Evaluation of Effectiveness, Quality and Inequalities in Health, Medical and Long-Term Care – Achievements and Challenges –

Katsunori Kondo

Professor of Social Epidemiology and Health Policy, Center for Preventive Medical Sciences, Chiba University

Abstract

The viewpoint of cost effectiveness analysis is essential for efficiently allocating limited social security resources. To do this, effectiveness must be measured together with costs. However, evaluation studies for effectiveness and quality of health, medical and long-term care in Japan have lagged behind those in other well developed countries.

We will hereby present studies that we have conducted regarding health (preventive care), rehabilitation medicine, long-term care (quality of care at special nursing homes for elderly people) and end-of-life care. In addition, we will discuss achievements in and potentials for "visualization" in the effectiveness and quality of health, medical and long-term care, and present future challenges.

In each study, some programs led to significant effects, while others did not. Simply suppressing costs uniformly without distinguishing between effective and ineffective programs will also inhibit activities that are effective. To encourage citizens to accept the cost burden associated with expansion of social security costs, it is critical to construct frameworks for "visualization" that lead to improved care quality and equity by utilizing increased health and long-term care costs while eliminating waste. This requires the construction of large-scale databases in each field, the accumulation of health services research and the development of a management system encompassing those findings.

Keywords: Care quality, health services research, visualization, database
JEL Classification: I00, I10, I12, I32

I. Background and purpose

With increased necessity for fiscal reform, debate on suppressing social security costs has been reignited. Although some have suggested "painful suppression of social security costs" as a means of combating the fiscal crisis, few of them would completely deny the necessity of social security. Thus, it is best to carefully proceed with cost reduction measures with no (or little) pain while maintaining the functions of social security as much as possible.

Cost performance analysis is widely used in regard to public works and suchlike in Japan and other countries when selecting policies and programs that offer significant benefits, effectiveness and utility (hereinafter referred to as "effectiveness") using limited resources. This involves comparatively analyzing the ratio of cost to effectiveness, such as simultane-

ously analyzing effectiveness and costs to determine whether effectiveness exceed costs, comparing two or more choices such as A and B to decide which choice offers the greatest effectiveness with the same costs or which choice lowers costs with the same effectiveness.

Measures and services with no effectiveness at all are a waste of funds. If this could be clarified and made it sufficiently known to the citizens of Japan, not only would inhibiting or discontinuing such measures be painless, but wasteful costs could also be reduced. Therefore, demonstrating which measures have no effectiveness makes painless waste elimination possible. Meanwhile, even if painful cost reduction is unavoidable, a means of reducing the pain or damage accompanying it is to retain the portions with excellent cost effectiveness and reduce costs from the parts with poor cost effectiveness. Thus, amid the tendency for tightening up fiscal constraints in the future, the importance and necessity of cost effectiveness analysis is likely to increase as well.

Some instances of cost effectiveness analysis have been accumulated in social security fields including health, medicine and welfare (caregiving). In many countries, they have been used as one basis for determining target technology covered by public medical costs. However, in Japan, cost effectiveness analysis regarding health and long-term care is lagging behind those of medication. As a reason for that, it can be indicated that there is little accumulated research regarding it due to a weak research basis for the evaluation of (clinical) effectiveness, which is the foundation for cost effectiveness analysis.

In consideration of increased need for such evaluation, we have previously been involved in research evaluating the effectiveness and quality of health, medical and long-term care (hereinafter abbreviated to "health care"). As part of this, we have also conducted rough cost estimates. In this report, we introduce this research and examine the following issues. Firstly, we discuss why the evaluation of cost effectiveness for health care is difficult. Secondly, we introduce some cases within these limitations and indicate to what extent the evaluation of programs—and the effectiveness and quality of services in health care—is possible. The cases we present are related to health (care prevention), rehabilitation medicine, long-term care and end-of-life care. Thirdly, we will discuss challenges in promoting cost-effectiveness analysis, which is becoming increasingly necessary, and the health services research upon which it is based in Japan.

II. Why is it difficult to research cost effectiveness?

Cost effectiveness analysis offers information with its basis in a consideration of optimal allocation or of priorities for introduction or abolition regarding limited resources. However, excluding some parts, such as medication, evaluation research regarding social security-related fields such as health care have been hardly conducted in Japan. The following difficulties may have contributed to this.

II-1. Lack of collaborative research between economics and health services research

Performing cost effectiveness analysis requires an interest in economics such as optimal allocation of resources, as well as academic knowledge of economics such as concepts of direct and indirect costs, benefits and effectiveness, marginal effects, methods of cost analysis, cost-benefit analysis, cost effectiveness analysis and cost utility analysis.

Meanwhile, professional knowledge regarding health care is also required. This includes knowledge of: medical needs for which health care technology should be applied (disease and pathology, degree of severity, common complications); teams and processes that provide health services; substitute intervention measures that could be used for comparison and contrast; costs that arise in those processes; characteristics or limitations of costs and receipt data; effectiveness of health care; utility and adverse effects; and, related confounding factors. (Hereinafter, research regarding such health care is described as "health services research" to distinguish it from medical research in the narrow sense, of which the basic science is biomedicine). In other words, cost effectiveness analysis requires knowledge of both economics and health services research. Gathering and analyzing data backed by such wide-ranging academic knowledge requires research project teams related to both fields. However, this is rare in Japan.

As a result, most medical data handled by economists in Japan has been receipt data. The receipt data includes invested medical resources (activity) and detailed data of costs. However, information on the clinical conditions of patients, which determines the validity of the application of medical activity, is mainly limited to Diagnosis Procedure Combination (DPC) information. Clinical information for evaluating treatment effectiveness only includes in-hospital death and does not include information, for example, in regard to whether blood pressure decrease or whether subsequent rehabilitation allows patients to walk. This means that the effectiveness of medical activity cannot be verified, so cost effectiveness is unknown even if the costs are calculated. Meanwhile, the main concern of medical researchers working on health service research is effectiveness and quality, and it is rare for them to collect information about costs. As a result, gathering data on both cost and effectiveness, which is essential for analysis of cost effectiveness, has been extremely rare, except in the case of medication (drugs), which must be submitted to the Ministry of Health, Labour and Welfare.

II-2. Complicated outcomes brought about by complicated intervention

There is a reason why cost effectiveness analyses have been conducted regarding medication but rarely in the case of other health care services. For example, in case of anti-hypertension drugs, short-term effectiveness can be evaluated by measuring blood pressure. Furthermore, in verification of the effectiveness of single elements such as medication, adjustments with a randomized controlled trial (RCT) using a control group administered

placebo can be conducted relatively easily in order to investigate other factors influencing results. RCTs can be easily performed if the following characteristics are present: (1) it is a relatively simple intervention, (2) there is little influence of other factors, (3) outcomes are clearly defined and measurable, and (4) short-term effectiveness is verified. Meanwhile, when evaluating a service provided by many factors, people or teams involved, the intervention itself is complicated, and it is not easy to standardize the quality of the intervention[1)]. For example, in the case of rehabilitation medicine mentioned below, many different professionals are involved, including doctors, nurses, physical therapists, occupational therapists and medical social workers. All of these people affect the quality of the health care provided. In the same team, some members of certain occupational categories are highly skilled, while other members of another category are less skilled. In addition, the presence of family caregivers can also affect treatment outcomes. Furthermore, when determining treatment effectiveness, improvement of post-discharge quality of life (QOL) is the endpoint, but it can often be difficult to measure it with a simple index.

II-3. No (available) basic database

When many factors are involved, data of these many factors must be gathered. Databases used for multifaceted analysis must have many items. Furthermore, as all patients are different, extremely large-scale data including diverse patient data is required to investigate effectiveness on such diverse patients. To examine the effectiveness further, longitudinal data that clarifies changes between before and after the intervention is also essential.

Administrative data could be used to construct a database including large-scale, multifaceted, cross-sectional and longitudinal data. For example, data on individuals requiring long-term care alone includes information on over 5 million people. However, in many cases, despite the fact that the use of data for research purposes is not subject to the Act on the Protection of Personal Information, and this is explicitly stated, using such data is often rejected by administrative officials for reasons of unintended use or to protect individual privacy.

If the secondary use of existing data is impossible, we have to make a new database by ourselves with a very limited number of evaluation targets and items. For example, as presented further below, a database is limited to only fall prevention activity participants, or only end-of-life care patients with six months of life expectancy. More limitation of subjects leads to greater subject homogeneity, but subjects are then limited to a few subjects or tens of subjects from one facility or one city, which may make data unable to withstand statistical evaluation.

To overcome it, data from multiple facilities or cities must be pooled. However, it heightens the hurdle involved due to the necessity of gaining cooperation from multiple facilities and cities that accept the significance of evaluation research. Even if such cooperation was given, the increase in target facilities and cities would lead to greater variation in service and data quality despite one wanting to verify the quality of a certain service. This

would then make even greater effort necessary in order to standardize such factors. Furthermore, when constructing independent databases, more answers to questionnaires and more items to be input into the database leads to an increase in the number of non-response or missing fields, restricting the number of variables that can be ascertained.

Thus, there are many limitations associated with even the construction of a database that can be used to evaluate the effectiveness and quality of health care. To overcome this, a significant amount of human resources, research grants and lucky encounters with cooperators is necessary. The current period for competitive research grants is approximately three to five years, and many of them are on such a small scale that they do not even cover personnel costs. Thus, it is no easy task to even begin to construct a database for this type of research.

II-4. Lack of cooperative research between the government and researchers

Data held by the government has great latent potential but it is rarely used for policy evaluation that can withstand academic criticism. There are a number of background factors for it. Firstly, the government and researchers have different priorities. Researchers, for example, prioritize adjustment by means of statistical methods that deduct the influence of factors such as age. In addition to age, data is required for many factors influencing effectiveness including gender, diagnosis and degree of severity. To examine effectiveness, an intervention group and control group need to be compared, and information from both before and after the intervention is required. Researchers believe that the mission is to determine what the "true effectiveness" of health care technology or factors are even after deducting the influence of the aforementioned factors by using advanced statistical methods.

Meanwhile, from the viewpoint of the government, providing individually-identifying information is difficult with regards to protecting personal information, and gathering information from before intervention and "evaluation for the sake of evaluation" whereby data is gathered on a control group that has not undergone the intervention is "inconceivable from the standpoint of administrative purposes." Moreover, strict research papers are difficult to understand and "cannot be used for policies." On the other hand, even if there are researchers who could help with evaluation or analysis that could be used for administration, the requirements for academic papers are not met so the work is not evaluated as being academic. Therefore, even if young research scientists get involved in this work, they would not be able to progress to higher academic posts, making it feel it a burden for them.

Such different priorities between the government and researchers means that there is rarely progression in collaborative research.

Thus, the involvement of many factors means that the construction of large-scale databases necessary for cost effectiveness analysis and collaboration between related parties is especially not proceeding in Japan compared to many other developed countries.

III. Case examples of health service research

Despite these limitations, we have engaged in evaluation research cases in the following fields: (1) health (care prevention), (2) rehabilitation medicine, (3) long-term care and (4) end-of-life care.

