Decay of impact after a health-education program for people with chronic diseases: preparing for reinforcement by analysis of prevalence, magnitude, timing, and predictors of decay

慢性疾患患者における健康教育プログラム実施後の decay of impact: decay のタイミング、割合、大きさ、および予測因子の分析から

MJ Park
朴　敏廷

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

13 February, 2013

Department of Health Communication
Graduate School of Medicine
The University of Tokyo

Table of contents

List of Tables

List of figures

List of appendices

Abstract

Background: Education for chronic disease self-management has benefits, but those benefits decrease over time. This "decay of impact" makes reinforcement programs necessary, but, for unknown reasons, those reinforcements are ineffective. This study's aims were to measure basic characteristics of the decay of impact, and to identify predictors of decay.

Methods: Adults with various chronic illnesses participated in a 6-week self-management educational program (n = 364). Before the program, and again 3, 6, and 12 months later, self-rated health, pain, coping, communication with medical doctors, self-efficacy, health-related distress, anxiety, and depression were measured. For each outcome, the prevalence of decay, its magnitude, and the timing of its onset were determined. Classification trees were used to identify predictors of decay.

Results: The prevalence of decay ranged from 7% (pain) to 26% (self-rated health). Its median magnitude ranged from 16.4% of the full-scale value (depression) to 39.5% (pain). The decay started 3 months after the program began in 27%-61% of the participants, depending on the outcome. For self-rated health, coping, and anxiety, the classification trees gave good predictions of the need for reinforcement. The length of disease history was a good predictor on 6 outcomes.

Conclusions and recommendations: The fact that decay of impact occurred only in some of the participants can explain why reinforcements appeared to be ineffective in previous studies. In future studies, people likely to need reinforcement can and should be identified prospectively. Soon after the main program, reinforcement programs should be offered to those who will need them

1. Introduction and literature review

1.1 The context, and the importance of self-management

The burden of chronic illnesses is increasing [1, 2]. As of 2008, "the global burden of disease is now, and will continue to be, dominated by chronic non-communicable diseases" [3]. Chronic diseases have a particularly large impact in Japan and other developed countries. In 2011, nearly half of all deaths in Japan were attributed to cardiovascular diseases, respiratory diseases, diabetes, and other non-communicable diseases (not including cancers) [4].

For people suffering with chronic diseases, medical treatment is obviously important, but it is not enough. Because patients with chronic diseases make important health-related decisions every day, by default they self-manage their illnesses. This requires a "patient-professional partnership, involving collaborative care and self-management education" [5]. In a 2010 report, the World Health Organization acknowledged that people with chronic conditions need not only medical interventions but also educational interventions promoting self-management [page 66 of reference 6]. (Even low-income countries will have this need [7].) Also very recently, an international group of experts proposed that the definition of health should emphasize "the ability to adapt and self-manage" [8]. This awareness of the importance of self-management is consistent with a major trend in health education: away from the mere provision of information and toward the promotion of skills and strategies for handling the daily problems caused by chronic illnesses [9].

1.2 The Arthritis Self-Management Program

One early example of the trend toward education for problem-solving skills is the Arthritis Self-Management Program (ASMP), which was developed at Stanford University [10]. The aim of that program is not to provide specific medical information. Instead, participants learn

> "1) techniques to deal with problems such as pain, fatigue, frustration and isolation, 2) appropriate exercise for maintaining and improving strength, flexibility, and endurance, 3) appropriate use of medications, 4) communicating effectively with family, friends, and health professionals, 5) healthy eating, 6) making informed treatment decisions, 7) disease related problem solving, and 8) getting a good night's sleep." [11].

This program was first offered more than 25 years ago in the United States, and since then it has been implemented widely (in, for example, Australia, Canada, China, Hong Kong, Lithuania, the Netherlands, New Zealand, the United Kingdom, and the West Indies).

Results of the initial research on the program indicated that it was beneficial [10]. Specifically, the program increased the participants' knowledge about arthritis and it increased the frequency with which they did health-related behaviors (recommended physical activities, relaxation, and arthritis-specific exercises). It also reduced their pain and their disability. However, further analyses revealed two unexpected findings: (1) the increases in knowledge were not correlated with the health benefits, and (2) the associations between health-related behaviors and health outcomes (pain, disability, and depression) were very weak [12]. Those findings indicated that the benefits of the program were not caused by the participants having more

knowledge about arthritis, and they were also not caused simply by the participants doing the recommended health-related behaviors.

