1 month by paired *t*-test; †$P < 0.01$ versus other types of institution by Newman-Keuls' test; ***$P < 0.05$; ‡0.01 versus Event (–) by Newman-Keuls' test.
Data are expressed as mean ± SD.

after admission ($P < 0.01$ versus on admission), but did not show an additional reduction 3 months after admission. A significant reduction was seen at 1 month irrespective of the type of institution that had managed the patients before admission, although the number of drugs on admission and the degree of reduction differed between the types of institutions. However, there was a large variation in the reduction of prescribed drugs between the facilities, presumably due to differences in the overall philosophy of the attending physicians and the disease and/or functional status of the patients. Patients with and without events during a 3-month period were analyzed separately (Table 2). They were taking a comparable number of medications on admission. The number of drugs in the patients with events was not significantly decreased at 1 month, and was rather increased at 3 months after admission because in many cases additional drugs were prescribed for treatment of events.

Discontinued drugs and inappropriate prescribing

Categorized by therapeutic class, discontinuation was frequent with neuropsychologic (121 cases), gastrointestinal (116 cases) and cardiovascular (94 cases) drugs, followed by metabolic/endocrine drugs (36 cases). Anti-ulcer drugs (44 cases) including H2 blockers and prostaglandin analogs, antipsychotics (35 cases), antihypertensives (33 cases) including calcium channel blockers, β blockers and angiotensin converting enzyme inhibitors, hypnotics (31 cases), laxatives (31 cases) and non-steroidal anti-inflammatory drugs (22 cases) were frequently withdrawn.

On admission, inappropriate prescribing was seen in 58 patients (10.0% of 581 patients taking drugs). Ticlopidine, digoxin at more than 0.125 mg/day and oxybutynin were prescribed in five or more cases (Table 3). Inappropriately prescribed medications were reduced by 33% 1 month after admission, and did not change

Table 3 Number of inappropriately prescribed drugs on admission and 1 month after admission

Medication	On admission	After 1 month
Ticlopidine	36	25
Digoxin†	11	8
Oxybutynin	5	4
Amitriptyline	4	2
Benzodiazepines‡	3	1
Disopyramide	1	1
Indomethacin	1	1
Total	61	41

†More than 0.125 mg/day; ‡Flurazepam, Chlordiazepoxide and Diazepam

thereafter (data not shown). The reduction was not restricted to specific drugs.

Events during admission

A total of 104 events were seen in 16.7% of the patients during a 3-month admission period. Frequent events (nine cases or more) were new occurrences or worsening of psychological disorders (14 cases); gastrointestinal symptoms (12 cases); respiratory problems, including aspiration, pneumonia and respiratory failure (10 cases); pyrexia and infection other than pneumonia (10 cases); and falls and fractures (nine cases).

Five cases of ADWEs were found in 2.2% of 230 patients with drug reduction. These included three cases of confusion following discontinuation of psychotropic drugs, a case of depression following discontinuation of antidepressants and a case of hyperglycemia following discontinuation of a sulfonylurea.

Subgroups analyses were performed to examine the bias effect on events. The rates of events by type of institutions before admission were 24.5% in the subjects from acute care hospitals, 18.1% in those from outpatient clinics (not significant compared to other groups) and 13.1% in those from other types of institutions ($P < 0.05$ versus the subjects from acute care hospitals). Specific types of events were not related to the higher rate of events in the subjects from acute care hospitals, suggesting that unstable conditions of these patients may play a role. Of five cases with ADWEs, three were found in the subjects from outpatient clinics, one from special nursing homes and one from sanitarium-type wards. Thus, it is likely that possible non-compliance in outpatients or types of institutions before admission did not influence the principal results concerning ADWEs.

The subjects in facilities A and B (Table 2), in which significant drug reduction was not observed, showed a higher rate of events than those in other facilities (28.1% versus 12.9%, $p < 0.001$). This result indicates that adverse drug reactions associated with polypharmacy would have been included in these events, or additional drugs would have been prescribed for treatment of events, although no specific type of events was noted regarding the difference between the facilities. No ADWEs were found in the subjects in facilities A and B, presumably relating to the continuation of medications.

Discussion

The present study showed that the number of prescribed drugs was significantly decreased within 1 month after admission to health service facilities for the elderly. Discontinuation was not limited to inappropriate prescribing, but a larger proportion of inappropriately prescribed medications were discontinued compared to the total reduction of prescribed drugs (33% versus 17%). ADWEs were not frequent, being found in only 2.2% of the patients with drug reduction, while unrelated events occurred in 16.7% of the total patients.

Reflecting on the high incidence of polypharmacy and adverse drug reactions in elderly patients,[1–3,14] the principal finding of the present study that prescribed drugs can be reduced safely in frail elderly patients provides important information on pharmacotherapy. Every physician may make an effort to prescribe the minimum number of drugs, but a patient's long history of illness results in the accumulation of prescribed drugs together with the uncertain efficacy of the drugs. Consequently, the necessity of each medication should be reviewed regularly according to evidence-based medicine and criteria for inappropriate prescribing.[5,6] There is a great opportunity to reconsider prescriptions when attending physicians and/or the insurance system change, as was the case with the present study.

The number of prescribed drugs on admission in this study was smaller than that found in the geriatric ward of our university hospital and that found in residential care/assisted living facilities in the USA.[2,9] This may be because nearly half of the subjects were admitted from long-term care hospitals or facilities, and thus, prescribed drugs had already been restricted. In fact, patients from acute care hospitals were taking more drugs than those from other types of institution. It is interesting that an older age was associated with a smaller number of prescribed drugs, and this did not change when the data were analyzed according to the type of institution from which the patients had come (data not shown). This finding is inconsistent with previous observations in hospitalized or community-dwelling patients,[2,14,15] but is reasonable to prevent non-compliance and adverse drug reactions. At the same time, however, the age-related decrease in medications may involve possible discrimination towards very old people. The smaller number of prescribed drugs in women and discontinuation of medications after

admission in this study are inconsistent with previous reports,[16,17] and may imply age and sex discrimination, although discontinuation seemed successful in this study. Thus, the discrimination issue should also be taken into consideration concerning pharmacotherapy in older people.

In the present study, ADWEs were fewer than the previously reported study in which 26% of cases of discontinuation led to ADWEs in elderly outpatients during a 4-month period.[10] One of the reasons that ADWEs were rare in the present study might be that the rate of reduction was relatively high for inappropriately prescribed medications, although most of the attending physicians did not know the criteria for inappropriate prescribing such as Beers' list.[5] Another reason is that consultant physicians or geriatricians made decisions on prescriptions, based on the disease and functional status of each patient. In fact, most of discontinued drugs were not on the list of inappropriate prescribing, implying that unnecessary drugs had been prescribed before admission to long-term care facilities. In addition, it is possible that we missed ADWEs that progressed very slowly and manifested after the follow-up period of 3 months. We also failed to address the effect of prophylactic medications such as antiplatelet and lipid-lowering agents.

It should be kept in mind that the present results were obtained in a frail elderly population admitted to long-term care facilities where most of the subjects were in a stable state with chronic illness.[12] However, as a model to investigate the effect of drug reduction in elderly people, the present findings will add new insight into pharmacotherapy in the elderly, and should be confirmed in different settings such as hospitals and outpatient clinics. Obviously, medications for acute illness should neither be decreased, nor should physicians hesitate to initiate them even in very old patients, and in fact, prescribed drugs were increased in the patients with events during admission in this study. To safely apply the findings of the present study to clinical practice, knowledge of the criteria for inappropriate prescribing should be widely distributed, and blanket discontinuation of drugs must be avoided. In the present study, we used the Beers' criteria to determine inappropriately prescribed medications because corresponding criteria do not exist in Japan.[5] Consequently, we failed to check many inappropriate drugs that are used in Japan but are not on the Beers' list or sold in the USA. Future investigation using the Japanese criteria for inappropriate prescribing, which the Japan Geriatrics Society is going to establish, will add more information. In Japan, the fixed payment insurance system has begun to cover elderly patients, with the expansion of the elderly population and medical expenses. Therefore, it is essential to establish an effective and safe way to refine the use of medication in elderly people in terms of prevention of adverse drug reactions and ageism.

Acknowledgment

This work was supported by a Research Grant for Longevity Sciences (14C-4) from the Ministry of Health, Labour and Welfare, Japan.