In evaluation of health policy, the "Threes E's" are important: (1) Efficiency or cost, (2) Effectiveness or quality, and (3) Equity or access. Adding to these, the extra "E" of (4) Empowerment for patients and citizens has led to the evaluation criteria known as the "Four E's", and good balance in these Four E's is considered to be a requirement for a good healthcare system[2].

However, in Japan, evaluation of quality and the advantages and disadvantages of access has lagged behind that in other developed countries as only cost has been focused on. Cost alone cannot be used to evaluate cost effectiveness. Firstly, methodology and frameworks need to be constructed for evaluating health care that offers a predictable effectiveness and quality. As mentioned above, this is not easy, but not impossible.

III-1. Health (care prevention)

One foundation of the reform of long-term care insurance system in 2006 was the introduction of a system emphasizing the prevention of functional decline called "care prevention". The results of this care prevention policy that was introduced in Japan, which is the most aged country in the world, is drawing attention from other countries including nations in Asia where the number of older people is increasing. The question remains as to whether care prevention is effective or not.

We investigated the situation and cause of it not leading to the desired results and will report on the necessity of "another strategy" and its basic directionality. Next, we will introduce the results of a community intervention study to verify the effectiveness of a new community-building type of prevention program, and give an overview of a cost analysis of the system.

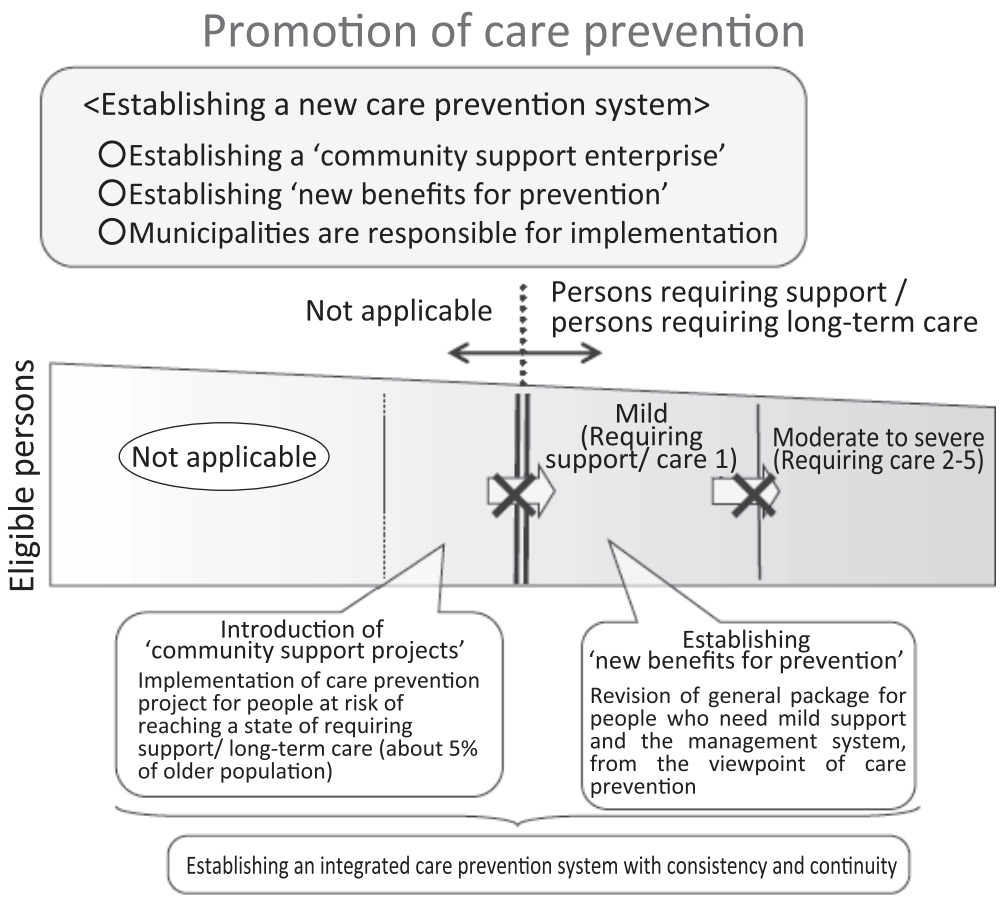
III-1-1. Care prevention activity framework

Care prevention is composed of two parts. The first is "benefit for prevention" targeting individuals who have already been certified as requiring long-term care. The second is "care prevention projects" performed as part of "community support projects" targeting individuals who have not yet been certified as being in need of long-term care (Figure 1).

The latter, "care prevention projects," is further divided into "specified elderly policies" and "general elderly policies." Specified elderly individuals are "persons likely to later require public insurance benefits for support or long-term care," and "general elderly individuals" refers to all other elderly individuals.

(Currently, the names have been changed, but here we use the expression "specified el-

Figure 1
Overview of care prevention-focused system



(Reference) Health, Labor and Welfare Ministry document—partly modified

derly individual" which was used when it was introduced.)

There are two strategies in preventive medicine: high risk strategy and population strategy. It is considered important to combine these two strategies[3)]. High risk strategy refers to specifying high risk individuals through screening and conducting intervention for such individuals. Meanwhile, population strategy targets the entire population overall without performing screening.

According to the "community comprehensive support center work manual" (December 19, 2005) by the Health and Welfare Bureau for the Elderly of the Ministry of Health, Labour and Welfare, high risk strategy is formulated for specified elderly individuals and population strategy is formulated for general elderly individuals for care prevention. In policies for specified elderly individuals, health checkup attendees undergo screening with a basic checklist to determine specified elderly individuals, and these people are recommended to participate in care prevention activities. Meanwhile, in policies for general elderly individu-

als, awareness-raising activities regarding care prevention knowledge involving the creation and allocation of pamphlets and lecture meetings are conducted in addition to the personnel training of volunteers engaging in care prevention, and the cultivation and support of community activity organizations.

III-1-2. Have newprevention provisions helped to prevent functional decline?

During the diet deliberation conducted during the introduction of care prevention, questions were raised regarding effectiveness, which were then verified 3 years later[4]. According to this verification, the newprevention provision for individuals who have been certified as requiring long-term care "was found to lead to a statistically significant reduction in the number of persons exhibiting an aggravated condition before and after the policy introduction, confirming the care prevention effectiveness." However, there were many limitations in this analysis method including the fact that the control group for this analysis was the same group before the introduction of the policy (historical control), and the report stated that "further investigation needs to be continued."

III-1-3. Problems related to policies for specified elderly individuals

Next, the policies for specified elderly individuals "reduced the onset rate of aggravation for requiring long-term care for specified elderly individuals before and after the introduction of the policy but statistically significant care prevention effectiveness could not be calculated[4]."

As shown below, a reason for the policies for specified elderly individuals not being effective is the fact that the three conditions necessary for effective preventive policies were not fulfilled. Firstly, the fact that high risk individuals who should be targeted for care prevention activities could not be screened. Secondly, the fact that these people did not participate in the care prevention program. Thirdly, the problem of the effectiveness of the care prevention program in which such individuals would participate.

(1) Screening problems

Frail elderly individuals who have risk factors making them prone to requiring long-term care were estimated to account for approximately 5% of the elderly population at the time. However, they actually only accounted for 0.14%[5]. One reason for this was the fact that screening was conducted for healthy participants.

Our investigation revealed that compared to individuals who do not undergo health checkups, those who do undergo such checkups are actually healthier[6,7]. Thus, healthy elderly individuals were unintentionally screened, making it impossible to understand "specified elderly individuals" according to the expected standards.

Another issue was the high cost involved in these policies. According to materials from the June 15, 2010 conference of national officials of Community-Comprehensive Support Centers, the costs required to determine specified elderly individuals with health checkups was approximately 50% of long-term care prevention activities (national expenditure base)

budget (17.6 billion yen). If cost-covered insurance was included, the cost rose to 34 billion yen (66.7%) of the 50.9 billion budget, meaning that more money was spent on ascertaining the target population than on providing care prevention programs.

(2) Problems related to participation refusal

There were also problems related to the second condition, "participation in care prevention programs by screened specified elderly individuals." A significant number of people refused despite being recommended to participate in a care prevention program.

According to the Community-Comprehensive Support Center/Long-Term Care Officers' Meeting materials (held on March 14, 2007), out of 112,124 specified elderly individuals (FY2006), over half did not participate in the program, including 27,025 (24.1%) who did not participate of their own accord.

In the year after its introduction, in consideration of the fact that activities were not proceeding as expected, the Ministry of Health, Labour and Welfare released notifications of relaxing standards, making it possible to substitute self-administered questionnaires rather than screening health checkups, and recommending the use of common names, which are easy to use in each municipality, because of complaints from specified elderly individuals about using the name "specified elderly individuals."

(3) Effectiveness of care prevention programs

Let us now consider the third condition, "Effectiveness of care prevention programs." Looking at individual studies, although some reported that programs were effective, others reported that there were no effects. Regarding fall prevention, on which the studies have been accumulated, systematic reviews comprehensively investigating previous studies or meta-analysis[8)] involving the gathering and re-analysis of data have been reported overseas. The results of the review of 19 controlled comparative studies (RCTs and pseudo-RCTs) indicated that the ratio of falls in the intervention group (relative risk ratio), when the control group was 1, was reduced to 0.91 to 0.90. However, no statistically significant difference was observed.

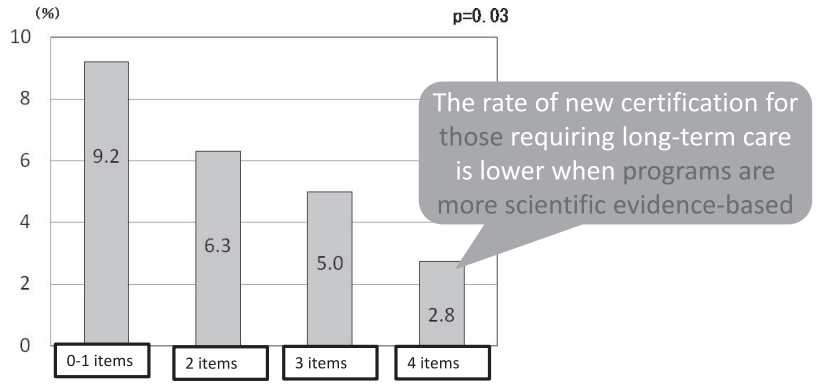
Let us now consider fall prevention programs in Japan. Four conditions were set based on evidence which had accumulated overseas, and the rates for requirement of long-term care one year later were compared between municipalities according to the extent of fulfilling the above conditions[9)].

The results indicated that of the 71 target municipalities (insurers), there were 8 municipalities with programs that fulfilled all four conditions. In these municipalities, the rate of certification for requiring long-term care was 2.8%. However, in the municipalities for which only 0 to 1 of the conditions were fulfilled, this rate nearly tripled to 9.2%. Thus, fall prevention programs were considered effective not in all municipalities, but in some municipalities where high quality programs were provided.