A hypothesis to explain those unexpected findings came from the results of a mixed qualitative-quantitative exploratory study [13]. In that study, patients who had participated in the ASMP were interviewed. Benefits of the program were defined as pain relief, decreased disability, and fewer arthritis-related visits to a physician. Patients who benefitted from the program were then compared with patients who did not benefit from the program. The main finding was that the patients who benefitted from the program had positive emotions and they believed "that they had more control over their disease," whereas the patients who did not benefit from the program had negative emotions and they believed that they had "a lack of control." That finding led to the hypothesis that confidence and perceived self-efficacy [14,15] are important determinants of the outcomes of this program. That hypothesis was tested [16, 17], and the results supported the idea that educational programs to promote self-management of chronic medical conditions are successful to the extent that they increase their participants' perceived self-efficacy.

1.3 The importance of perceived self-efficacy

The concept of perceived self-efficacy was developed mainly by Albert Bandura [14, 15]. As defined by Bandura, "Perceived self-efficacy refers to beliefs in one's capabilities to organize and execute the courses of action required to manage prospective situations."[15, page 2]. As applied to education for chronic illness self-management, it means that patients need something more than knowledge about their disease. They also need something more than knowledge about specific actions they could take

(things they could do) to keep their disease under control. In addition to knowledge, they need to believe that they will be able to do those things at the appropriate time and place, and they also need to believe that doing them will cause results that they desire. Promotion of self-management will be successful only to the extent that patients have such beliefs. The beliefs referred to in self-efficacy theory can change, so self-efficacy can increase, and the most effective way to increase self-efficacy is to directly and personally *experience success* in taking specific actions that cause desired results. Self-efficacy can also be increased by watching others succeed in those actions, by verbal suggestion and persuasion, and by being in an appropriate physiological state (e.g., not feeling anxious about doing something new) [18].

For example, a patient with a chronic illness might be told that she should make a list of questions about her illness and about its treatment. She might also be told that she should use that list and ask those questions during her next clinic visit. However, according to self-efficacy theory, if she has never actually made such a list and used it in the past, then she will be less likely to believe that she can do it, and she will be less likely to do it in the future. To increase her self-efficacy in this area, she can begin by writing only one or two questions, and by seeing how other people made lists and how they asked their doctors questions. She can also be verbally encouraged to make and use such a list, and if she is worried or anxious about having such a discussion with her doctor then she can be taught relaxation techniques to relieve her anxiety when she is in that situation. Once her self-efficacy for this activity has increased, she will be more likely to make a list of questions and to use it when she visits her doctor. She might then be able to control her disease better.

Communicating effectively during a clinic visit is of course not the only thing patients can do to improve their self-management of chronic illness. Other such actions include doing daily physical exercise, using community services and support, taking medications appropriately, communicating effectively with family members and others, managing nutrition and diet, and controlling the psychological problems that can accompany chronic illness [19]. Aiming to increase self-efficacy in those and related areas, the group at Stanford University modified the ASMP. They incorporated new activities specifically designed to increase self-efficacy [17]. Then they compared the effects of the modified program with the effects of the original program. They found that reductions in pain, disability, and depression were from 1.4 to 12 times greater after the modified program than after the original program [17]. As a result, they adopted self-efficacy theory as the conceptual foundation for their further development of self-management programs.

In this context, increasing self-efficacy means building confidence to deal with disease-related problems [20] and enhancing one's motivation to implement strategies required for managing one's chronic disease and for reducing its effects on daily life [21]. One of the most consistent successes of programs such as the ASMP is that they improve self-efficacy [20, 22, 23-28].

1.4 Disease-specific programs and multimorbidity

As the success of the ASMP became known, other self-management educational programs were designed and implemented for people with other chronic medical conditions. Among them were disease-specific programs for people with diabetes [29, 30], asthma [31], chronic obstructive pulmonary disease [32], inflammatory bowel disease [33], heart failure [34, 35], macular degeneration [36], other chronic vision problems [37], HIV / AIDS [38], chronic

pain [39], osteoporosis [40], and cardiac pain caused by chronic stable angina [41].