References

1 Hanlon JT, Shimp LA, Semla TP. Recent advances in geriatrics: drug-related problems in the elderly. *Ann Pharmacother* 2000; **34**: 360–365.
2 Akishita M, Toba K, Nagano K, Ouchi Y. Adverse drug reactions in older people with dementia. *J Am Geriatr Soc* 2002; **50**: 400–401.
3 Rothschild JM, Bates DW, Leape LL. Preventable medical injuries in older patients. *Arch Intern Med* 2000; **160**: 2717–2728.
4 Phillips DP, Christenfeld N, Glynn LM. Increase in US medication-error deaths between 1983 and 1993. *Lancet* 1998; **351**: 643–644.
5 Beers MH. Explicit criteria for determining potentially inappropriate medication use by the elderly. An update. *Arch Intern Med* 1997; **157**: 1531–1536.
6 Hanlon JT, Schmader KE, Samsa GP *et al.* A method for assessing drug therapy appropriateness. *J Clin Epidemiol* 1992; **45**: 1045–1051.
7 Spore DL, Mor V, Larrat P, Hawes C, Hiris J. Inappropriate drug prescriptions for elderly residents of board and care facilities. *Am J Public Health* 1997; **87**: 404–409.
8 Hanlon JT, Schmader KE, Boult C *et al.* Use of inappropriate prescription drugs by older people. *J Am Geriatr Soc* 2002; **50**: 26–34.
9 Sloane PD, Zimmerman S, Brown LC, Ives TJ, Walsh JF. Inappropriate medication prescribing in residential care/assisted living facilities. *J Am Geriatr Soc* 2002; **50**: 1001–1011.
10 Graves T, Hanlon JT, Schmader KE *et al.* Adverse events after discontinuing medications in elderly outpatients. *Arch Intern Med* 1997; **157**: 2205–2210.
11 Hanlon JT, Weinberger M, Samsa GP *et al.* A randomized, controlled trial of a clinical pharmacist intervention to improve inappropriate prescribing in elderly outpatients with polypharmacy. *Am J Med* 1996; **100**: 428–437.
12 Hirakawa Y, Masuda Y, Uemura K *et al.* Current admission policies of long-term care facilities in Japan. *Geriatrics Gerontol Int* 2003; **3**: 73–78.
13 Naranjo CA, Busto U, Sellers EM *et al.* A method for estimating the probability of adverse drug reactions. *Clin Pharmacol Ther* 1981; **30**: 239–245.
14 Stewart RB, Cooper JW. Polypharmacy in the aged. Practical solutions. *Drugs Aging* 1994; **4**: 449–461.
15 Linjakumpu T, Hartikainen S, Klaukka T, Veijola J, Kivela SL, Isoaho R. Use of medications and polypharmacy are increasing among the elderly. *J Clin Epidemiol* 2002; **55**: 809–817.
16 Martin I, Hall J, Gardner T. Prescribing for patients aged 65 years and over in New Zealand general practice. *N Z Med J* 2002; **115**: U221.
17 Koopmans RT, van der Borgh JP, Bor JH, Hekster YA. Increase in drug use after admission to Dutch nursing homes. *Pharm World Sci* 2003; **25**: 30–34.

Geriatrics and Gerontology International 2005; 5: 37–47

ORIGINAL ARTICLE

Reliability of a geriatric assessment instrument with illustrations

Jiro Okochi,[1] Tai Takahashi,[2] Kiyoshi Takamuku,[3] Shinya Matsuda[1] and Yasuo Takagi[4]

[1]University of Occupational and Environmental Health, Kitahyushu, [2]International University of Health and Welfare, Ohtawara, [3]Souseien Geriatric Health Services Facility, Nakatsu and [4]University of Kyushu, Fukuoka, Japan

Background: The Typology of the Aged with Illustrations is a four-scale geriatric assessment measure of elderly function and activity of daily living. Illustrations are incorporated to facilitate the understanding of elderly function and activity status. The purpose of the study was to test the reliability and validity of the instrument and the effectiveness of the illustrations in improving performance.

Methods: Reliability of the typology was investigated using the test–retest procedure, with a sample of 240 elderly persons. The effectiveness of the illustrations was measured, using the instrument with and without the illustrations, by kappa statistics and the χ^2 test. The results were stratified to assess differences in reliability between experienced and non-experienced evaluators, and for elderly persons in residential institutions compared with day care facilities. Both constructive and concurrent validities were also evaluated.

Results: The intra-rater reliability (re-assessment by the same evaluator 1 week later) of the instrument was increased significantly on all four scales (average kappa statistics: 0.82 with illustrations, 0.65 without illustrations) with the inclusion of illustrations vs no illustrations, and the inter-rater reliability (re-assessment by a second evaluator within 24 hours) was significantly increased for the mobility and eating scales. There was no difference in reliability for experienced and inexperienced evaluators, or for different care situations.

Conclusion: The inclusion of illustrations in the instrument facilitated the understanding of elderly persons' activity of daily living and increased its reliability. This instrument has wide applicability and can be used for communication between health care professionals and for on-the-job training in geriatric functional assessment.

Keywords: activity of daily living, geriatric assessment, reproducibility of results, validity, visual aids.

Introduction

With the implementation of the long-term care insurance law (LTCI) in Japan in 2000, geriatric assessment became an obligation for care service providers as a means to improving the quality of care.[1] Assessment instruments incorporating a large number of measuring scales have been translated from other languages, or developed by professional associations, and have been adopted in care-plans for usage in LTCI care management.[2] According to Ministry of Health, Labour and Welfare (MHLW) figures, the average time required for a geriatric assessment just prior to the implementation of LTCI was 1.8 h per person. With the prospect of an increasing workload under the LTCI, the Typology of

Accepted for publication 22 July 2004.

Correspondence: Dr Jiro Okochi, Department of Preventive Medicine and Community Health, University of Occupational and Environmental Health, 1-1 Iseigaoka Kitakyushu 807-8555, Japan. Email: PXU14045@nifty.ne.jp

the Aged with Illustrations (TAI) was developed to provide for effective and efficient care-management.[3,4] TAI measures function and activity status for mobility (and transfer), mental (cognitive) status, toileting and eating, through observation and interview of the elderly person (see appendices). Each scale is scored on a six-point scale from 0 to 5, ranging from total loss of function to no limitation, respectively. Each level of TAI corresponds to a different level of care, and thus the care-manager can immediately predict the type of care the elderly person requires. For example, an elderly person with a TAI mobility scale score of 3 requires indirect assistance but does not require direct help from others, as can be seen from the relevant illustration in the Appendix Figure 1.

A TAI based computer system incorporating this aspect has been developed and is widely used by Japanese care-managers.

The instrument has several further advantageous characteristics. First, it uses only four basic scales to stratify activity and functions, based on the logic that an elderly person who can perform a difficult task is also able to perform an easier task. For example, the elderly person who can climb stairs can also walk on a flat floor, and this elderly person can immediately be classified as 5 (see Appendix Fig. 1); the evaluator does not have to check for function at lower levels in most cases. A significant advantage of the use of only a few scales is the ease of repetition for the collection of longitudinal data (Jiro Okochi, unpubl. data).

Second, TAI has high predictive validity as regards resource utilization in the institution. A large scale time study using 1260 institutionalized elderly persons, which combined the results of the four scales with a simple assessment of medical service use, showed an overall explanatory variance of 0.63.[3] This figure is as good as that achieved by the case-mix classification system used in the Japanese LTCI (0.62, according to MHLW figure), which employs 79 variables.

Finally, the illustrations accompanying each level of the instrument seem to facilitate an understanding of elderly function assessment.

These are all significant advantages, but a reliability and validity study of TAI remained to be undertaken. This study was designed to answer three questions: First, what is the reliability and validity of the TAI instrument? Second, do the illustrations in fact increase the reliability of the instrument? Third, does experience in its administration, or institutional vs day care of the elderly person, affect its reliability?

Methods

This study uses three methodologies, namely test–retest procedures in the reliability study, constructive validity testing and concurrent validity testing.

The assessment protocol

The 'assessment with illustrations' protocol is presented in the Appendix Figures 1–4. This assessment protocol represents a revision by the authors of the original version,[3,4] in accordance with the principles of the International Classification of Functioning, Disability and Health (ICF).[5] It was initially translated by one of the authors (JO) into English and then back translated by different author (KT) into Japanese to test for differences between the English and Japanese versions. The English version presented in this report incorporated minor modifications after back translation. The non-illustrated protocol is identical apart from the deletion of all illustrations.