These results above suggest that policies for specified elderly individuals have many problems: screening methods and costs, small number of individuals participating in activi-

Figure 2
Evidence-based fall prevention programs help decrease the proportion of new certifications

(Hayashi T, Kondo K, 2013)



・Survey of 71 municipalities throughout Japan: We calculated rates for certification of those requiring long-term care according to the number of items that meet the following "four characteristics."
・"Four characteristics" for which there is evidence proving they are effective in fall prevention programs:
(1) Staff specialized in exercise guidance are involved in fall prevention intervention, such as physiotherapists and occupational therapists.
(2) The exercise program includes not only muscle strength enhancing, but also balance training.
(3) The total time of intervention involving instruction from an instructor and home exercise is at least 2 hours per week.
(4) A follow-up program after conclusion of the fall prevention workshop is provided.

ties despite being screened, program effectiveness and the scale (coverage ratio) of subjects for the care prevention program[10].

III-1-4. Problematic points related to policies for general elderly individuals-missing program evaluation

Let us now consider an "alternate care prevention strategy," comprising policies for general elderly individuals (=population strategies).

Most policies for general elderly individuals involve awareness-raising activities through lecture meetings or similar, and few have organized care prevention programs for which the intervention theory and methodology has been indicated. Although the system in place does include activities to evaluate the effectiveness of such programs, most of the results of such assessment have not been reported. Even investigations of the number of participants, which is a basic form of evaluation, have only indicated rough estimates, and the actual number of participants had not even been investigated until 2014. Furthermore, as lists of participants are rarely made, participants cannot be followed up in order to evaluate effectiveness using the currently-available information.

III-1-5. Novel pilot program=choices for care prevention

Thus, evaluation and verification is currently insufficient for the aforementioned three policies, "new benefits for prevention," "policies for specified elderly individuals" and "policies for general elderly individuals," and these have not been refined into effective systems. As a result, the effectiveness cannot be evaluated, let alone the cost effectiveness of these activities. Despite the limited information that is available, it is highly likely that both effec-

tive and ineffective policies are in place. If an attempt is made to restrict costs without a system for evaluating effectiveness, the scale of all current policies will have to be reduced equally. This would mean that even effective activities would be discontinued. However, what we need to restrict is ineffective activities only.

The next issue that we face is enhancing evaluation. If things progress randomly without it, "exorbitant waste" on a scale of tens of billions of yen such as that observed for target screening costs is likely to continually occur. Considering our limited social insurance resources brought about as the declining birth rate and aging society becomes more serious, it is unacceptable to simply maintain the current status.

There are three future options for Japanese society. The first is continuing to improve the problems associated with policies for specified elderly individuals while expanding such activities by 10-fold or more. The second choice is developing and evaluating programs that equate to an "alternate care prevention strategy," comprising policies for general elderly individuals (=population strategies), and also developing effective strategies and methods. The third option is to quickly come to the conclusion that "efforts to prevent functional decline are essentially attempting to achieve 'eternal youth and longevity,' making these efforts unrealistic and fruitless," and abandon such efforts.

In the field of program evaluation, the term "type III error (mistake)" is used[11]. These errors, including "type I errors" that mistakenly assume effectiveness (despite lack of statistical significance), "type II errors" that overlook effectiveness (despite statistical significance), and "type III errors", which refers to discontinuing activities in the early stages despite the possibility of effectiveness if improvements are made. If an evaluation were conducted, we could determine whether certain care prevention programs were highly effective or non-apparently effective. Type III error refers to completely stopping activities due to effectiveness not being apparent overall. The discontinuation of care prevention activities that are still under discussion is likely to fall under this type of error.

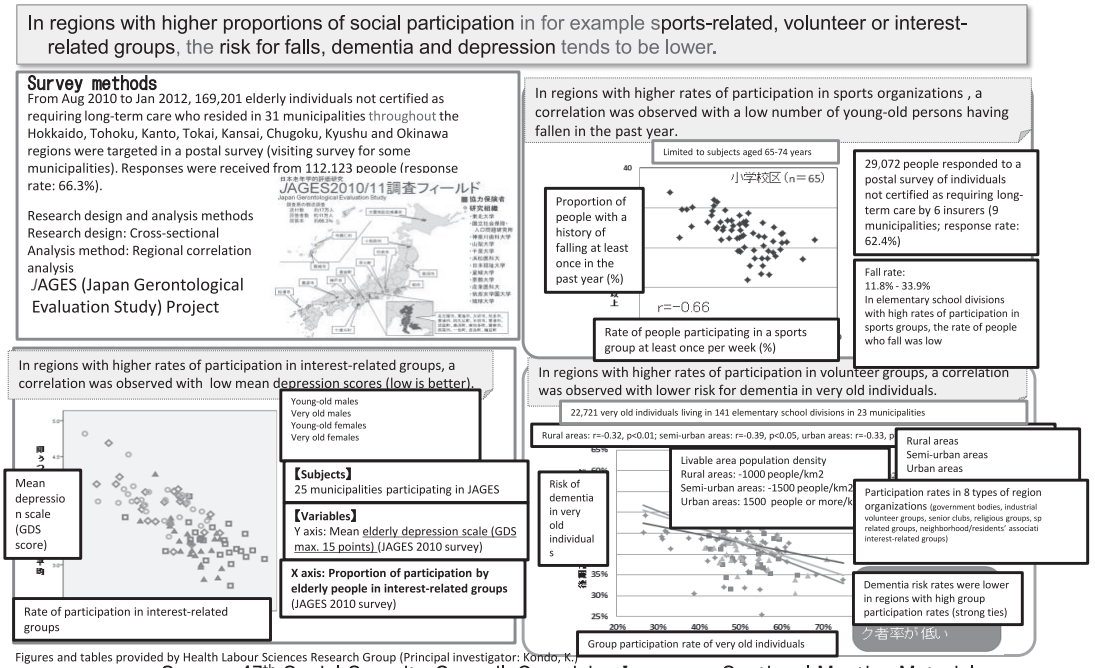
III-1-6. Social capital and Taketoyo Project

With the aim of obtaining basic evidence for policies for elderly care, we have been involved in the Aichi Gerontological Evaluation Study (AGES) project since 1999[12;13]. As part of this, we have verified the relationship between community social capital and the health of residents in the community. Social capital is the richness of "social cohesion between residents in the community" that Robert Putnam, a political scientist, has defined as "the characteristics of the social structures of trust, norms and networks that can improve social efficiency by vitalizing coordinated activities[14]."

Various research results have indicated that social capital is related to health. Ecological correlation analysis with municipalities or elementary schools as aggregate units (Figure 3) found that communities with many older people who participated in sports or hobby groups had fewer individuals with a history of falling, low depression scores (good mental health) and a low proportion of individuals at risk for dementia[15].

Analyses using self-rated health[16], which has predictive power of mortality, as a health

Figure 3
The Relationship between Social Participation and Care Prevention Effects



Source: 47th Social Security Council, Caregiving Insurance Sectional Meeting Materials

index have also indicated the same results. The correlation coefficients of social capital and the proportion of people who answered that their self-rated health was not good were calculated for each of the 25 old villages. This led to a result of -0.50 (p=0.02), revealing a relationship whereby richer social capital indicated fewer individuals with "poor health levels." However, ecological studies such as this of simple community level correlations between values aggregated for community units can overestimate the relationship.

Therefore, we performed multi-level analyses, which can simultaneously analyze correlations between individual level and community level variables[17;18]. These results indicated that in a community with 1% higher values for general sense of trust, the rate of people who responded that their self-rated health was poor, was reduced by 4%. Although this figure appears very low, the denominator was all elderly individuals living in that region, and the figure accounts for 40 out of every 1000 individuals, resulting in showing that this effect is not low at all.

The health indices used also include mental health indices[19] and more objective index of the number of remaining teeth[20]. It follows that the "introduction of the concept of social capital" was included in the 2014 Ministry of Health, Labour and Welfare white paper.

(1) A population strategy in the social capital theory – Taketoyo Project

If population strategies targeting general elderly individuals could enrich social capital

by encouraging them to participate in social activities, it may likely improve psychological status and social lifestyles overall, and increase the health level of community-residing elderly individuals. Though many programs targeting them have been introduced nationwide, few programs has been evaluated in regard to effectiveness on health.

Therefore, we have been evaluating effectiveness of care prevention programs based on social capital theory. In Taketoyo town, Aichi Prefecture, there is a community interventional trial called the Taketoyo Project.[21;22;23;24)]

Three sites called "salons" at which older people enjoy social activities programs were established in 2007, and by 2014, the salons had increased to 11 sites. Interim evaluation has indicated that these programs have promoted social support and participation, which are compositional elements of social capital. We also found that healthier people are likely to participate in the program. Therefore, even when a participant group exhibits the anticipated improvement in health indices, it might be the effects of better health at baseline. To take account of these effects and unobservable differences between individuals who participated in the program and those who did not, we used the instrumental variable method, in which the distance from each resident's home to the nearest salon was used as an instrumental variable. As a result, a favorable outcome of a 2.5-fold improvement in self-rated health was observed in individuals who participated in the program. Then, the Taketoyo Project was introduced on the homepage of the Ministry of Health, Labour and Welfare (http://www.mhlw.go.jp/bunya/shakaihosho/seminar/02_88.html, http://www.mhlw.go.jp/bunya:shakaihosho/seminar/dl/02_88-24.pdf).

As it mainly targeted healthy individuals, no statistically significant differences were de-

Figure 4
Case example of community-building utilizing long-term care prevention activities

http://www.mhlw.go.jp/topics/2009/05/tp0501-1.html

Taketoyo City, Aichi Prefecture — Long-term Care Prevention Manual Revised Version, p.4

For example, in Taketoyo City, Aichi Prefecture, Japan, regional salons for learning and exchange that could be easily accessed by elderly individuals at any time were introduced as a primary prevention activity in each district and they are run autonomously by residents. When preparing to establish these regional salons, municipalities called for resident volunteers and gave explanations of planning to them. Subsequently, the volunteers came up with issues for the region, as well as discussing the type of salon necessary for each region, leading to the opening of the salons. The diverse plans drawn up as part of these regionally-autonomous activities were created based on each region's needs, and results have been seen such as a great increase in the number of individuals participating in long-term caregiving prevention activities compared to the previous year.

日本老年学的評価研究
Japan Gerontological Evaluation Study

Taketoyo Project (Regional intervention research to prevent long-term care)

http://square.umin.ac.jp/ages/taketoyo.html

tected in rates of certification for requiring long-term care insurance in short-term follow-up. However, 5-year follow-up indicated significant suppression of the rate of certification for requiring long-term care insurance. Analysis with instrumental variables etc. taking account of background factors such as the fact that healthier people can more easily attend such salons also indicated the same results. The rate of certification for requiring long-term care insurance for participants was 7.7%, which was roughly half of that of non-participants (14%)[25].