However, many adults with those diagnoses have multimorbidity. That is, they have more than one chronic condition [42, 43, 44]. This poses a challenge to researchers who study disease-specific self-management programs: Who should be included in those studies? The problem can be illustrated with an example: If some ASMP participants have arthritis only, and they participate together with others who have multimorbidity including arthritis, then the results of research on that program will be generalizable to the population of all patients with arthritis because the sample is representative of that population. It is representative because many patients in that population do have multimorbidity. However, those results will be influenced by the program's effects on the participants who have other conditions in addition to arthritis, and therefore with regard to *arthritis alone* the results will be less valid. (Even if every participant in an ASMP program had only one diagnosis, results of research on that program would still reflect more than one diagnosis, because in the ASMP the definition of "arthritis" is not strict: "People with different types of rheumatic diseases, such as osteoarthritis, rheumatoid arthritis, fibromyalgia, lupus, and others, attend together [11]."") As in all clinical research, there is a trade-off between generalizability (i.e., representativeness, external validity) and interpretability (i.e., internal validity) [45].

If a study of the ASMP excluded patients with multimorbidity, then the results would be irrelevant to many people with arthritis and there would be a danger of selection bias [46]. As noted by Berger, "Bench scientists and animal researchers tend to emphasize higher internal validity, while social

and behavioral scientists and epidemiologists tend to prefer enhanced representativeness" [45]. Research on chronic disease self-management follows that tendency: the trade-off between representativeness and internal validity is usually resolved by sacrificing some internal validity in order to increase representativeness. Specifically, researchers in this field generally do not exclude patients who have multiple chronic conditions. It is important to remember that, because multimorbidity is common and because researchers strive for representativeness, *research on disease-specific programs generally does not give disease-specific results*.

Multimorbidity is important not only for research but also for practice. Should patients with more than one chronic disease participate in more than one disease-specific program? How many different disease-specific programs should be implemented? If resources are limited, which diseases should receive attention and which should be ignored? These problems can be avoided if patients with different diseases have important self-management challenges in common. If patients with different chronic conditions have self-management needs that overlap, then it should be possible to design a generic program – a single intervention that will benefit patients with different diagnoses, and will also benefit patients with multimorbidity.

1.5 The generic program

Around the same time that the early disease-specific programs were being studied, research was also being done on the similarities of self-management needs across many chronic conditions. Of course there are differences related to diagnosis: for example, people whose only medical problem is chronic pain caused by arthritis do not need to remember to monitor their blood-sugar level or to avoid specific triggers of an asthma

attack. Nonetheless, by 1991 it was clear that people with asthma, arthritis, diabetes, heart disease, and chronic obstructive pulmonary disease in fact do share important self-management needs [47]. Specifically, all of them need self-management skills for using medications, exercise and other physical activity, and smoking cessation. Most of them also need self-management skills for managing diet and nutrition, acute episodes and emergencies, relations with significant others, and emotional and psychological responses to illness. They also need skills for recognizing and responding to signs and symptoms, and for controlling triggers. Across different chronic diseases "there are many commonalities in the nature of self-management tasks" [48, page 369]. In short, people with different chronic conditions have many self-management needs in common. That finding, together with the increasing prevalence of multimorbidity, led to the development of the generic Chronic Disease Self-Management Program (CDSMP) [49]. People with different diagnoses attended the same CDSMP sessions together. This program has been implemented even more widely than the ASMP.

The aims of the CDSMP are, as with the ASMP, to increase the participants' skills and confidence, that is, their self-efficacy, for chronic disease self-management in general. As described by its developers, the content of the CDSMP focuses on six areas:

> "1) techniques to deal with problems such as frustration, fatigue, pain and isolation,
>
> 2) appropriate exercise for maintaining and improving strength, flexibility, and endurance,
>
> 3) appropriate use of medications,
>
> 4) communicating effectively with family, friends, and health

professionals,

5) nutrition, and,

6) how to evaluate new treatments." [50]

Even though this program is considered to be educational and it uses a textbook [51], there are no classes and there is no teacher. There are workshops and each workshop has two lay leaders. These leaders are not medical professionals. Most lay leaders either have a chronic disease or have personal experience with a chronic disease in one of their family members. The leaders underwent at least 30 hours of training [50], as specified by the original developers of this program The leaders' job is not to give medical information or instructions. It is to facilitate and manage discussions based on the textbook. The people who attend are not considered to be students. Instead, they are participants. The topics of their discussions are introduced in the textbook, and through their discussions they realize how others have experienced and responded to problems similar to their own, even if their diagnoses are different. They talk about how to manage those problems. They learn some self-management skills from the textbook and they also learn from each other. Then they practice what they learned, to try to make those skills into new habits. They focus less on what is difficult, and more on what is possible.