Reliability testing

The participants in the reliability study were 240 elderly persons (aged over 60) resident in four long-term care facilities or attending four day-care services in Kyushu, Japan. Thirty elderly persons were randomly selected from the client lists of each of the eight establishments. The elderly persons or their proxies were asked to provide written informed consent to participate in this study because of the heavy burden of repeated measurement. In case of participation refusal, a replacement was newly recruited until each facility provided data for 30 elderly persons.

Assessment was carried out by 28 volunteer institutional nurses, social workers, physiotherapists and occupational therapists selected from the various departments of the service providers in such a way that no prior contact with the elderly subject had occurred, to avoid the possibility of bias from accumulated knowledge.

In the inter-rater reliability study, each elderly person was assessed by two evaluators using the illustrated protocol and two separate evaluators using the protocol without illustrations. All measurement was performed independently without referring to the result of other measurers. The evaluations (observation and interview) were all carried out within 24 h of each other. Therefore each of the 240 participants was measured twice with the illustrated instrument and twice with the instrument without illustrations.

In the intra-rater reproducibility study, each subject was re-assessed a week after the first assessment by one of each of his or her two evaluators who used the illustrated and non-illustrated instrument in the inter-rater reliability study. Each evaluator used the same protocol as before, that is, either with or without illustrations. Thus all 240 elderly participants were reassessed using both the illustrated ($n = 240$) and the non-illustrated instrument ($n = 240$).

To permit the test of the influence of previous experience with the TAI on reliability, we recruited

evaluators from one residential ($n = 40$) and one day-care ($n = 40$) institution where the TAI instrument had not been used previously, and additionally, new healthcare workers who did not have previous experience with TAI from one residential ($n = 40$) and one day-care ($n = 40$) institution where TAI had been used routinely. In another two residential and two day-care institutions, TAI had been used routinely. Thus half the evaluators had previous experience of TAI and the other half did not. Each evaluator was given a brief written explanation of the assessment protocol, which they were required to use without seeking assistance. No special introduction to or explanation of the protocol, by lecture or personal instruction, was given in advance of the assessment procedure.

Cohen's kappa and weighted kappa values were computed to compare the reliability of the illustrated and non-illustrated versions of the measure,[6] using STATA for Windows.

To test the effects of the testing location (institution or day-care) and of the evaluator's experience with the TAI instrument, we stratified the results on these parameters.

The relative reliability of the illustrated and non-illustrated protocols was also tested in terms of accordant and discordant cases in the individual scales, using a 2×2 table and the χ^2 test of significance.

Validity testing

Constructive validity was examined with the participation of 18 volunteer healthcare professionals, including nurses, occupational and physiotherapists and social workers.

For each of the four scales, the levels of performance were disconnected and randomly reordered. The volunteers were asked to reorganize the disconnected levels in the four scales in order of the severity of functional limitation. The rate of correct completion and the patterns of failure were recorded.

Concurrent validity was tested by the assessment of 3046 elderly recipients of the care service in Kitakyushu city, using both the Functional Independence Measure (FIM)[7] and TAI, and comparing the outcomes using Spearman's correlation coefficient.

Results

Reliability

The volunteers for assessment in the reliability study were 157 women and 83 men (total = 240), with average ages of 79.4 (SD 8.9) and 83.8 (SD 7.8), respectively.

Table 1 shows the results of the inter-rater reproducibility study. The average Cohen's kappa and weighted kappa values were 0.47 and 0.65, respectively, for the illustrated protocol. Although these values were superior to those for the protocol without illustrations, with the exception of the toileting function, the 95% confidence intervals overlapped and the difference did not reach statistical significance ($P > 0.05$).

Table 2 shows the results for intra-rater reliability (re-rating by the same evaluator a week later). Data for 26 of

Table 1 Inter-rater reproducibility of the Typology of the Aged with Illustrations instrument with and without illustrations ($n = 240$)

Function	Kappa		Weighted kappa	
	With illustrations	Without	With illustrations	Without
Mobility	0.47 (0.40–0.54)	0.43 (0.36–0.50)	0.68 (0.60–0.76)	0.64 (0.56–0.72)
Mental	0.41 (0.34–0.48)	0.36 (0.29–0.42)	0.64 (0.56–0.72)	0.55 (0.46–0.64)
Eating	0.57 (0.48–0.66)	0.41 (0.32–0.50)	0.69 (0.60–0.78)	0.59 (0.50–0.68)
Toileting	0.42 (0.35–0.49)	0.50 (0.43–0.57)	0.59 (0.51–0.67)	0.65 (0.56–0.74)

Table 2 Intra-rater reproducibility of the Typology of the Aged with Illustrations instrument with and without illustrations ($n = 214$)

Function	Kappa		Weighted kappa	
	With illustrations	Without	With illustrations	Without
Mobility	0.82 (0.75–0.89)	0.63 (0.56–0.70)	0.88 (0.80–0.96)	0.77 (0.69–0.85)
Mental	0.76 (0.69–0.82)	0.59 (0.52–0.66)	0.83 (0.74–0.93)	0.74 (0.65–0.83)
Eating	0.86 (0.79–0.93)	0.71 (0.63–0.79)	0.86 (0.77–0.95)	0.81 (0.71–0.90)
Toileting	0.83 (0.75–0.90)	0.67 (0.60–0.74)	0.87 (0.78–0.96)	0.78 (0.69–0.86)

Kappa values for the illustrated and non-illustrated protocols were significantly different on all four scales ($P < 0.05$).

the original participants could not be obtained due to worsening of their health status or absence from day care on the re-test day.

The average kappa and weighted kappa values were 0.82 and 0.86, respectively. Reliability was better with the illustrated protocol than with the non-illustrated protocol on all scales: the 95% confidence intervals for kappa values did not overlap, and the differences were significant ($P < 0.05$). All four weighted kappa values were at a highly acceptable level, but the confidence intervals for the illustrated and non-illustrated protocols over-lapped and statistical significance was not achieved.

Table 3 shows the 2 × 2 table of concordant and discordant cases for the alternative test of inter-rater reliability of the illustrated protocol vs the non-illustrated protocol. The mobility and eating scales appeared to perform better with illustrations, but for the mental and toileting scales results were similar.

Inter-rater reproducibility for the illustrated and non-illustrated protocols was also compared using kappa statistics. Four hundred and eighty (240 + 240) pairs of results were combined to calculate the kappa values. The kappa values (with 95% confidence interval in parentheses) for the four scales were: mobility, 0.44 (0.35–0.54); mental, 0.37 (0.31–0.44); eating, 0.51 (0.44–0.59); and toileting, 0.48 (0.41–0.54). The weighted kappa values were: mobility, 0.65 (0.60–0.71); mental, 0.57 (0.50–0.63); eating, 0.66 (0.60–0.73); and toileting, 0.48 (0.59–0.72).

These results lie between the kappa statistics values for the illustrated and non-illustrated protocols shown in Table 1, suggesting little difference between the illustrated and non-illustrated protocols.

Tables 4 and 5 show the results of the comparisons of inter-rater reliability for residential and day-care settings, and for level of experience of the evaluator (experienced vs inexperienced). Although none of the kappa statistics indicated statistical significance, reliability tended to be higher in residential than in day-care settings.

Validity

In the test of constructive validity, all the volunteer health care professionals reconstructed the mobility scale correctly. For the volunteers, six of the 18 reconstructed the mental function scale wrongly, and their

Table 3 2 × 2 Table of evaluator concordance in illustrated and non-illustrated protocols ($n = 240$) in the inter-rater reliability study

Function	With illustration	Without illustration Disagreed	Agreed
Mobility	Disagreed	52	41 ($P < 0.001$)†
	Agreed	45	102
Mental	Disagreed	59	48
	Agreed	59	74
Eating	Disagreed	26	32 ($P < 0.05$)†
	Agreed	53	129
Toileting	Disagreed	47	59
	Agreed	44	90

†Pearson's χ^2 test.

Table 4 Inter-rater reliability stratified by care setting (day care $n = 120$, institutional care $n = 120$)

		Kappa	Weighted kappa
Mobility	Day care	0.34 (0.21–0.48)	0.52 (0.40–0.63)
	Institutional care	0.56 (0.45–0.66)	0.75 (0.63–0.87)
Mental	Day care	0.34 (0.23–0.44)	0.55 (0.44–0.67)
	Institutional care	0.44 (0.33–0.56)	0.65 (0.52–0.77)
Eating	Day care	0.63 (0.48–0.77)	0.62 (0.27–0.97)
	Institutional care	0.51 (0.39–0.64)	0.69 (0.61–0.78)
Toileting	Day care	0.41 (0.29–0.53)	0.50 (0.30–0.69)
	Institutional care	0.39 (0.28–0.49)	0.58 (0.46–0.69)

The illustrated version of the Typology of the Aged with Illustrations was employed.

pattern of failure was the same, namely mistaking the order of levels 1 and 2. In the eating scale, two volunteers made errors, both mistaking levels 1 and 2. For the toileting scale, three mistook the order of levels 2 and 3.