(2) Taketoyo Project cost analysis

The costs used for the salons program in Taketoyo Town were analyzed in 2010. Costs were calculated to include fees paid by the town, public long-term care insurance such as administrative outsourcing expenses, and human resource costs for personnel at Health Divisions, Welfare Divisions, community comprehensive support centers and the social welfare council. Costs for full-time employees were calculated based on the Survey on Wages of Local Government Employees, with the part-time hourly wage calculated as 800 yen. Opportunity costs for volunteers were not included as costs. These were calculated based on government materials and similar[26].

The total cost for one year was 6,327,077 yen. The breakdown of (1) personnel costs and (2) actual total annual expenditure was 3,197,460 yen and 3,129,617 yen, respectively (Table 1).

Actual total annual expenditure included venue operating costs provided by the town, gratuities for instructors, expendable supply costs and casual employment wages (Table 2). Venue operating costs totaled 2,724,590 yen for eight sites in 2010. This figure was approximately 4% lower than that of the previous year despite the number of salons increasing by one (Table 3).

In 2010, there were eight salons, 727 actual participants (counting participants who went to multiple salons once, excluding guest participants and those who only participated in the opening ceremony), 177 registered volunteers, 6,441 total participants (including volunteers, opening ceremony and free participants and invited guests), and 109 open days.

Thus, the annual cost was 8,703 yen per actual participant and 982 yen per total participant (Table 1). Costs per salon were calculated by dividing total cost by the eight salons, resulting in 790,885 yen. The cost per salon open day was calculated by dividing total cost by salon open days, resulting in 58,047 yen.

The nationwide annual mean amount of benefits for each individual who certified as requiring long-term care insurance is approximately 1.92 million yen. Although it depends on hypothetical conditions, if at least 4 persons per year could be inhibited from being certified as requiring long-term care insurance, this program would offer excellent cost effectiveness. For 5 years, approximately 46 individuals (7.7%[25] of the approximately 600 participants in 2009) and 9 individuals per year could be anticipated to not be certificated as requiring long-term care as a result of this program, it appears to offer excellent cost effectiveness.

Table1
Total costs and breakdown for Taketoyo Project

Year of 2010

Cost item	Breakdown/unit	Cost (yen)
Total cost		6,327,077
①Personnel		3,197,460
②Total actual expenses	Venue operating costs, remuneration for instructors, irregular employment wages, etc.	3,129,617
	Per actual participant (727 people)	8,703
	Per total participants (6,441 people)	982
	Per salon location (8 locations)	790,885
	Per salon day open (109 days)	58,047

(2012 Outline of actual survey for long term care benefits) Annual long term care benefits per person is 1.92 million Yen＝It pays for itself when reducing the number of people who newly certified as requiring long-term care at a rate of 4 people a year vs an average of 600 people in 5 years × 7.7%=46 people

Table 2
Breakdown of total annual expenditure

Cost item	Cost of venue management, etc.	Gratuity for lecturers	Other consumables	Postage fees	Wages for casual employment	Total
Annual expenditure	2,724,590	190,000	29,987	5,840	179,200	3,129,617

Guidelines for each venue, including expenses for commission, consumables, rental charge for venue-including gratuities for lecturers included in expenses for venue commission
Reference:Taketoyo City documents

Table 3
Summary of performance of the program

Each venue's performance	Number of actual participants	Number of registered volunteers	Total number of participants	Number of open days	Registration fee (income)	Cost of venue management, etc.	Increase and decrease of allowance	Remarks
2009 total	637	155	5,223	91	409,500	2,827,780		7 venues
2010 total	727	177	6,441	109	519,200	2,724,590	△ 103,190	8 venues
Comparison with the previous year	114%	114%	123%	119%	116%	96%		

Reference:Kondo, Takeda & Suzuki (2014)-the author developed from the briefing paper, etc.

III-2. Empirical research elucidating more effective rehabilitation

Next, we will consider more effective stroke rehabilitation programs. Strokes were targeted as they will likely become more important in health and long-term care for our super-aged society due to being the greatest cause of becoming bedridden, and being the target disorder for planning medical coordination in the 2008 Revised Medical Care Law.

Here, we will discuss our empirical research, which has been cited in the stroke treatment guidelines jointly published by five academic societies[27]. This shows the effects of increasing early rehabilitation and the amount of training, involvement with rehabilitation specialists, regularly holding case conferences (care management), and case registration databases used to verify these effects. These research projects were health services research rather than being medical research, or research on technological systems rather than individual technology. This research should be enhanced in Japan in the future.

III-2-1. Early rehabilitation

The guidelines[27] recommend and advise starting rehabilitation from an early period. This is because disuse of physical and mental functions reduces overall functions including muscle and cardiopulmonary functioning. Disuse syndrome is even more dangerous than it is generally considered. However, disuse syndrome may occur or recovery speed may be slowed even for patients who engage in rehabilitation due to an insufficient amount of rehabilitation training[28]. Furthermore, one can also confirm the relationship whereby a greater amount of walking leads to faster recovery time (muscle cross-sectional area variation)[29].

Many observational studies have also found that even in consideration of age, degree of independence in activities of daily living (ADL) and amount of daily training, shorter days from onset to the first day of rehabilitation leads to higher rates of ADL improvement and daily improvement[30;31;32;33]. However, these results alone could probably be explained by the fact that rehabilitation is started earlier after onset when recovery is faster. Therefore, some studies utilizing data about which calendar day/holiday patients were hospitalized have been conducted. For example, patients who are hospitalized on the weekend, during the Golden Week holiday or the New Year's holiday, when there are few staff available due to it being a holiday, tend to start rehabilitation later or undergo less training. When this fact is utilized to verify the effects of early rehabilitation, it has been found that patients hospitalized on weekends with low early rehabilitation implementation rates exhibited lower ADL improvement rates[34]. Furthermore, patients who underwent little rehabilitation during the early stage of hospitalization not only had poor outcomes as evaluated by the modified Rankin Scale (mRS), but also a higher mortality rate[35].

Thus, despite this evidence being accumulated, the date of hospitalization in recovery rehabilitation wards currently exceeds 30 days after onset[36], indicating that the care currently being offered is far from "early rehabilitation." A study of acute cerebral infarction patients in Diagnosis Procedure Combination (DPC) participating hospitals (n=38,947 pa-

tients) found that the rehabilitation implementation rate was 47.8%, indicating that approximately half of patients do not undergo rehabilitation[30].

It is anticipated that policies to promote and achieve early rehabilitation such as making it mandatory to place rehabilitation staff at acute hospitals will reduce the number of individuals who require long-term care.

III-2-2. Will increasing the amount of training improve outcomes?

One means of enhancing rehabilitation care is increasing the amount of training. Increasing the amount of training has empirically been said to increase effects. However, as more evidence has been gathered, this issue has also been discussed at educational lectures held by the Japanese Association of Rehabilitation Medicine[37].

(1) Three background factors for the amount of training being focused on[37]

The first factor is "rediscovering" disuse syndrome. The second factor is questions regarding the validity of medical fees. In Japan, the maximum daily limit for medical fees for training is 2 hours. In contrast, guidelines in the United States indicate that to be eligible to participate in a full-scale rehabilitation program, the patient must be able to withstand at least 3 hours of physical exercise and training[38]. Thus, the upper limit in Japan does not even reach the lower limit in the United States. The third factor is the demonstration of "greater recovery brought about by increased training" as evidence-based medicine (EBM)[38;39;40].

(2) Enhancing rehabilitation effects by increasing the amount of training

Let us give a simple introduction of the evidence. Firstly, progressive resistance training (PRT), in which the amount of resistance is gradually increased, improves muscle strength and walking speed[41]. Increasing the amount of hemiplegic arm training of stroke patients by constraint induced movement therapy which forces them to use their hemiplegic arm is effective for functional recovery in the hemiplegic arm of both acute phase patients[42] and chronic phase patients[43]. Effectiveness for both acute and chronic patients has been demonstrated in randomized controlled trials (RCTs), which are considered to offer the highest quality evidence.

The effectiveness of increasing the intensity of training for stroke patients has already been investigated with many RCTs. Meta-analysis of the results has demonstrated that good outcomes for nervous and muscular function[44] as well as for ADL were achieved after 3 months in a group that underwent intensive rehabilitation.

In Japan, a program involving high density (large amount) training on consecutive days including weekends and holidays significantly improved ADL during hospitalization compared to before the program had been introduced[45;46]. Rehabilitation has also been demonstrated to have psychosocial effects such as for depression[47;48].

(3) Empirical research of the effectiveness of the amount of training and rehabilitation

There is clearly a relationship between the amount of training and effectiveness of rehabilitation (within a certain range). Thus, is the amount of training in Japan truly sufficient? In 2003, the Japanese Association of Rehabilitation Medicine conducted a "study on effectiveness of rehabilitation and medical treatment fees"[32]. With the data provided by this survey, we independently conducted a secondary analysis of the relationship between daily amount of training and ADL improvement rate[49;50].

Subjects for analysis comprised 1,446 patients (mean age: 66.0 years). Mean Barthel Index (BI; ADL scale with a maximum score of 100 points, with 0 indicating total dependent) score of 50.4 points upon hospitalization and 72.4 points upon discharge, with a home discharge rate of 70%. The mean number of days after onset for the initial day of rehabilitation was 55.6 days and the mean daily amount of training was 56 minutes when calculated as "PT + OT + training exceeding the maximum allowed for medical fees + amount of voluntary/additional self-exercise by the patient or family."

Multiple regression analysis was performed to clarify whether the ADL improvement rate (daily) increases with a greater amount of training even when controlled for confounding factors. The results suggested that the more the amount of training increased, the greater the ADL improvement rate.

III-2-3. Involvement of specialists and regular conferences

We also investigated whether "the involvement of rehabilitation specialists (board-certificated physiatrist of the Japanese Association of Rehabilitation Medicine) enhances outcomes such as ADL improvement rates or home discharge rates"[51]. The results of multivariate analysis (multiple regression analysis/logistic regression analysis) indicated that, for example, if cerebral infarction patient groups are limitedly analyzed, groups for which there was specialist involvement exhibited significantly good outcomes for daily ADL improvement rates and home discharge rates.

Another hypothesis verification is "whether hospital groups at which case conferences are held regularly have better outcomes"[52]. Matched pair analysis was also conducted in addition to the aforementioned multivariate analysis. Pairs with similar conditions (matches) for a maximum of eight confounding factors (including age, primary disease, hospital days after onset until first day of rehabilitation and Barthel Index upon hospitalization, etc.) were created and compared in the regular conference group and non-regular conference group. Comparison under the eight conditions with altered matching conditions indicated that for all conditions, outcomes were better for the regular conference group. Statistically, for example, for the degree of ADL improvement (during hospitalization) improved significantly with six of the eight conditions and daily ADL improvement rate improved with four of the conditions in the regular conference group.

III-2-4. Database development

The aforementioned analyses demonstrating that early rehabilitation, training amount,

specialist and regular conferences bring about good outcomes are only possible if patient data at multiple facilities where different programs are conducted is pooled. Large-scale data is required to verify more multilateral hypotheses and to constantly monitor care quality.