1.6 Effectiveness of the generic program

The effects of this generic program were tested in a randomized controlled trial with a six-month waiting-list control [49]. There were 15 outcome measures in three areas: four measures of health-related behaviors, eight measures of health status, and three measures of health-service utilization. The participants were adults with arthritis, stroke, lung disease,

heart disease, or combinations thereof. People with multimorbidity were not excluded. The results showed that for the group as a whole the CDSMP reduced the number of hospitalizations and inpatient days. The intervention group also benefitted with regard to fatigue, disability, physical exercise, self-reported health, health-related distress, limitations on social activities, communication with physicians, and cognitive symptom management.

This trial was unusual because of its size. It had 391 patients in the control group and 561 in the intervention group. Because so many patients were involved, the data from those with multimorbidity could be analyzed separately. The control group had 225 patients with multimorbidity and the intervention group had 311. As compared with the multimorbidity control group, the multimorbidity intervention group improved on 10 of the 15 outcome measures. The improvements were in all three areas: health-related behaviors, health status, and health-service utilization.

Similar to the data on patients with multimorbidity, data on participants who had only one diagnosis could also be analyzed separately. The numbers of patients in these single-diagnosis subgroups were as follows, for control and intervention, respectively: arthritis only, 62 and 86; heart disease only, 31 and 45; and lung disease only, 60 and 107. The results showed that, in general, for all three of the diagnoses, the patients who participated in the CDSMP improved more than did those in the corresponding control subgroup. Specifically, as compared with their respective controls, the heart-disease-only intervention subgroup improved on 13 of the 15 outcome measures, the arthritis-only intervention subgroup improved on 8 outcome measures, and the lung-disease-only intervention subgroup improved on 7 outcome measures. For all three single-diagnosis subgroups, the intervention

resulted in improvement on at least one outcome measure in each of the three areas. For all three diagnoses, the intervention reduced disability, it improved self-reported health status, and it increased both aerobic exercise and the use of cognitive techniques for symptom management. Two of the outcomes concerned physical activity, and the authors noted that patients who began the study with relatively high levels of physical activity maintained those high levels, while the intervention increased the level of physical activity in patients whose levels had been low at baseline.

Those findings for the three single-diagnosis subgroups are positive, but exactly how they should be interpreted is not completely clear. The reason is that, as mentioned above, the researchers could obtain a representative sample only by sacrificing interpretability. Even though those participants had only one diagnosis, they participated together with people who had diagnoses different from their own, and together with people who had multimorbidity. Would the participants who had (for example) only arthritis have benefitted more if all of the other participants in their discussion group had also had arthritis only? That question is interesting, but answering it from the available data is impossible. The authors pointed out that the program could accommodate the different disease-related challenges faced by different participants. They emphasized that activities in the CDSMP included "aiding patients to identify their own individual needs and problems and then assisting them to work most intensively in those areas" and that the program was also designed to meet the needs of people with multimorbidity.

The results of this study showed how a generic program can be useful. With regard to the practicality and the effectiveness of generic programs, the authors concluded that "These results indicate that it is possible to educate

patients with different chronic diseases successfully in the same intervention at the same time" [49]. Since then, two other studies of patients with arthritis have directly compared the generic program (the CDSMP) with the arthritis-specific program (the ASMP) [52, 53].

In the first of those two studies, both programs had benefits [52]. Some of the differences between the programs were statistically significant, but they were not large, and those differences had diminished by the time of the 1-year follow-up. One conclusion of that study was that the generic program is a reasonable alternative to the disease-specific program. Another conclusion was that the generic program is likely to be particularly useful, because "more individuals can be reached at less cost" and "a generic program may reach larger numbers of persons with arthritis because arthritis is one of the most common comorbid conditions" [52].