For the test of concurrent validity, the TAI measure was completed for 2756 of the sample of 3048 elderly participants (90%), and the FIM measure for 2631 (86%). Both the TAI and FIM observations were completed for 2483 (81%). Their average age was 75 (SE 10) for men ($n = 691$), and 78 (SE 9) for women ($n = 1796$). All of the Spearman's rank coefficients for the correlation of the TAI and the FIM subscales, shown in Table 6, were statistically significant ($P < 0.01$) but ranged from 0.23 to 0.72. The TAI mobility scale showed high correlation with FIM subscales of wheelchair, stairs, bed/chair, tub/shower, bathing. The TAI mental scale correlated strongly with the FIM subscale of memory. The TAI eating scale correlated strongly with the FIM subscale of eating. In TAI, toileting correlated strongly with the FIM subscale of toilet, dress lower body, toileting and bladder management.

Table 5 Inter-rater reliability stratified by experience of evaluator (experienced evaluators $n = 120$, inexperienced evaluators $n = 120$)

		Kappa	Weighted kappa
Mobility	Experienced	0.48 (0.39–0.57)	0.69 (0.58–0.80)
	Inexperienced	0.46 (0.36–0.56)	0.67 (0.57–0.78)
Mental	Experienced	0.40 (0.32–0.48)	0.63 (0.51–0.74)
	Inexperienced	0.44 (0.34–0.53)	0.64 (0.52–0.77)
Eating	Experienced	0.64 (0.51–0.76)	0.74 (0.65–0.83)
	Inexperienced	0.51 (0.38–0.64)	0.69 (0.51–0.77)
Toileting	Experienced	0.33 (0.23–0.43)	0.52 (0.40–0.63)
	Inexperienced	0.52 (0.42–0.61)	0.66 (0.54–0.77)

The illustrated version of the Typology of the Aged with Illustrations was employed.

Table 6 Spearman's rank correlation coefficients for comparison of the Typology of the Aged with Illustrations (TAI) and Functional Independence Measure (FIM) ($n = 2487$)

FIM subscales		TAI scales			
		Mobility	Mental	Eating	Toileting
Locomotion	Walk/wheelchair	**0.70**	0.25	0.45	0.52
	Stairs	**0.72**	0.23	0.41	0.50
Transfer	Bed/chair	**0.65**	0.29	0.50	0.57
	Toilet	**0.61**	0.31	0.57	**0.62**
	Tub/shower	**0.67**	0.31	0.51	0.58
Communication	Comprehension	0.28	0.50	0.35	0.36
	Expansion	0.34	0.41	0.41	0.39
Social cognition	Social interaction	0.35	0.53	0.39	0.43
	Problem solving	0.34	0.59	0.38	0.43
	Memory	0.29	**0.62**	0.35	0.41
Self-care	Eating	0.44	0.35	**0.68**	0.52
	Grooming	0.52	0.40	0.58	0.59
	Bathing	**0.64**	0.33	0.51	0.58
	Dress upper body	0.57	0.35	0.61	0.59
	Dress lower body	0.59	0.35	0.58	**0.62**
	Toileting	0.57	0.37	0.60	**0.68**
Sphincter control	Bladder management	0.50	0.38	0.54	**0.67**
	Bowel management	0.50	0.37	0.56	**0.67**

All correlations were statistically significant ($P < 0.01$).
Values that exceed 0.6 are shown in **bold type**.
The illustrated version of TAI was employed.

Discussion

The use of illustrations has progressively evolved in medical education and practice.[8] For example, illustrations are regularly employed in the assessment in young children.[9–11] An instrument with illustrations for more extensive functional measurement, to be used in the physician's office, has also been developed.[12]

A trial to incorporate illustrations in health outcome measurement showed that illustrations improved 1 week test–retest reliability.[13]

Our results indicate that Typology of the Aged with Illustrations is a reliable instrument. The inclusion of illustrations increased intra-rater reliability significantly on all four scales, as indicated by kappa values. Inter-rater reliability was also improved by the use of illustrations in the scales of mobility and eating function. These results indicate that the inclusion of illustrations in measuring instruments can improve the reliability of geriatric assessment.

However, their inclusion produced no significant difference in inter-rater reliability for the mental status and toileting scales. For these scales the evaluator must collect information on elderly orientation, problem behavior and toileting function. This information is not obtained by observation only, but also requires direct questioning of the elderly person and/or his or her proxy. This might be the reason why the reproducibility of mental status and toileting scales did not improve in inter-rater reliability study. Therefore, a more elucidating text attached to the illustration is probably necessary to improve reliability.

Although any direct comparison with other studies in different settings must be interpreted with caution, the average kappa value of 0.47 for the illustrated instrument seems low compared with the results of another instrument measuring function and activity,[7] in which the FIM kappa ranged from 0.53 (memory) to 0.66 (stair climbing). However, in this study the evaluators were clinicians who already knew the patient well. In a study using the Barthel index, reliability was compared between nurses and non-clinical researchers, producing kappa values ranging from 0.27 (bowels) to 0.68 (bathing).[14] Our study, which was performed under strict controls, employing both experienced and nonexperienced evaluators, without prior introduction to the instrument and avoiding possible effects of prior knowledge of the subject, still showed an acceptable level of reliability.

Because health care professionals with various backgrounds are engaged in long-term care of the elderly, this characteristic of the TAI is advantageous to its use as an in-house and external communication tool. Health care professionals who are not accustomed to the functional assessment of elderly persons can measure their status without prior education. The TAI could also be usefully employed as an educational tool for on-the-job training. However, although this study showed that inexperienced health care professionals perform reliably with the TAI, we still lack evidence as to whether it could be successfully used by family members or the elderly persons themselves.

This study failed to satisfactorily answer the question as to whether the inclusion of illustrations was essential to the success of the instrumental design. The next step towards answering this question is to set a gold standard with which to compare the illustrated and non-illustrated protocols.

The constructive validity of the instrument, while primarily confirmed by the test of its reproduction by health care professionals, failed on some items. For the mobility scale, there was perfect reproduction, but for the mental status, eating and toileting scales, there was different re-ordering of two items by a minority of participants, with the same error made in every case. The pattern of failure could therefore be considered systematic. Errors were most frequent for the mental status scale. This may be explained by the fact that the burden of care is more intense when elderly persons exhibit problem behavior, but that in the course of mental deterioration with Alzheimer's disease, low cognitive performance is decreasingly associated with impulsive behavior after 2 years.[15] These variations indicate that the TAI scales must be interpreted carefully in terms of the severity of dysfunction indicated. The thresholds described in the appendices should, nevertheless, be an aid to the correct assessment of elderly function.

In the test of concurrent validity, some FIM subscales correlated highly with one or other of the TAI scales. The areas of correlation indicate the major areas of TAI coverage, namely the TAI mobility scale with the FIM subscales of wheelchair, stairs, bed/chair, tub/shower, bathing. The TAI mental scale correlated strongly with the FIM subscale of memory. The TAI eating correlated strongly with the FIM subscale of eating. TAI toileting correlated strongly with the FIM subscale of toilet, dress lower body, toileting and bladder management. TAI, however, is not an interval scale describing level of function as a percentage, but rather, an index of functional and activity status.

We are currently undertaking a large-scale cross-sectional study with the ICF, with a view to improvements.

Finally, the largest advantage of this instrument is its simplicity. In the longitudinal studies, many ADL and function related assessments were employed,[16,17] which resulted in high-cost to repeat measurement. The authors are currently performing longitudinal studies to reveal trajectories of functional decline among community dwelling elders using this advantage of the instrument.

Conclusion

TAI is a simple instrument for use in the long term care of the aged, employing only four basic scales.

The addition of illustrations to the instrument increases the intra-rater reliability of all four scales, and the inter-rater reliability of the movement and eating scales. No difference was found in the reliability of experienced and inexperienced users. Therefore, the TAI can be used as a basic language of communication among the health care professionals, and also as an educational tool for the functional and ADL assessment of elderly persons.