In countries such as the United States, large-scale databases (or data banks) have been developed since the 1980's: for example, the Uniform Data System for Medical Rehabilitation (UDSMR)[53], for which 13 million rehabilitation patients have been registered, and the NINDS Stroke Data Bank.

In Japan from the end of the 1990's, various large-scale data banks of case data have been jointly developed by multiple facilities led by academic societies and similar. These data banks contain a large number of cases, including acute phase stroke[54], unruptured cerebral aneurysms[55], external head trauma[56], ICU-acquired infections[57] and influenza[58] in addition to rheumatism (approximately 6000 cases at 25 facilities)[59], acute myocardial infarction (1183 cases at 20 facilities)[60] and perinatal period (16,299 cases at 40 facilities)[61].

We formed a research team with the aid of specialists from the Japanese Association of Rehabilitation Medicine and engaged in the development of a multicenter data bank for rehabilitation patients.

It required approximately 5 years to develop this data bank. With a 2005-06 Health Labour Sciences Research Grant, we first developed a data bank for stroke rehabilitation patients[62;63]. We recruited participating facilities and provided a generally available database that could be downloaded from the Internet at http://rehadb.umin.jp/[64]. In 2012, four rehabilitation-related groups established the Japanese Association for Rehabilitation Database (JARD), and there are currently over 30,000 registered patients from approximately 80 hospitals. Evaluation conducted by comparing expected functional independence measure (FIM)[65] predicted from status upon hospitalization with actual FIM score upon discharge indicated reproducibility in the finding that at facilities with specialists or high training amounts, there were more patients exhibiting improvement than predictive value, and a significant difference for treatment outcomes was noted between hospitals (Figure 5)[33].

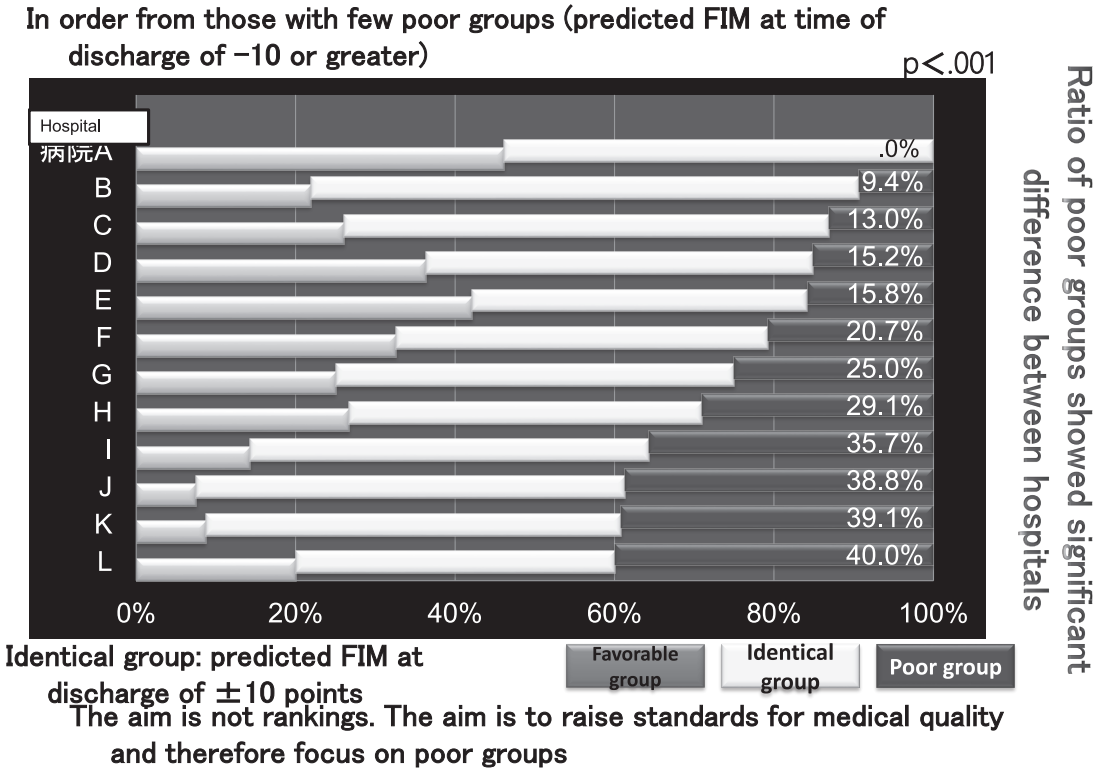
III-3. Achievements in evaluation of care quality in long-term care services

Next, we will consider long-term care services. As the amount of long-term care increased, evaluation of care quality has become required. Here, we will introduce and discuss an evaluation study of care quality in both visiting care (day care/services) and facilities care (special elderly nursing homes).

III-3-1. Changes in use frequency of visiting care services before and after introducing new preventive services, and subsequent outcomes

With the revision of the Long-Term Care Insurance Act in 2006, some people limited their use due to, for example, converting to bundled payment for visiting care services. Although some people in the field voiced concerns that more people might become "homebound", actually increasing the number of elderly with exacerbated long-term care require-

Figure 5
Treatment outcome benchmarks at recovery-phase rehabilitation hospital wards



ment, a committee of the Ministry of Health, Labour and Welfare reported that compared to before the introduction of new preventive services, groups of patients with exacerbated long-term care requirement decreased and effectiveness of care prevention was noted. However, critically reviewing the method of analysis, the possibility of "a reverse relationship of cause and effect" could not be denied. Therefore, we used a method of analysis that we believed to be more appropriate for such and conducted a follow-up study to verify the effectiveness of the new care prevention programs.

From approximately 288,532 person-months' worth of long-term care insurance benefit payment receipt data over 2 years from seven insurers (municipalities) between April 2005 and March 2007 (approximately 12,000 people × number of months of observation of each person), we extracted data from 432 people who had utilized the preventive day care/services since before the introduction of new preventive programs and switched from requiring support to requiring support level 1 during the observation period. These subjects were divided into three groups according to whether the number of times services were used decreased, did not change, or increased after the introduction of new preventive services, and elements such as rates of exacerbation of the required care level were compared.

The results showed that after switching to requiring support level 1, the number of times

of preventive services used decreased to 42% (180/432) and the mean number of usages also significantly decreased. Looking at preventive services overall, only a small percentage of subjects in the decreased usage group exhibited exacerbation, however, unlike in the report by the committee of Ministry of Health, Labour and Welfare, this was not a statistically significant difference.

In contrast to the results of the committee, the exacerbation rate (15.2%) and exacerbation onset rate (0.025) of the decreased usage group was 4.5- to 5-fold that of the increased/maintained usage group, which exhibited results of 3.4% and 0.005 respectively, demonstrating a statistically significant difference.

Thus, it appears that decreased usage of day care/services may have lowered the amount of physical activity, leading to progression of disuse and an increase in the number of individuals who exhibited exacerbation. It is calculated that the decreased usage of day care/services has led to exacerbation in approximately 20,000 individuals throughout Japan. It is roughly estimated that an increased provision of care by the reform could be amounting to approximately 13 billion yen. When using a different method of analysis from that utilized by the committee, different results were obtained. We anticipate stricter program evaluation such as reproducible verification made possible with the utilization of big data owned by insurers and multilateral program evaluation research.

III-3-2. Evaluation of care quality in special elderly nursing homes

We will also present a care quality evaluation that we have implemented for special elderly nursing homes, which are one of the most important facility care services and have the longest history in care quality evaluation.

(1) Achievements of care quality evaluation[66)]

Activities for care quality evaluations for special elderly nursing homes in Japan began in the late 1980's. The stages of development can be broadly divided into three parts: (1) the beginnings of service evaluation—late 1980's, (2) systemization of service evaluation—1990's, and (3) the diversification of service evaluation—2000's and onward.

Most of the previous evaluation methods were check sheet-based using evaluation items. The evaluation items focus more on structure/progress elements than outcomes, therefore, the achievement of organizational level activities such as "the presence or absence of implementation of training" has been evaluated rather than user level elements such as "the presence or absence of pressure ulcers."

Outside of Japan, the evaluation of care quality using quantitative indices has been progressing[67;68;69)]. For example, in the United States, the objective and quantitative indices of quality measures (long-term care: 14 items, short-term care: 5 items) and quality indicators (24 items) were developed from the assessment tool, Minimum Data Set (MDS), and inter-facility comparison of care quality in nursing homes has been conducted. The results have been made widely available to the general public, and steady improvements in quality have also been observed since the introduction (American Health Care Association 2011).

Looking at the state of activities being conducted overseas, it appears that also in Japan, 1) enhanced outcome evaluation, 2) the development of objective and quantitative indices and 3) the operation as a system are required.

(2) Care quality evaluation at special elderly nursing homes—attempts to use data of certification of requiring long-term care

In the revision of the Long-Term Care Insurance Act, care quality evaluation has been the focus of attention. Therefore, we aimed to develop care quality evaluation indices that could be used to make comparisons between special elderly nursing homes. In consideration of costs for data collection, we created indices from existing official data, which is made up of individual certifications of requiring care level and of eligibility for benefits from long-term care insurance.

Firstly, we developed 11 indices that could be created from the data of the certification and verified the reliability and validity of these indices. Next, we verified criterion-related validity between evaluation using these indices and the evaluation by three visiting surveyors who were blinded to evaluation results of indices.

A. Care quality index development

The certification data of requiring long-term care provided by 40 insurers was combined with benefits provision data, by which used facility is detected. We analyzed 4,923 people (91 facilities) who were using special elderly nursing homes at the two time points of June 2007 and November 2008. Care quality was evaluated using 11 indices according to the rate of maintenance or improvement in required care level one year later. We investigated the extent of differences between facilities and the correlations between indices[70].

The results indicated that the maintenance or improvement rates for the more comprehensive indices of required care level, degree of being bedridden and degree of dementia independence were all approximately 76%. Difference between maximum values (facilities with high maintenance or improvement rates) and minimum values (facilities with many individuals exhibiting exacerbation) ranging from approximately 21 points (the rate of walking) to approximately 74 points (the rate of contracture) were observed. Many indices showed a difference of approximately 40 points. Excluding the index of proportion of those without pressure ulcers in both of the two time points, the other 10 indices were highly correlated with each other and showed significant positive correlations.

Of the 11 indices created, the maintenance or improvement rate of required care level could be used as a comprehensive index, and the rate of meal intake and urination/excretion could be used as specific indices.

The following items need to be discussed. When creating indices, who should be included into the subject of the evaluation and to what extent should users' attributes be adjusted? When utilizing existing data for care quality evaluation, there is the advantage of no additional efforts needed in gathering data for evaluation. However, it requires the development of simplified data handling methods for accumulating and creating longitudinal data, and

furthermore, to construct a management system for the effective utilization of evaluation results within nursing homes.