In the second study directly comparing the ASMP with the CDSMP, both programs had benefits, but the results favored the generic program [53]. At the time of the 4-month follow-up, on the measures of pain and of disability the generic program was superior to the disease-specific program. At the time of the 12-month follow-up the two programs did not differ, except that the participants in the generic program had slightly fewer visits to a physician. Overall, the generic program was at least as beneficial as the disease-specific program, and in some ways the generic program was better. The researchers commented that patients with multimorbidity "may see the CDSMP as being most relevant to their lives because of its generic focus" [53].

A very recent publication reports the results of a search for predictors of the outcomes of the CDSMP [54]. In data from 1,385 patients, there was no evidence of any consistent association between diagnosis and the

effectiveness of the generic program. The conclusion in 2011 was essentially the same as that published 12 years earlier: the generic program is "useful to a wide range of people with chronic illness."

1.7 Further research on the effectiveness of self-management education

As lay-led self-management programs have become more common, some of them have been studied in controlled trails. The results of 17 randomized controlled trials of such programs were meta-analyzed in a 2009 Cochrane review [22].

Only two of those 17 studies maintained both control and intervention groups for more than six months, so the reviewers found that there was not enough evidence to determine how long these programs' effects are maintained. They described the limitations as follows: "Thus there is insufficient information to state whether any benefits seen would be sustained over time, or indeed whether there might be benefits which only become apparent in the long term" [22]. Of those two limitations, the latter refers to the phenomenon known as "delay of impact" and the former refers to the one known as "decay of impact," which is the main topic of this thesis [55].

As for short-term effects, the meta-analysis results indicated that these programs have statistically significant benefits on seven of 11 outcomes that were studied. Those benefits are reductions in pain, fatigue, disability, and psychological depression, and also improvement in self-rated general health. Regarding health-related behaviors, the benefits are increases in the frequency of aerobic exercise and increases in the frequency of the use of cognitive techniques for symptom management. There were no statistically significant differences between intervention and control groups on the measures of psychological well-being or health-related quality of life, and none in the

number of visits to a physician or in the number of days or nights in hospital. To express the sizes of the effects, the authors of that review used the standardized mean difference (SMD) or the weighted mean difference (WMD) between control and intervention groups. The largest effect was for symptom management: WMD = -0.55 (95% CI -0.85 to -0.26). Judged using Cohen's guidelines [56, pages 24-26], the size of that effect would be considered to be "medium." Among the six other outcomes with statistically significant differences, all of the effect sizes were smaller: pain, SMD = -0.10 (-0.17 to -0.04); disability, SMD = -0.15 (-0.25 to -0.05); fatigue, SMD = -0.16 (-0.23 to -0.09); depression, SMD = -0.16 (-0.24 to -0.07); self-rated general health, WMD = -0.20 (-0.31 to -0.10); aerobic exercise, SMD = -0.20 (-0.27 to -0.12). Therefore, the conclusion of the review was that none of the effects are large, and some of them are "not clinically important" [22].

Other studies done recently had similar results. Five such studies and their findings are described briefly below:

1. Gitlin, et al. [57] measured changes over four months after a modified CDSMP for elderly African Americans. They found statistically significant improvements regarding fatigue, exercise, self-efficacy, health-related distress, illness intrusiveness, and the use of cognitive coping strategies. The relief of health-related distress was of "medium" size, and all of the other effects were small.

2. In an evaluation of short-term outcomes of the CDSMP in Hong Kong, Siu, et al. [58] found increases in self-efficacy, in exercise, and in the use of cognitive techniques for coping with symptoms. While those improvements were statistically significant, all of them were small.

3. Barlow, et al. measured the effectiveness of the CDSMP in patients with multiple sclerosis [59]. They found that the program increased self-efficacy, and it reduced the physical and psychological impact of multiple sclerosis. All of the effects were small.

4. Lorig, et al. [60] studied an Internet-based version of the CDSMP. They found that the program relieved pain and shortness of breath, reduced fatigue and health-related distress, and increased the frequency of stretching and strength-building exercises. Nonetheless, all of those effects were small.

5. Jerant, et al. [61] measured the effects of an in-home version of the CDSMP. They found that the program increased self-efficacy, but the effect was small and it lasted for less than one year.

In summary, the research results now available indicate that chronic-disease self-management education does have some benefits, but those benefits are small and they are restricted to only a few of the measured outcomes. Also, for effects lasting longer than 6 months the evidence is weak.