Acknowledgments

This study was supported by grants from the Japanese Ministry of Health (Medical Frontier Strategy Research-H13-Chihou/Kossetu-019), Labour and Welfare, and the Japanese Ministry of Education and Science (14770190).

The authors are grateful to the health care workers at geriatric facilities for the elderly and in their affiliated day care services, namely Altenheim Yokokura, Oyakuen, Sunlife Minori and Souseien, for data collection. The authors are also grateful for the methodological and statistical support from Dr Kouhei Akazawa of Niigata University, School of Medicine, Department of Medical Informatics and Dr Kenji Toba of Kyorin University, School of Medicine, Department of Geriatric Medicine and to Mr Kazumi Nakae of Tokyo Design Inc. for the artwork.

References

1 Matsuda S. The health and social system for the aged in Japan. *Aging Clin Exp Res* 2002; **14**: 265–270.
2 Masuda Y, Kuzuya M, Uemura K *et al.* The effect of public long-term care insurance plan on care management and care planning in Japanese geriatric hospitals. *Arch Gerontol Geriatr* 2001; **32**: 167–177.
3 Takahashi T, Okochi J, Takamuku K, Matsuda S. The introduction of typology of the aged with illustrations. *Casemix Q* 2001; **3**: 3–14.
4 Takahashi T, Takamuku K, Okamoto S, Abe N. *[Typology of the Aged with Illustrations.]* Tokyo: Nikkei BP, 1997. (In Japanese.)
5 World Health Organization. *International Classification of Functioning, Disability and Health.* Geneva: WHO, 2001.
6 Fleiss J, Levine B, Cho Paik M. *Statistical Method for Rates and Proportion*, 3rd edn. New Jersey: Wiley and Sons, Inc, 2003.
7 Hamilton BB, Laughlin JA, Fiedler RC, Granger CV. Interrater reliability of the 7-level functional independence measure (FIM). *Scand J Rehabil Med* 1994; **26**: 115–119.
8 Ansary MA, el Nahas AM. Medical illustration in the UK: Its current and potential role in medical education. *J Audiov Med Med* 2000; **23**: 69–72.
9 Warren SL, Dadson N. Assessment of anxiety in young children. *Curr Opin Pediatr* 2001; **13**: 580–585.
10 Praver F, DiGiuseppe R, Pelcovitz D, Mandel FS, Gaines R. A preliminary study of a cartoon measure for children's reactions to chronic trauma. *Child Maltreat* 2000; **5**: 273–285.
11 Valla JP, Bergeron L, Berube H, Gaudet N, St-Georges M. A structured pictorial questionnaire to assess DSM-III-R-based diagnoses in children (6–11 years): Development, validity, and reliability. *J Abnorm Child Psychol* 1994; **22**: 403–423.
12 Nelson EC, Landgraf JM, Hays RD, Wasson JH, Kirk JW. The functional status of patients. How can it be measured in physicians' offices? *Med Care* 1990; **28**: 1111–1126.
13 Hadorn DC, Hays RD, Uebersax J, Hauber T. Improving task comprehension in the measurement of health state preferences. A trial of informational cartoon figures and a paired-comparison task. *J Clin Epidemiol* 1992; **45**: 233–243.
14 Richards SH, Peters TJ, Coast J, Gunnell DJ, Darlow MA, Pounsford J. Inter-rater reliability of the Barthel ADL index: How does a researcher compare to a nurse? *Clin Rehabil* 2000; **14**: 72–78.
15 McCarty HJ, Roth DL, Goode KT *et al.* Longitudinal course of behavioral problems during Alzheimer's disease: Linear versus curvilinear patterns of decline. *J Gerontol A Biol Sci Med Sci* 2000; **55**: M200–M206.
16 Ho HK, Matsubayashi K, Wada T, Kimura M, Saijoh K. Factors associated with ADL dependence. A comparative study of residential care home and community-dwelling elderly in Japan. *Geriatr Gerontol Int* 2002; **2**: 80–86.
17 Anderson RT, James MK, Miller ME, Worley AS, Longino CF Jr. The timing of change: Patterns in transitions in functional status among elderly persons. *J Gerontol B Psychol Sci Soc Sci* 1998; **53**: S17–S27.

Appendix

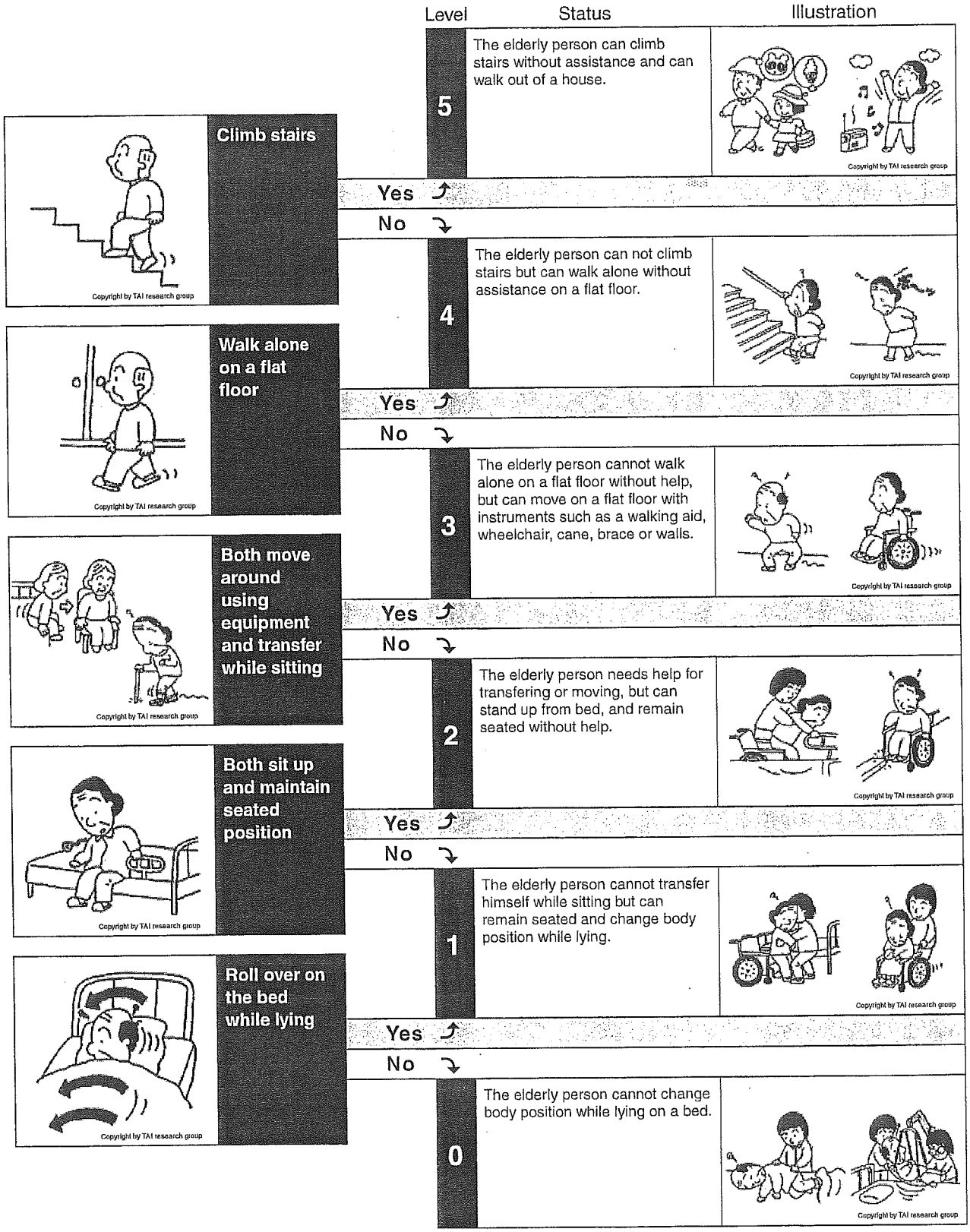


Figure 1 Mobility.

5

Status	Illustration
No disturbance in cognitive function mentioned below. 1. inappropriate behavior for two weeks 2. memory or decision making disturbance 3. Orientation disturbance	 Don't waste your money! I can read paper Copyright by TAI research group

Any disturbance in cognitive function

- No ↗
- Yes ↘

Problem behavior

4

Status	NO	YES
The elderly person has disturbance of decision making or disturbance of living caused by memory loss but shows no problem behavior	↵	↳

Illustration

Apparent forgetfulness

Copyright by TAI research group

3

Status
The elderly person does not have orientation disturbance but has behavior(s) that embarrass his family or neighbors.