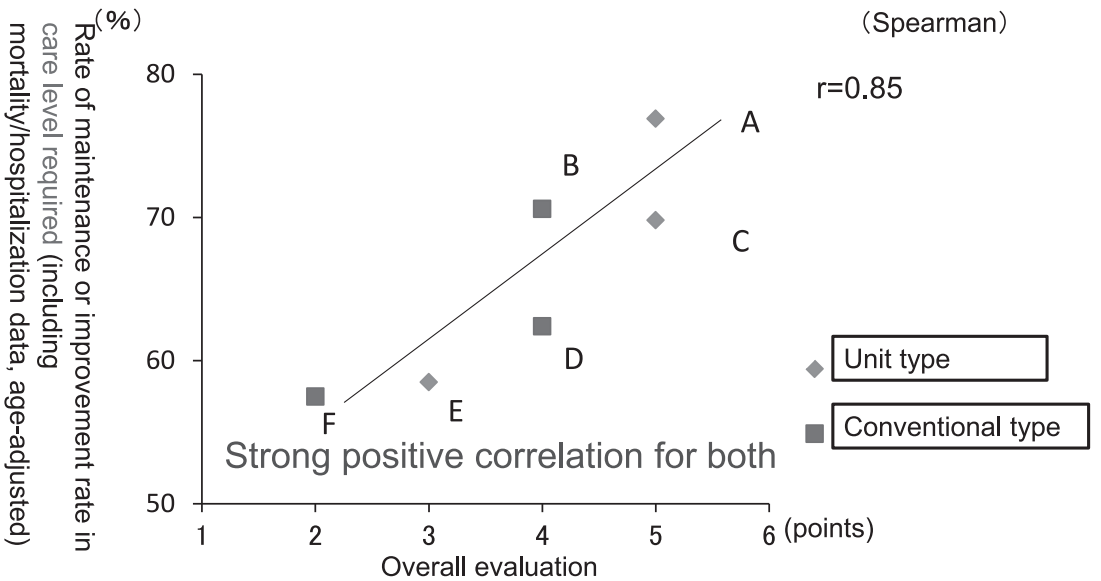
B. Comparison of care quality with blind visit surveys

To investigate whether the index, maintenance or improvement rate of required care level could actually be used to evaluate the care quality of special elderly nursing homes, we investigated criterion-related validity. Evaluation results of 4-hour visit surveys by three individuals who have a detailed knowledge of elderly care were used as a reference for the criterion-related validity test, in which the correlations between these results and the rates of required care level indices created from the existing official data are examined [71]. Six special elderly nursing homes for which insurers provided data related to long-term care insurance were targeted for the analysis. The indices of the rate of required care level for a total of 33 facilities was calculated and these facilities were divided into high, medium and low score groups. We then selected three facilities each from the high, medium and low score groups. The three visiting surveyors, who were blinded to (did not know) the indices score, evaluated care quality through 4 hours of observation and interviews at the target facilities. Evaluation by the surveyors covered care-related items such as meals, bathing, excretion, mobility and functional training care, along with other items such as the atmosphere between employees and users (elderly), and overall evaluation. These were assessed according to six ranks, which were very poor, poor, somewhat poor, somewhat good, good and very good. We then examined the correlation between the results of "overall evaluation" of care quality based on consensus among the surveyors and the rate of maintenance or improvement in requiring care level.

We found a moderate to high correlation coefficient ranging from 0.53 to 0.85 between care quality evaluation by surveyors and the indices of the rate of requiring care level (Figure 6).

As data related to certification for long-term care insurance has already been standardized and digitized, it is theoretically possible to calculate the indices, the rate of maintenance or improvement in requiring care level at all special elderly nursing homes throughout Japan. However, although we actually ask many insurers to provide data for long-term care insurance, we were not able to obtain the data for reasons such as the protection of personal information. New measures such as the utilization of data from general databases of long-term care insurance constructed by the government are necessary. If "big data" could be obtained, it would be simple to create multiple indices from such data, however, it would be necessary to confirm the validity of whether it reflected care quality. Based on the present results, it appears that when care quality evaluation by visiting surveyors is used as a reference there is high criterion-related validity for the index in regard to the rate of maintenance or improvement in requiring care level. As it is important that care quality evaluation indices are accepted by industrial bodies, we believe that further joint research involving facilities and validity test using data on more cases at more facilities is necessary.

Figure 6
Correlation between "overall evaluation" by visiting surveyors and "rate of maintenance or improvement in care-required level"



III-4. End-of-life care quality

Finally, we will present end-of-life care quality evaluation.

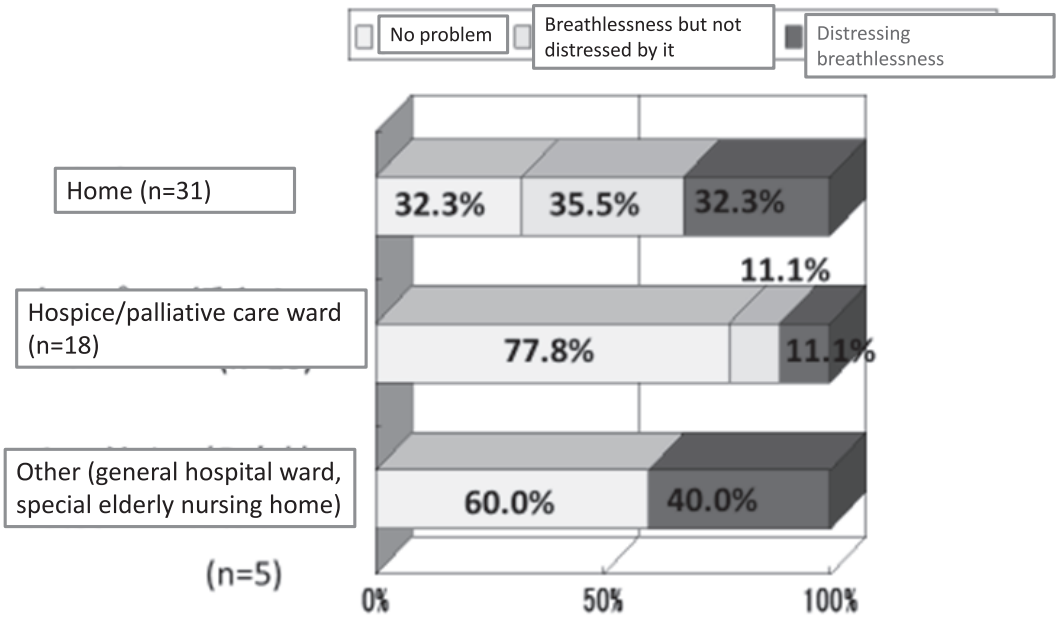
As the aging of society further progresses, the number of people who pass away is also increasing. In 2008, approximately 1.15 million people died in Japan. 30 years from then it is expected that this figure will increase by 1.5 times to approximately 1.7 million people per year[10;72].

It is difficult but not impossible to evaluate end-of-life care quality. For example, the Minimum Data Set for palliative care (MDS-PC) is utilized to demonstrate that even amongst cancer patients, the proportion of patients who experience breathlessness and distress during exertion differs according to the place of care[73]. If this type of evaluation were made mandatory and databases were constructed, comparisons of different methods of care and inter-facility comparisons would also be possible.

IV. Possibilities and Future Challenges

We have reviewed our achievements in evaluation research of care quality from (1) health (care prevention) to (2) rehabilitation medicine, (3) long-term care and (4) end-of-life care. Based on the above, we will summarize the possibilities and future challenges in these fields.

Figure 7
Breathlessness upon exertion – Cancer (n=54)



IV-1. Possibilities for large-scale databases

Many of the studies that we have introduced were conducted using databases constructed with the participation of multiple municipalities or facilities. These databases had several advantages which cannot obtained from databases created from single municipalities or facilities[10].

Firstly, as multiple municipalities or facilities participate, the recorded data is of a large scale, increasing statistical power. Secondly, it can be examined whether the results from single municipalities or facilities has enough universality to show reproducibility also in the results from the other municipalities or facilities. Thirdly, if comparisons are made between municipalities or facilities in consideration of the differences between the subjects such as of the resident, patient and admitted persons, treatment and care outcomes along with efficiency benchmarks (comparison using numerical indices) can be examined. Fourthly, controlled studies can be performed. In general, because the same intervention technology or program (protocol) is often commonly used at a certain municipality or facility, it is difficult to find control groups for single municipality or facility databases. To verify the effectiveness or efficiency, outcomes need to be compared between different programs (protocols) as controls. The acquisition of data from multiple municipalities and facilities makes it possible to conduct comparisons with such control groups.

IV-2. Five viewpoints for effectiveness evaluation

However, attention must be paid to the fact that simply establishing a large-scale database does not ensure meaningful evaluation. Based on our experience in many evaluation research projects, we will present five important viewpoints for the evaluation of health and long-term care policies and care: (1) evaluation to promote the management cycle, rather than ranking, (2) support of management-led bottom-up type evaluation, (3) frameworks for evaluation that mutually complements with multi-level analysis (macro, mezzo, micro) and multiple elements, (4) multi-dimensional/multilateral evaluation with multiple evaluation criteria and methods, and (5) management of comparisons and evaluations between multiple municipal bodies or agencies by means of benchmarks created using databases[10)].

IV-2-1. Evaluation to promote the management (PDCA) cycle, rather than ranking

Firstly, the objective is visualization. The aim of evaluation is to manage the methods of intervention (involving policies, programs etc.) and increase effectiveness, efficiency and equity. Evaluation equates to the checking part of the Plan-Do-Check-Action (PDCA) management cycle. The aim of evaluation should not be ranking, but rather "visualizing," leading to challenges for actions toward new improvements and keys for overcoming the challenges.

IV-2-2. Support of management-led bottom-up evaluation

Next, let us consider who will lead these "visualization" initiatives. Even if forcing goals on staff with a top-down approach and making objective assessment using indices by external agencies such as researchers or central government, the results of evaluation which are unlikely to be accepted or utilized by management who may in fact exhibit opposition, would be of little significance. "Visualization" requires not merely evaluation, but management-led initiatives, problem-solving or ideas and attitudes that lead to subsequent quality improvement.

IV-2-3. Frameworks for evaluation that mutually complement multi-level analysis (macro, mezzo, micro) and multiple elements

Thirdly, evaluation frameworks must be comprehensive in each of the macro-mezzo-micro levels. In each level, frameworks grasping correlations between many components are necessary. If there is no framework that can grasp elements and interactive relationships as evaluation targets, there is the risk of conducting incorrect evaluation such as assessment indicating that "home care is economical," which does not take into consideration cost burden to families ("cost shift").

IV-2-4 Multi-dimensional/multilateral evaluation with multiple evaluation criteria and methods

The fourth issue is comprehensiveness in evaluation criteria and autonomy. As shown in figure 8, elements and criteria that are targeted for evaluation must be multifaceted, including (1) input (investment, resources), (2) processes, (3) environment, (4) individuals/action, (5) outcomes (effectiveness, results), (6) efficiency and (7) equity.