1.8 Effects in subgroups

Even before the most recent results of the Cochrane meta-analysis were published, the original developers of the CDSMP acknowledged that the program's effects are not large. They explained that fact by saying that it is a consequence of the composition of the population of patients with chronic diseases: "It should be noted that the population is very heterogeneous for disease, age, education and symptom distribution. Thus group changes and mean effect sizes tend to be modest." [60]. That explanation implies that the effect sizes will not be small if they are not whole-group means. It implies that there are some subgroups with very small effects or perhaps zero effects, which dilute (partly efface or obscure) large effects in other subgroups. If that

explanation is true, then larger effects would be measured if homogeneous subgroups were analyzed separately.

Some research on subgroups has already been done. As noted above in section 1.6, the CDSMP was found to be beneficial in three different single-disease subgroups [49]. In addition, there is evidence that self-management programs are more beneficial in women [62], in younger patients [62, 63], in subgroups with better cognitive status and lower educational levels [64], in patients with certain personality characteristics [65], in Vietnamese-speaking and Chinese-speaking patients in Australia [66], and in patients with lower baseline levels of self-efficacy and health-related quality of life [63]. In the study by Nolte, et al. [62], the subgroups were defined by their pattern of change after the program, and data from pattern-defined subgroups are also essential to the present thesis, as described below.

1.9 Duration of effects, decay of impact, and reinforcement programs

Among all of the results that were meta-analyzed in the Cochrane review [22], none referred to effects lasting longer than 6 months. For such long-term effects no results could be meta-analyzed because only two long-term studies met the inclusion criteria. (This probably reflects the practical difficulty of maintaining a group randomized to receive no educational intervention for more than 6 months.) Thus, there are important questions about the long-term effectiveness of these programs. One study of very-long-term (8-year) outcomes has been reported [67]. However, that study was observational, and for practical and ethical reasons it may now be impossible to conduct randomized trials lasting longer than a few months.

Nonetheless, the benefits of these programs are generally believed to be short-lived. Writing about self-management programs, Riemsma, et al. [68]

noted that "Short-term effects are rarely maintained over long intervals" and according to Mulligan and Newman such programs' "effects tend not to be maintained." [69] This phenomenon is not universal [59, 70], but it has been seen to various extents in many studies [65, 67, 71-75], and it has been given various names: attenuation [65, 67, 71], deterioration [72], relapse [73], backsliding [55], and decay of impact (which is the name used here) [55] (Figure 1).

Considering the decay of impact to be common and important in health education, LW Green wrote more than 30 years ago that it should be prevented, or its effects mitigated, by reinforcement: "We know that reinforcement is as important to [health] education as booster shots are to sustained immunization." [55] Reinforcement is commonly used to maintain the effects of treatments for addictions [76]. With regard to education for self-management among people with chronic illnesses, reinforcement programs (booster sessions, telephone follow-up, etc.) are recommended in order to realize long-term benefits [77].

Some reinforcement programs have been tested, but the results of those tests have been mixed, and some are counterintuitive. In one study of arthritis self-management education, there seemed to be little or no decay after the program, and reinforcement did not alter its effects [78]. In a separate study also of patients with arthritis, no benefit of booster sessions was found [68]. Similar results have been reported with regard to other chronic conditions. In a study of a diabetes self-management intervention [79], telephone follow-up did result in improvement on a biological measure (lipid ratio) but it "did not generally produce meaningful incremental effects". In the same study, reinforcement had the opposite effect on psychosocial measures (and

particularly on the Chronic Illness Resources survey), that is, it "appeared to produce less improvement … than conditions not receiving the telephone follow-up". In studies of a different diabetes self-management program, "automated telephone reinforcement did not improve the effectiveness" of the program [29], and neither did reinforcement via a discussion group for peer support [30]. In the latter study, two results were directly contrary to expectations: patients who were randomized to the non-reinforcement arm of the study reported greater relief from health-related distress, and the reduction in their depression was also greater. In a pilot study of internet-based support for self-management of dyspnea, people with chronic lung disease did not benefit from a booster [80].

In summary, although reinforcement is needed, direct comparisons between reinforcement and non-reinforcement have given the former almost no empirical support. Reinforcement programs are necessary, but they are ineffective.