Illustration

Someone took my wallet!

Copyright by TAI research group

Orientation questions
1. What month is it now?
2. Where are we now?
3. Who is this person?

- All correct ↗
- Any failure ↘

Problem behavior

2

Status	NO	YES
Severe orientation disturbance but no problem behavior	↵	↳

Illustration

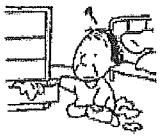
Sitting all day long

Copyright by TAI research group

1

Status
Both severe orientation disturbance and problem behavior(s)

Illustration



Restlessness and /or agressiveness

Copyright by TAI research group

Responsiveness

- Yes ↗
- No ↘

0

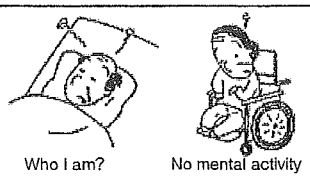
Status	Illustration
No mental activity and no problem behavior	 Who I am? No mental activity Copyright by TAI research group

Figure 2 Mental status.

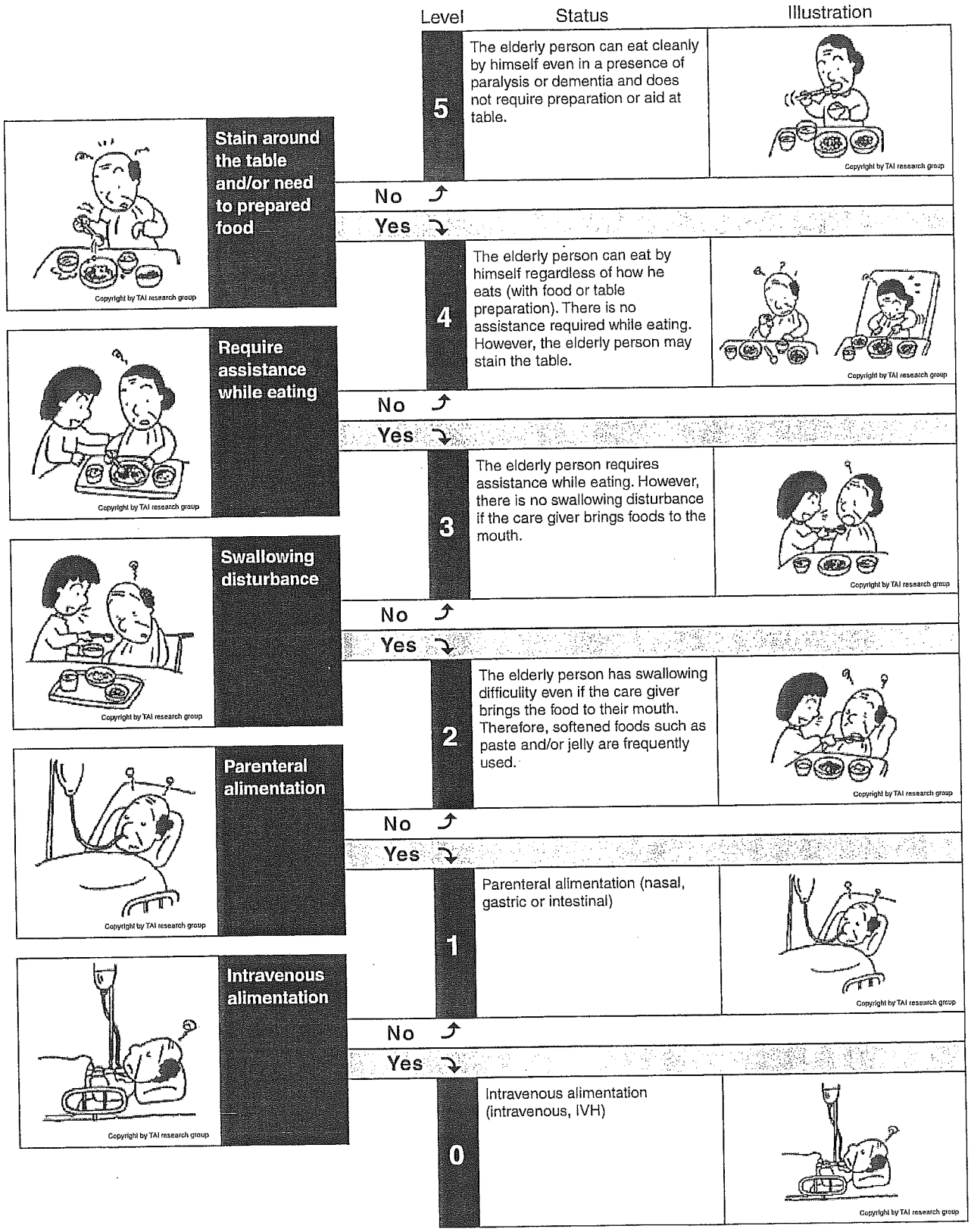


Figure 3 Eating.

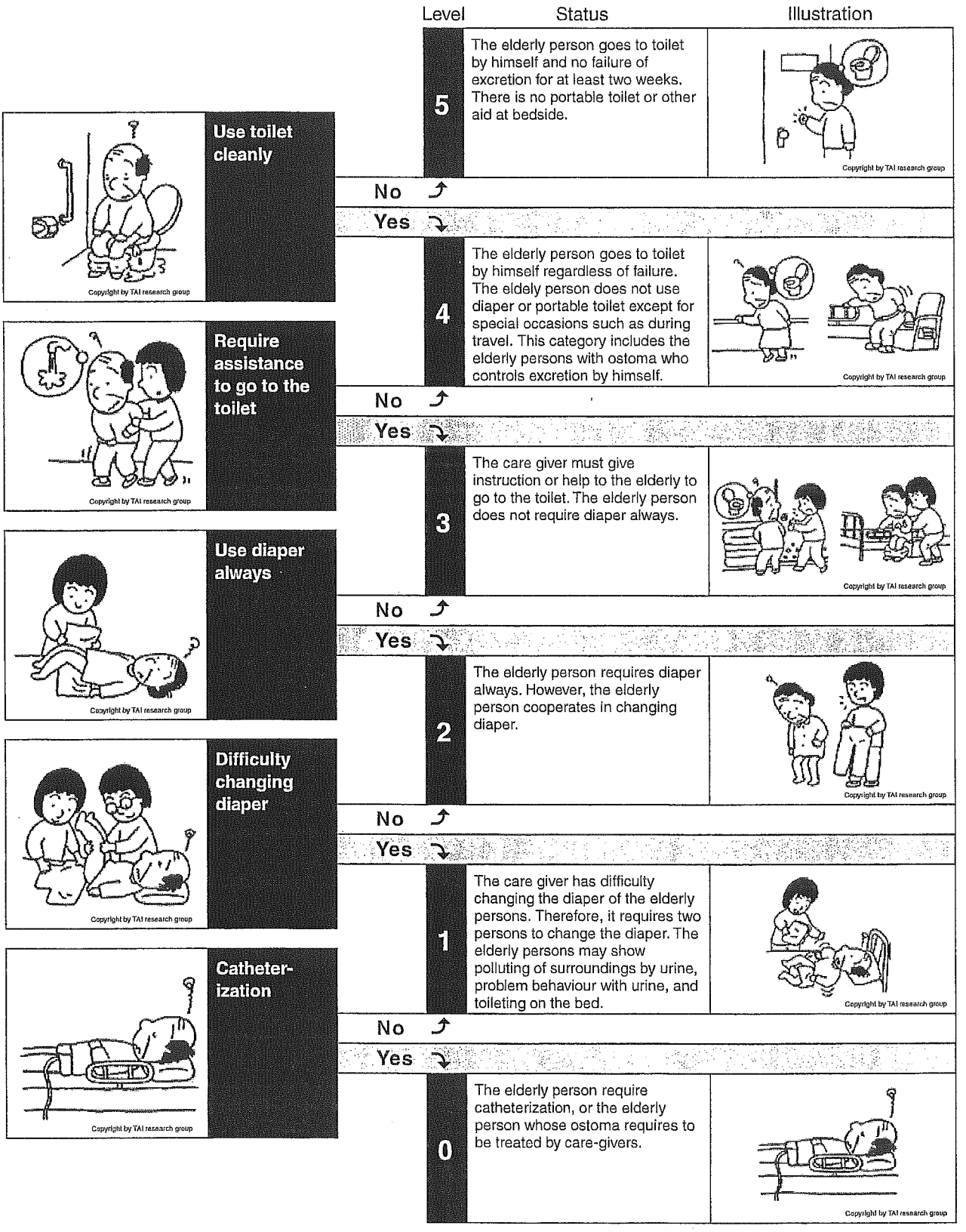


Figure 4 Toileting.