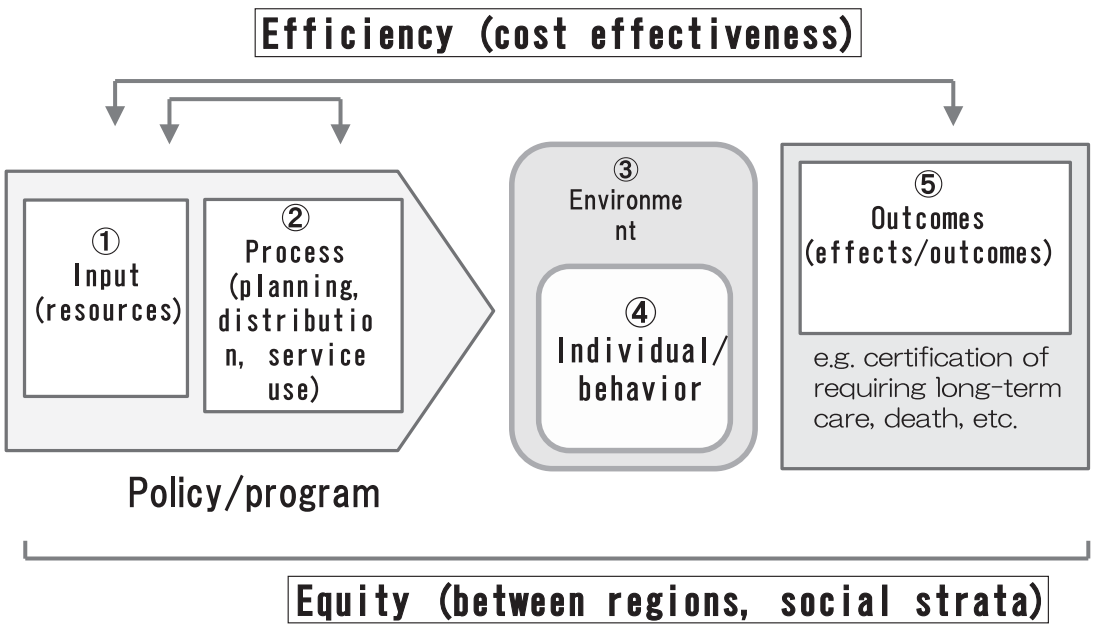
Furthermore, different evaluation standpoints lead to different quality or criteria for evaluation being emphasized. Accordingly, multidimensional evaluation with different viewpoints is also necessary.

IV-2-5. Management of comparisons and evaluations between multiple municipal bodies or agencies by means of benchmarks created using databases

An infrastructure for program evaluation based on the above principles could be provided by the database in which many management bodies such as facilities and municipal bodies participate. The construction of such a database would not necessarily offer perfect evaluation, but it could make multilateral evaluation possible. Furthermore, the participation of management in data collection or the process of development of evaluation indices could increase the likelihood of utilizing the evaluation with the agreement and understanding of management.

If such research is continued in addition to repeatedly revising indices or systems based on evidence and opinions from management to frontline staff, it would be possible to set tar-

Figure 8
Five elements and two aspects of policy evaluation indices groups



gets or carry out planning based on evidence and to evaluate current status, process, outcome, cost effectiveness and "visualization."

IV-3. Challenges in large-scale database development

Large-scale databases have the aforementioned significance and possibilities. However, there are several reasons why data banks have only been previously developed in some areas. These challenges also need to be overcome for the development of more large-scale databases, data management systems and the accumulation of health services research or program evaluation.

We will consider these challenges separately in the four categories: (1) development, (2) data collection, (3) utilization after data collection and (4) challenges related to human resources.

(1) Challenges related to development stage include development bodies and development costs. Independently gathering data involves cost, and it takes 10 years to gather data, develop indices, verify its validity and improve it. Therefore, it is difficult to conduct such continuous research for research groups funded by competitive short-term research grants. Bodies with long-term stable resources will be required to continue development after a prototype system is developed. In countries such as the United Kingdom, such development is conducted by the government.

(2) There are challenges related to data collection such as not being allowed to obtain such even in the case of existing data or individual identifiers necessary for longitudinal follow-up and measures for reducing the cost burden for participating facilities and municipalities[59]. To effectively, efficiently and fairly use limited social security resources, the usage of data owned by municipal bodies etc. will allow little progress without legal support for data use even for public purposes.

(3) If data is not utilized after being gathered, data collection and evaluation will not continue. Good practice examples for utilization need to be gathered.

(4) With regard to challenges related to human resources, there are hardly any postgraduate schools that cultivate human resources capable of conducting evaluation research, and research funds are also scarce. For example, even in university medical schools, the number of faculty posts in the field of social medicine, which is responsible for health services research, is decreasing. Under such circumstances, human resources are not being cultivated, and research is not being conducted.

From my experiences of health services research, I consider that the current status in which evaluations are lacking is extremely poor in terms of cost efficiency compared to societies with evaluation systems established under initiatives by the national government. It is hoped that the challenges mentioned in this report will be overcome.

V. Conclusion

To allocate limited social security resources efficiently, the viewpoint of cost effectiveness analysis is essential. It requires the measurement of effectiveness as well as costs. However, health services research evaluating effectiveness (quality) of health, medical and long-term care has lagged in Japan. We have presented examples of health services research that we have been involved in and discussed achievements, potentials and future challenges of the "visualization" of quality and effectiveness.

Through health services research achievements, by utilizing existing data and adding independent data, the sort of quality evaluation presented in this report is possible. It was obvious that many programs for care prevention, rehabilitation medicine, long-term care and end-of-life care were moderately effective at best and sometimes less so. For acceptance of further burdens in increasing public healthcare and social security costs by Japanese citizens, it is essential to establish a framework that eliminates waste and improves care quality and equity by increasing health and long-term care costs. To do this, the "visualization" of effectiveness (quality), efficiency and equity is required. If costs are simply reduced uniformly without distinguishing between effective and ineffective programs/policies though evaluative research, it will engender the absence of care prevention and good rehabilitation medicine which will lead to an increase in people requiring long-term care and cause older individuals to suffer more during the end-of-life period. In healthcare, "pain" is different to the metaphorical pain of, for example, resisting the urge to buy a desired luxury item. It refers to literal "pain," that deprives people of the most basic human right of "life."

The full-scale development of "the era of evaluation and accountability" and "the era of visualization" in Japan requires the construction of large-scale databases in each field, accumulation of evaluation and management research using these databases and the development of management systems utilizing these visualizations and evidence produced.

Acknowledgement: This is a report compiling research results supported by many study grants. Gratitude is expressed for a Health and Labour Sciences Research Grant (comprehensive research regarding longevity) (2013 – Choju – Ippan – 003), a grant-in-aid for scientific research for basic research A (23243070) and many other research grants (see 2012 preface by Katsunori Kondo).

References

1) Victoria, Cesar G., J. P. Habicht, and J. Bryce (2004), "Evidence-Based Public Health: Moving Beyond Randomized Trials", Am J Public Health, Vol. 94 No. 3, pp. 400–405
2) Kondo, K. (2004) "Moving beyond the 'age of suppression of medical fees'—Medical and Social Welfare Policies in the United Kingdom" Igaku Shoin
3) Rose, G. (1992), The Strategy of Preventive Medicine (supervised by Soda K and Tanaka H: Preventive medicine strategies—Countermeasures for lifestyle-related diseases and

health promotion, Igaku Shoin, 1998), Oxford University Press
4) Ministry of Health, Labour and Welfare/Care prevention Continuous Evaluation Analysis Planning Conference (2008), "Quantitative Effect Analysis of Care Prevention Services (secondary analysis results) (draft)"
5) Health and Welfare Bureau for the Elderly of the Ministry of Health, Labour and Welfare (2007), "Outline of Investigation Regarding Revision of the Method of Determining Specified elderly Individuals"
6) Hiramatsu, M., Kondo, K. and Hirai, H. (2009), "Background Factors for Targets of Care prevention Measures not Undergoing Checkups: Focusing on Socioeconomic Factors", Public Welfare Guidelines, Vol. 56, No. 3, pp. 1–8
7) Hirai., H., Kondo, K., Ojima, T. (2009), "Examination of risk factors for onset of certification of long-term care insurance in community-dwelling older people: AGES project 3-year follow-up study", Journal of the Japanese Society of Public Health, Vol. 56, No. 8, pp. 501–12
8) Gates, S., J.D. Fisher, M. W. Cooke et al., (2008), "Multifactorial Assessment and Targeted Intervention for Preventing Falls and Injuries among Older People in Community and Emergency Care Settings: Systematic Review and Meta-Analysis", Bmj, Vol. 336, No. 7636, pp. 130–33
9) Hayashi, T. and Kondo, K. (2013), "The relationship between the implementation of evidence-based fall prevention programs and the prevalence of certification for the eligibility of the long-term care insurance benefit", Comprehensive Rehabilitation, Vol. 41, No. 4, pp. 359–65
10) Kondo, K. (2012), "Overcoming the 'medical crisis'—Directions for medicine and caregiving in the United Kingdom and Japan" Igaku Shoin
11) Rossi, P. H., H. E. Freeman, and M. W. Lipsey (2005), Evaluation: A Systematic Approach, Sage Publications
12) Kondo, K. (2007), "Verification of the 'health disparity society'—Large-scale epidemiological study for long-term care prevention" Igaku Shoin
13) Kondo, K. (ed) (2010), Health Inequalities in Japan—An Empirical Study of the Older People, Trans Pacific Press (Melbourne)
14) Putnam, RD. (1993), Making Democracy Work, Princeton University Press
15) Kondo, K., JAGES Project (2014), "Visualization of health disparity and the social determinants of health—JAGES 2010–11 Project", Iryo to Shakai, Vol. 24, No. 1, pp. 5–19
16) Idler, E. L., and Y. Benyamini. (1997), "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies", J Health Soc Behav, Vol. 38, No. 1, pp. 21–37
17) Ichida, Y., Yoshikawa, G. and Hirai, H., et al., (2005), "The Health of the Aged and Social Capital: Multilevel Analysis", Transactions of Rural Planning (Journal of Rural Planning Association), No. 7 (separate volume 24), pp. 277–82
18) Ichida, Y., Yoshikawa, G. and Matsuda, R., et al., (2005), "Elderly people in Japan: A large-scale social epidemiological study for care prevention—Social Capital and Health", Japanese Journal of Public Health, Vol. 69, No. 11, pp. 914–19