1.10 Questions about reinforcement and decay of impact

Why are reinforcements ineffective? One possible answer is related to the findings discussed in section 1.8 above. That is, it is possible that reinforcements are needed only by a subgroup of participants – a subgroup (or subgroups) with decay of impact. If that is true, then reinforcements appear to be ineffective because only some of the people who receive them in fact need them, so their benefits are diluted by whole-group analyses that include their lack of effect among people who do not need them.

It stands to reason that reinforcements can be optimized on the basis of a clear and accurate understanding of the phenomenon that they are intended to prevent or mitigate. Continuing LW Green's analogy with immunization

[55], one might say that the effects of booster shots can be optimized only if one knows when, by how much, and in whom immunity will decay if a booster is not given. Similarly, if one wants to use reinforcement to prevent or reduce the decay of impact, then the characteristics of that decay should be studied. For example, in education for preventing the spread of HIV / AIDS, information about when the decay of the educational impact begins can be used to decide when to begin reinforcements [75]. On the basis of the research done up to now, reinforcements after the CDSMP cannot be planned rationally, because previous studies give no information about what proportion of the participants have decay of impact and therefore need reinforcements, about how strong that need is (i.e., the magnitude of the decay), or about when reinforcements are needed (i.e., when the decay begins). In addition, to minimize the costs of planning and implementation, reinforcements should be offered preferentially to the participants who are most likely to need them (i.e., to those who are predicted to have decay). All of that information about the decay of impact would have important practical implications, but it is not available from previous research. In that context, the present thesis addresses two main issues:

1. What are the prevalence, magnitude, and timing of the decay of impact after the CDSMP?
2. Is it possible to predict which participants are most likely to have decay of impact? That is, is it possible to know, before the decay begins, which participants are at higher risk?

To the extent that those questions can be answered, the design and implementation of reinforcement programs will begin to have an empirical foundation.

1.11 Summary, and aims of this study

This is a study of patterns of change after a particular educational program for people with chronic diseases. It is focused on the one pattern called "decay of impact." The aims of this study were to acquire information about the basic characteristics of the decay of impact (prevalence, magnitude, timing, etc.), and to find a way of predicting, before the decay begins, which participants are most likely to have it and thus to need reinforcement.

2. Methods

2.1 Recruitment of participants, ethical approval, and informed consent

The participants were adults with chronic diseases who joined a program aimed at enhancing their ability to self-manage their chronic illnesses. They were recruited using an announcement on the Internet homepage of the Japan Chronic Disease Self-Management Association [81], and by referrals from flyers left in public service centers. The program was made available to women and men equally.

This study was approved by the Research Ethics Committee of the Graduate School of Medicine at the University of Tokyo (IRB#: 1472-(3)) (Appendix 1). Participation in the program and in this research were voluntary, and informed consent was obtained in writing before the study began (Appendix 2).

2.2 The educational program

The program comprised group-discussion sessions with 5 to 13 participants. There was one session each week for six consecutive weeks. Seventy six programs of six sessions each were held between August 2006 and April 2010. The first programs were held in Kumamoto Prefecture, and later programs were held in Tokyo and in Aichi, Chiba, Fukuoka, Hokkaido, Hyogo, Kagoshima, Miyagi, Okayama, Osaka, Saga, and Saitama Prefectures.

Each discussion group had two lay leaders. Their function was not to teach, but to facilitate and manage the discussion.

2.3 Measurements

All data were collected via self-administered questionnaires. The measurements of health status, self-management behaviors, and

psychological health (all described below) were based on those used by Lorig, et al. [82], and are the same as those used by Yukawa et al. [83].

There were 2 measures of health status: (i) Overall health status was self-evaluated on a 5-point scale, with 1 = excellent and 5 = poor. (ii) Pain during the previous 2 weeks was measured on an 11-point scale, with 0 indicating no symptoms and 10 indicating severe symptoms.

There were 2 measures of self-management behaviors: (i) The use of 6 different cognitive techniques to cope with symptoms was measured on a 6-point scale, with 0 = never and 5 = always (Cronbach's alpha = 0.72). (ii) The use of 3 different proactive methods for improving communication with medical doctors was measured on a 6-point scale, with 0 = never and 5 = always (Cronbach's alpha = 0.78).