Osteoporos Int (2005) 16: 273–279
DOI 10.1007/s00198-004-1669-3

ORIGINAL ARTICLE

Trunk deformity is associated with a reduction in outdoor activities of daily living and life satisfaction in community-dwelling older people

Toshiaki Takahashi · Kenji Ishida · Daisuke Hirose
Yasunori Nagano · Kiyoto Okumiya
Masanori Nishinaga · Kozo Matsubayashi
Yoshinori Doi · Toshikazu Tani · Hiroshi Yamamoto

Received: 9 September 2003 / Accepted: 7 May 2004 / Published online: 2 July 2004

Abstract We have evaluated the association between trunk deformities of the sagittal plane and functional impairment of daily living in community-dwelling elderly subjects. The analysis involved a detailed assessment of indoor and outdoor activities of daily living, satisfaction with life, and mental status. The participants in this study were 236 community-dwelling older adults, aged 65 years and older, living in Kahoku district of Kochi in Japan. The participants were classified based on their posture, which was assessed using photographs of the subjects, and interviewed to assess their basic activities of daily living (BADL), instrumental ADL (IADL), and cognitive well-being in the cross-sectional study. The statistical analysis was performed using the Mann-Whitney *U*-test. The lumbar kyphosis group received significantly lower BADL and IADL scores than the normal group. The trunk deformity group which were defined as kyphosis, flat back, and lumbar lordosis groups exhibited decreases in activities that included going out, shopping, depositing and withdrawing money, and visiting friends in the hospital. These activities require going outdoors; thus, this study showed that the trunk deformity group had limitations in outdoor activities. There was no significant difference between the geriatric depression score (GDS) and the pattern of posture. The abnormal trunk deformity groups tended to score lower than the normal group with regard to subjective healthiness and life satisfaction measures, including subjective health condition, everyday feeling, satisfaction with human relationships, satisfaction with economic condition, and satisfaction with present life.

Keywords Activities of daily living · Kyphosis · Life satisfaction · Trunk deformity

T. Takahashi (✉) · K. Ishida · D. Hirose · Y. Nagano
T. Tani · H. Yamamoto
Department of Orthopaedics, Kochi Medical School,
Oko-cho, 783-8505 Nankoku, Kochi, Japan
E-mail: takahast@kochi-ms.ac.jp
Tel.: +81-88-8802387
Fax: +81-88-8802388

K. Okumiya · M. Nishinaga · Y. Doi
Department of Medicine and Geriatrics,
Kochi Medical School, Kochi, Japan

K. Matsubayashi
Center for Southeast Asian Studies, Kyoto University,
Kyoto, Japan

Introduction

Several studies have reported on the relationship between trunk deformity and lumbago [1,2]. It is predictable that patients with abnormal posture would be at increased risk for falling, as their balance is perturbed by the posture abnormality [3,4]. Loss of distal lumbar lordosis is the main cause of sagittal imbalance in individuals who do not maintain sagittal alignment [5]. This abnormal posture could lead to the limitation of daily activities.

There have been several evaluations of posture and functional activities to date [6]; however, very few involve elderly subjects. Ettinger et al. [7] reported that kyphotic women did not have greater back pain, disability caused by back problems, or poorer health than non-kyphotic women. Another study showed a poor correlation between quality of life and abnormal findings on radiography or densitometry [8].

Vertebral body compression fractures have been shown to be associated with the severity of kyphosis [9]. Ryan et al. [10] reported that there was a significant association between scores of osteoporosis severity and limitations in functional activity. Vertebral compression fractures associated with osteoporosis can be self-limiting, causing considerable pain and disability [8].

Vertebral compression fractures are associated with significant impairments in physical, functional, and psychosocial performance in the elderly [11,12,13]. It is crucial to improve the mental status of the elderly. However, there have been few reports regarding the

association between trunk deformity and psychogenic activity in elderly patients.

In this study, we have evaluated the daily activities and mental condition of community-dwelling elderly subjects with regard to the severity of trunk deformity in the sagittal plane.

Materials and methods

Subjects

Participants who applied in 1999 included 236 community-dwelling older adults, aged 65 years and older, living in Kahoku district in Kochi prefecture, who had been enrolled in one of several studies involving annual medical check-ups (beginning in 1994). The population of Kahoku rural area is 5800 people, 50% of whom are engaged in agricultural work.

Study participants were observed from July to August 1999, and then classified based on their posture, which was assessed using photographs taken by researchers. In total, there were 145 females (mean; 79.0 years) and 91 males (mean; 80.3 years) with a mean age of 80 years (range, 65–94 years), and a mean height of 149.1 cm. Functional status of the lumbar spine [14] and knee [15] were measured using the assessment of the Japanese Orthopaedic Association (JOA). In this study, JOA scores for assessing treatment of low back pain was calculated without incorporating urinary bladder function. Comorbidities were hypertension (31.6%), cardiac arrhythmia (6.1%), diabetes mellitus (5.7%), cerebrovascular disorder (4.2%), coronary artery disease (3.2%), senile dementia (2.1%), and Parkinson's disease (0.4%).

Methods

Interviews and examinations, conducted at the community plaza of Kahoku, consisted of a questionnaire covering physical health, functional status, and mental status. The analysis of trunk deformity was examined by photography. The presence or absence of disease was based on the subjects' self-report of a doctor's diagnosis. Informed consent was obtained from all participants.

Postural analysis

In order to protect the participants' privacy when undergoing the community health check-ups, we constructed a device with which to assess trunk posture without requiring that the subjects disrobe. Participants wore clothes typical of the summer season, and the device was equipped with a band that, when twisted, would reveal the alignment of the body. Each participant had reflective surface markers attached at various locations, including C7, T6 (xyphoid process level), L4 (Jacob line level), the left greater trochanter, the left lateral chondyle of the femur, and the left lateral malleous, as shown in Fig. 1. For the photograph, participants were positioned carefully and asked to remain relaxed while standing up straight. Posture in the sagittal plane was classified into one of the following five types: I normal, II thoracic

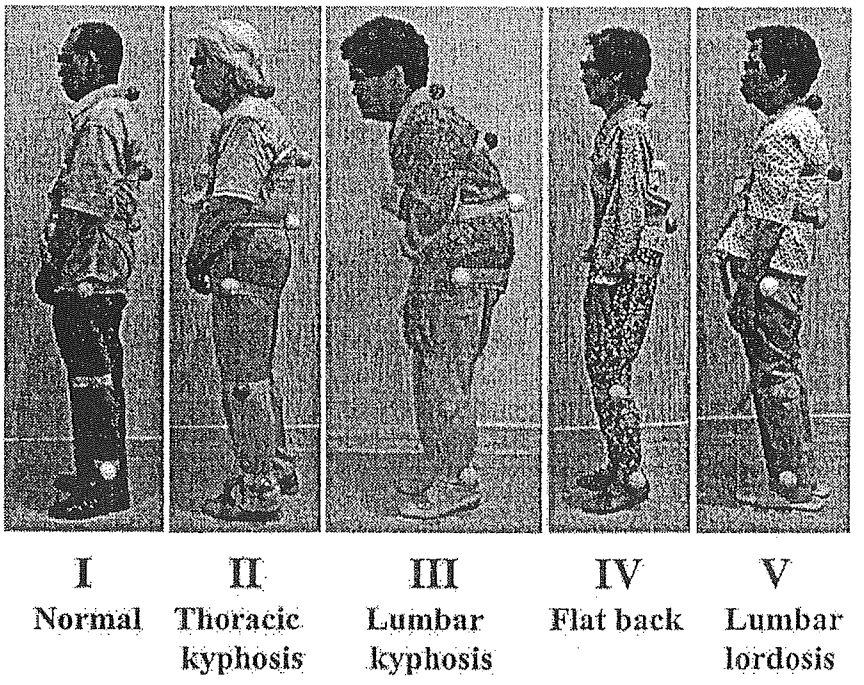


Fig. 1 Classification of trunk deformity from photographs

kyphosis, III lumbar kyphosis, IV flat back, and V lumbar lordosis, as described by Ando et al. [16,17,18]. Three orthopaedic doctors independently determined the classification, and we adopted the classification given by at least two doctors. In cases in which none of the doctors' classifications agreed, we discussed the case and jointly decided on the classification. The classification system is shown in Fig. 1, which includes photographs of patients representing each of the posture groups.