19) Harpham, T. and K. Mckenzie (2006), Social Capital and Mental Health, Jessica Kingsley Publishers
20) Aida, J., T. Hanibuchi, M. Nakade, et al. (2009), "The Different Effects of Vertical Social Capital and Horizontal Social Capital on Dental Status: A Multilevel Analysis", Soc Sci Med, Vol. 69, No. 4, pp. 512–18
21) Hirai, H. and Kondo, K., (2008), "Development and evaluation of a long-term care prevention program—Taketoyo model for activities to prevent seclusion" Nihon Fukushi University 21st Century COE Program Study of Welfare Society Development, Minerva Shobo, pp. 174–82
22) Kondo, K., Hirai, H. and Takeda, T., et al. (2010), "Social capital and health", The Japanese Journal of Behaviormetrics, Vol. 37, pp. 27–37
23) Hirai, H. and Kondo, K., "Evaluating a care prevention program utilized community salon", Quarterly of Social Security Research, Vol. 46, No. 3
24) Hirai, H. (2010), "Change in Individual-level Social Capital Indicators Among Elderly Salon Program Participants", Journal of Rural Planning Association Vol. 28, pp. 201–206
25) Hikichi, H., N. Kondo, K. Kondo, et al. (2015), "Effect of Community Intervention Program Promoting Social Interactions on Functional Disability Prevention for Older Adults: Propensity Score Matching and Instrumental Variable Analyses, Jages Taketoyo Study", Journal of Epidemiology and Community Health, doi:10.1136/jech-2014-205345
26) Kondo, K., Takeda, T. and Suzuki, K. (2014), "Cost effectiveness analysis of city-building care prevention 'consideration salon activities'', 9th Conference of the Japan Health Economics Association, compilation of general abstracts, p. 20
27) Joint Committee on Guidelines for the Management of Stroke (2009) 'Stroke Treatment Guidelines 2009'
28) Kondo, K. and Oota, T. (1997), "Changes with Time in Cross-Sectional Areas of Leg Muscles in Early Stroke Rehabilitation Patients: Disuse Muscle Atrophy and its Recovery", The Japanese Journal Rehabilitation Medicine, Vol. 34, No. 2, pp. 129–33
29) Kondo, K. (2007) 'Training in getting up, walking and climbing stairs to enhance the healthy side' edited by Kondo, K. and Oi, M. "Stroke rehabilitation—From early rehabilitation to care management" Ishiyaku Publishing, pp. 204–215
30) Matsuda, S. (2010) "Analysis of the state of rehabilitation in the acute phase of cerebral infarction using DPC data", Journal of Clinical Rehabilitation, Vol. 19, No. 6, pp. 607–611
31) Kondo, K. (2004), "Recovery phase rehabilitation wards", Sogo Rehabilitation Vol. 32, No. 4, pp. 305–11
32) Japanese Association of Rehabilitation Medicine (2003), "Fact-finding survey of rehabilitation patient treatment outcomes and medical fees"
33) Jeong, S., K. Kondo, N. Shiraishi, et al. (2010), "An Evaluation of the Quality of Post-Stroke Rehabilitation in Japan", Clinical Audit, Vol. 20, pp. 59–66

34) Matsui, H., H. Hashimoto, H. Horiguchi, et al. (2010), "An Exploration of the Association between Very Early Rehabilitation and Outcome for the Patients with Acute Ischemic Stroke in Japan: A Nationwide Retrospective Cohort Survey", BMC Health Services Research, Vol. 10, No. 1, p. 213
35) Hasegawa, Y., Y. Yoneda, S. Okuda, et al. (2005), "The Effect of Weekends and Holidays on Stroke Outcome in Acute Stroke Units", Cerebrovasc Dis, Vol. 20, No. 5, pp. 325–31
36) Kaifuku Rehabilitation Ward Association (2010), "Survey report on the state of and issues related to recovery phase rehabilitation wards"
37) Kondo, K. (2004), "Rehabilitation Intensity and Its Effects", Journal of the Japanese Association of Rehabilitation Medicine Vol. 41, No. 12, pp. 849–53
38) AHCPR (Agency for Health Care Policy and Research) (1995), Post-Stroke Rehabilitation. Ahcpr Archived Clinical Practice Guidelines. Health Services/Technology Assessment Text (Hstat), The National Library of Medicine (NLM)
39) Royal College of Physicians (2004), National Clinical Guidelines for Stroke. 2nd Edition, Royal College of Physicians
40) The Joint Committee on Guidelines for the Management of Stroke (2004) 'Stroke Treatment Guidelines 2004'
41) Latham, N. K., D. U. Jette, M. Slavin, et al. (2005), "Physical Therapy During Stroke Rehabilitation for People with Different Walking Abilities", Arch Phys Med Rehabil, Vol. 86, No. 12 Supple 2, pp. S41–S50
42) Dromerick, A. W., D. F. Edwards, and M. Hahn (2000), "Does the Application of Constraint-Induced Movement Therapy During Acute Rehabilitation Reduce Arm Impairment after Ischemic Stroke?", Stroke, Vol. 31, No. 12, pp. 2984–88
43) van der Lee, J. H., R. C. Wagenaar, G. J. Lankhorst, et al. (1999), "Forced Use of the Upper Extremity in Chronic Stroke Patients: Results from a Single-Blind Randomized Clinical Trial", Stroke, Vol. 30, No. 11, pp. 2369–75
44) Langhorne, P., R. Wagenaar, and C. Partridge (1996), "Physiotherapy after Stroke: More Is Better?", Physiother Res Int, Vol. 1, No. 2, pp. 75–88
45) Nagai, S., Sonoda, S., et al. (2003), "Effects of the Full-Time Integrated Treatment (Fit) Program", Sogo Rehabilitation, Vol. 31, pp. 175–83
46) Saito, E. and Sonoda, S. (2003) "The Fit program—Working to implement it in comprehensive high intensity rehabilitation wards" Igaku Shoin
47) Singh, N. A., K. M. Clements, and M. A. Fiatarone (1997), "A Randomized Controlled Trial of Progressive Resistance Training in Depressed Elders", J Gerontol A Biol Sci Med Sci, Vol. 52, No. 1, pp. M27–35
48) North, T. C., P. McCullagh, and Z. V. Tran (1990), "Effect of Exercise on Depression", Exerc Sport Sci Rev, Vol. 18, pp. 379–415
49) Kondo, K. (2004), "Impact of Recovery Phase Rehabilitation Ward: A Policy Evaluation", Journal of the Japanese Association of Rehabilitation Medicine, Vol. 41, No. 4, pp. 214–18

50) Kondo, K. (2004), "Recovery phase rehabilitation wards", Sogo Rehabilitation, Vol. 32, No. 4, pp. 305–11
51) The Japanese Association of Rehabilitation Medicine (2004), "Patient outcomes depending on the presence of a rehabilitation specialist—Degree of ADL improvement, Relationship between ADL improvement rates and home discharge rates"
52) The Japanese Association of Rehabilitation Medicine (2004), "Implementation of regular conferences and outcomes for rehabilitation patients—Relationship between degree of temporary improvement and ADL improvement rates"
53) Uniform Data System for Medical Rehabilitation <http://www.udsmr.org/>
54) Kobayashi, S. (2005) "Stroke Data Bank 2005", Nakayama Shoten
55) The Japan Neurosurgical Society (2004) "Unruptured Cerebral Aneurysms Study (UCAS), Japan)"
56) Tanaka, K. and Aruga, T. (2005), "Updated knowledge 1 regarding cranial injuries: data bank" edited by Ishida, A. "Cranial injuries for care staff, patients and their families—Guidelines for understanding the disease and overcoming impairment" Ishiyaku Publishing, pp. 15–17
57) Suka, M., Yoshida, K. and Takezawa, J. (2004), "Evaluation of the medical burden of ICU-acquired infections", Japanese Journal of Infection Prevention and Control, Vol. 19, pp. 389–94
58) Kawai, N., Ikematsu, H., Iwaki, N. (2004) "Investigation of influenza at facilities throughout Japan for the 2003/2004 season—Survey of use of an online database for prevention, diagnosis and treatment", Japan Medical Journal, No. 4204, pp. 26–32
59) Matsui, T. and Tohma, S. (2005) "Rheumatoid arthritis: From causative research to a new age for treatment—A treatment and research support system" Japanese Journal of Clinical Medicine, Vol. 63, pp. 699–705
60) Kawaguchi, M., Yagi, M., Kimura, H., et al. (2001) "Effects of reperfusion therapy on in-hospital death in acute phase myocardial infarction patients—Based on hijc multi-institutional joint research in 1999", Cardioangiology, Vol. 50, pp. 504–509
61) Nakamura, Y (2001) "Perinatal statistics at 40 national hospitals (2000)", Iryo, Vol. 55, pp. 375–81
62) Kondo, K. and Yamaguchi, A. (2005), "Possibilities and issues related to large-scale data banks to build up evidence", Sogo Rehabilitation, Vol. 33, No. 12, pp. 1119–24
63) Health Labour Sciences Research Grant for Research Projects on Aging and Health (Principal investigator: Yamaguchi, A.) 2005 General Study Report (2006) "Research regarding the construction of a community rehabilitation system for elderly individuals"
64) Health Labour Sciences Research Grant for Research Projects on Aging and Health (Principal investigator: Kondo, K.) 2007–2009 General Study Report (2009) "Research regarding the development of a rehabilitation patient data bank (DB)"
65) Jeong, S., Inoue, Y., Kondo, K., Matsumoto, D. and Shiraishi, N. (2014), "Formula for predicting FIM for stroke patients at discharge from an acute ward or convalescent rehabilitation ward", Japanese Journal of Comprehensive Rehabilitation Science, Vol. 5,

pp.19–25
66) Ito, M. and Kondo, K. (2012), "Care quality achievements and issues—Focusing on evaluation at special elderly nursing homes", Quarterly of Social Security Research, Vol. 48, No. 2, pp. 120–32
67) Ito, M. (2007), "State of and evaluation of systems for care quality management in US nursing homes: Based on interviews of managers", Japanese Journal of Social Welfare, Vol. 48, No. 1, pp. 153–66
68) Sawada, Y., Kondo, K. and Ito, M. (2007), "Care quality management systems in US nursing homes—Based on a review of the literature and empirical experience in clinical settings", Journal of the Japan Society for Healthcare Administration, pp. 293–301
69) Nagasawa, M. (2012), "Development of care quality evaluation indices and related issues—Focusing on international trends and outcome indices in the United Kingdom", Quarterly of Social Security Research, Vol. 48, No. 2, pp. 133–51
70) Ito, M., Kondo, K., Izumi, M. et al. (2013), "Attempt at quality evaluation of care in special elderly nursing homes using data related to certification for requiring long-term care—Creation of 11 indices and comparison between facilities", Public Welfare Guidelines, Vol. 60, No. 5, pp. 14–23
71) Ito, M., Kondo, K., Shimizu, Y., et al. (2013), "Verification of the validity of care quality evaluation indices created from data related to certification for requiring long-term care—Through a visiting survey of special elderly nursing homes with a blind study", Health Labour Sciences Research Grant (Research representative: Kondo, K.), edited by Development of benchmarks and a system for comprehensive policy evaluation of long-term caregiving insurance (2010—Research Projects on Aging and Health—Designated-008) Health Labour Sciences Research Grant (Research Projects on Aging and Health) Research Team, pp. 282–88
72) Kondo, K. (2008), "Issues in end-of-life care at home: Based on a national survey and initiatives in the United Kingdom", Journal of the Japan Society of Hospice & Homecare, Vol. 16, No. 3, pp. 244–56
73) Sugimoto, H., Kondo, K., Higuchi, K., et al. (2007), "Reliability and usefulness of the Japanese version of mds-Pc (Minimum data set for palliative care)", Journal of the Japan Society for Healthcare Administration, Vol. 44, No. 3, pp. 243–51