ADL analysis

The subjects were asked questions regarding basic activities of daily living (BADL) (walking, ascending and descending stairs, feeding, dressing, using the toilet, bathing, grooming, and taking medicine) and instrumental ADL (IADL) (using public transportation, shopping for groceries, preparing meals, paying bills, depositing and withdrawing money, writing, reading newspapers, reading magazines or books, taking an interest in news of health, visiting friends, giving advice to family or friends, visiting friends in the hospital, and talking to young people) [19]. We assessed the ADL score using a 4-point scale, based on the help required for each activity: 3 completely independent; 2 some help needed; 1 much help needed; and 0 completely dependent.

Mental state analysis

Geriatric Depression Scale

The Geriatric Depression Scale (GDS) [20,21], a measure of depressive symptomatology assessed on scale of 0–30, was administrated. We assessed the short form of 15 items, it is interpreted that a score >5 points is suggestive of depression, a score >10 points is almost always depression.

Visual analog scale (VAS)

Each year, we conducted an assessment of subjective quality of life (QOL), especially subjective healthiness and life satisfaction, using a validated self-reported visual analogue scale (VAS) [22]. The components of questions were subjective health condition, everyday feeling, satisfaction with human relationship to others, satisfaction with human relationship to family, satisfaction with economic condition, satisfaction with present life, and subjective happiness. The VAS questionnaire ended with a summing-up graph in the form of a 100 mm bar, graded with the subjectively worst condition on the left and the best one on the right. The subject was asked to place a mark on the 100 mm bar based on his or her condition. We defined the distance (mm) from the left to the marked position as the VAS score (0–100), with high scores indicating a high QOL [23].

Statistical analysis

For the classification of posture, Cohen's kappa coefficients were used to test statistical reliability. To determine inter-observer reliability, each reviewer's responses were compared with those of the other reviewers.

Data concerning ADL, GDS and life satisfaction were expressed by mean, SD, and SEM. The differences among the pattern of trunk deformities were evaluated using Kruskal-Wallis test, between with (II–V) and without trunk deformities (I) were evaluated using Mann-Whitney test. Differences were considered significant at $P<0.05$.

Results

The classification of trunk deformity resulted in five groups: I normal group (109 subjects; 46.2%), II thoracic kyphosis group (47 subjects; 19.9%), III lumbar kyphosis group (41 subjects; 17.4%), IV flat back group (28 subjects; 11.9%), and V lumbar lordosis group (11 subjects; 4.7%). There was a mean inter-observer kappa coefficient of 0.47 for both observation times, with a mean inter-observer agreement of 60.2%. We calculated a mean intra-observer kappa coefficient of 0.55 for the two observation times, with a mean inter-observer agreement of 68.3%. Table 1 shows the baseline characteristics in each group. There was no significant difference in age, sex, and overall health status such as comorbidities among the groups.

The mean BADL score of abnormal trunk posture (II–V) was 23.1; that of the normal (I) group was 23.6. The lumbar kyphosis group had significantly lower BADL scores than the normal group ($P=0.017$) (Table 2). With regard to BADL, walking was more likely to be limited in the abnormal trunk posture group (II–V) than in normal participants (I) ($P=0.02$).

The mean IADL score of abnormal trunk posture (II–V) was 10.3, that of the normal (I) group was 11.2. There was no significant difference in IADL among these groups ($P=0.1$) (Table 3). However, the abnormal posture groups (II–V) had lower IADL scores that differed significantly from the normal group (I) ($P=0.047$) (Table 3).

The achieved ratio of transportation of IADLs was associated with trunk deformity ($P=0.04$) (Table 4). The group with trunk deformity group had significant disturbances in certain IADLs, including transportation, shopping for groceries, depositing and withdrawing money, and visiting friends in the hospital (Table 4). Subjects with lumbar lordosis did not exhibit significant differences from the normal group, because of the small size of this group.

There was no significant difference between GDS and the pattern of trunk deformity ($P=0.70$) (Table 5). Measures of subjective healthiness and life satisfaction (Table 6), assessed using a validated, self-reported, visual analogue scale (VAS), were not significantly dif-

Table 1 Baseline characteristics of participants. All data are expressed as mean (95% confidence interval). *I* normal, *II* thoracic kyphosis, *III* lumbar kyphosis, *IV* flat back, and *V* lumbar lordosis

	I	II	III	IV	V	Total
Number	109	47	41	28	11	236
Age	78.4 (68.9, 87.8)	81.3 (71.3, 91.4)	80.8 (70.0, 91.8)	80.2 (70.9, 89.5)	80.6 (71.8, 89.5)	80
Gender (Female, Male)	55, 54	30, 17	34, 7	18, 10	7, 4	144, 92
Height	152.0 (133.5, 170.6)	145.7 (128.8, 162.7)	142.8 (126.3, 159.2)	150.8 (134.8, 166.9)	151.6 (133.4, 169.8)	149.1
Weight	54.3 (33.7, 74.9)	47.3 (33.1, 61.5)	47.4 (29.3, 65.5)	48.1 (31.0, 65.3)	53.0 (28.5, 77.5)	50.9
JOA score (lumbar)	25.6 (18.2, 33)	25.2 (18, 32.4)	24.1 (14.7, 33.5)	24.4 (15.4, 33.4)	27.5 (23.5, 31.5)	25.2
JOA score (knee)	92.0 (67.03, 116.9)	88.9 (59.18, 118.65)	88.0 (58.53, 117.45)	89.6 (66.81, 112.45)	93.3 (63.11, 123.49)	90.5
Coexisting illness						
Hypertension	34	15	17	6	2	74
Cardiac arrhythmia	5	4	4	1	0	14
Diabetes mellitus	5	2	2	3	1	13
Cerebrovascular disorder	3	2	2	1	0	10
Coronary artery disease	4	1	2	0	0	7
Senile dementia	1	2	0	2	0	5
Parkinson disease	0	0	1	0	0	1

Table 2 Total BADL score (points) by the classification of trunk deformity in comparison with the normal trunk group (I). *I* normal, *II* thoracic kyphosis, *III* lumbar kyphosis, *IV* flat back, and *V* lumbar lordosis

	Mean	SD	95% confidence intervals
I	23.6	0.9	21.8; 25.4
II	23.3	1.56	20.18; 26.42
III	22.7	3.81	15.08; 30.32
IV	23.3	1.76	19.78; 26.82
V	23.7	0.65	22.4; 25.0
II–V	23.1	2.51	18.1; 28.12

Table 3 Total IADL score (points) by the classification of trunk deformity in comparison with the normal trunk group (I). *I* normal, *II* thoracic kyphosis, *III* lumbar kyphosis, *IV* flat back, and *V* lumbar lordosis

	Mean	SD	95% confidence intervals
I	11.2	3.13	4.94; 17.46
II	10.6	3.09	4.42; 16.78
III	9.8	4.09	1.62; 17.98
IV	10	4.42	1.16; 18.84
V	12	1.41	9.18; 14.82
II–V	10.3	3.68	2.94; 17.66

ferent from normal in participants with trunk deformity ($P=0.08$). However, the abnormal trunk deformity group tended to have lower scores with regard to subjective health condition ($P=0.03$), everyday feeling ($P=0.007$), satisfaction with human relationships to family ($P=0.035$), satisfaction with economic condition ($P=0.03$), and satisfaction with present life ($P=0.051$) than those of the normal group.

Discussion

Trunk posture in the elderly, especially kyphosis, is known to be associated with vertebral compression fractures. Measurement of kyphosis may be useful in assessing the severity of spinal osteoporosis [9]. The high prevalence of back pain demonstrates the importance of pain management in the treatment of osteoporosis [24]. The number of recent vertebral fractures was also a significant predictor of poor performance in functional reach and walking speed tests [25]. Women with multiple vertebral deformities had significantly greater impairment of ADL function than women without such deformities [26].

Lyles et al. [12] showed that patients with vertebral compression fractures had reduced levels of functional performance, pain with activity, and difficulty in activities in comparison with patients that did not have fractures. Kyphosis is associated with qualitatively and quantitatively diminished function, especially with regard to the performance of mobility tasks [27]. Our results also showed that the walking activity of the