There were 4 measures of psychological health: (i) The number and frequency of symptoms of anxiety in the past week was measured using 7 items on 4-point scales (Cronbach's alpha = 0.83) [84]. (ii) Symptoms of depression were measured on a similar scale (Cronbach's alpha = 0.72) [84]. (iii) Health-related distress in the past month was measured using 4 items on 6-point scales to measure health-related discouragement, fear of the future, worries, and frustrations (Cronbach's alpha = 0.92). (iv) Self-efficacy to manage chronic conditions was measured with 6 questions (Cronbach's alpha = 0.92). Four of those 6 questions asked about participants' confidence in their ability to do things they want to do despite chronic-disease symptoms. The other 2 questions asked about managing their chronic conditions in order to reduce the number of doctor visits, and about reducing the effects of their chronic conditions on daily life without taking medicines.

Follow-up questionnaires also included one measure of perceived positive change attributed to the program. That scale comprised 7 items, which were developed from conversations with patients about the effects of the program. The possible range of scores was from 0 to 28, and higher scores indicated greater perceived benefits from the program (Cronbach's alpha = 0.88).

The first (i.e., the baseline) questionnaire also had items asking about age, schooling, civil status, diagnoses, etc. Copies of the questionnaires are in Appendices 3 and 4.

2.4 Study design and data collection

For this study, the target population was all people with chronic diseases who participated in a CDSMP workshops held in Japan between August 2006 and April 2010. Some family members of chronic-disease patients also participated in workshops, but they were not included in this study. This was a longitudinal cohort study in which data were collected four times over one year. Baseline data were collected before the first group-discussion session, and follow-up questionnaires were sent by postal mail 3 months, 6 months, and 12 months later (Figure 2). A self-addressed post-paid envelope was included. If a follow-up questionnaire was not returned within two weeks, a reminder postcard was sent. Data were also collected on the participants' attendance at each of the program's sessions.

2.5 Analysis

2.5.1 Definition of decay of impact

The terms "relapse", "attenuation", "decay of impact", and "backsliding" are relatively easy to understand. Nonetheless, it was necessary

to have a single operational definition. The operational definition used in the present study is new, because a search of the literature revealed no relevant previous studies. That is, the concept of decay of impact after self-management education has been discussed [55, 65, 67, 71, 73], but it has never before been operationally defined in terms of individuals' scores on scales such as those used in this study.

In the present study, for each outcome measure, a participant was categorized as having decay of impact on a given measure only if that participant's data met both of the following two conditions: (#1) the best value was better than the baseline value, that is, there was improvement after the baseline value was measured, and (#2) the best value was also better than the last measured value, that is, there was decay after the aforementioned improvement. This definition has three advantages: first, it is consistent with LW Green's illustration of decay of impact as improvement followed by deterioration (Figure 3B in [55], which is reproduced here as part "B" of Figure 1); second it is consistent with the common definition of relapse as "to go back into a previous condition or into a worse state after making an improvement" [85]; and third, it could be applied objectively to the data collected in this study, as described in the next section.

2.5.2 Estimating true scores and constructing confidence intervals

Before the operational definition of decay of impact could be applied, the first step was to compute an estimated true score (t') from each observed score (i.e., from the measured values of each outcome). This was done by a 8method of "shrinkage" that is well-known both in psychometrics [86, page 153] and more generally as Stein's estimator [87, 88]. For computing estimated true scores, the shrinkage factor is the reliability coefficient. The coefficients of

test-retest reliability reported by Lorig, et al. [Table 2.4 in reference 82] were used. According to Nunnally & Bernstein [89, page 259-260] and to Furr & Bacharach [86, page 153], that computation requires the mean score, the observed score, and the reliability coefficient, as follows:

t' = (reliability × (observed score – mean score)) + mean score.

As an example, consider scores on the scale measuring self-efficacy. The test-retest reliability for that scale was 0.89. For one of the participants (ID: N00023), the observed baseline score, best score, and last score were 22, 45, and 24, respectively. The corresponding mean scores were 32.21, 35.36, and 35.54, respectively. Thus,

$t'_{\text{baseline}} = (0.89 \times (22 - 32.21)) + 32.21 = 23.12$

$t'_{\text{best}} = (0.89 \times (45 - 35.36)) + 35.36 = 43.94$

and

$t'_{\text{last}} = (0.89 \times (24 - 35.54)) + 35.54 = 25.27.$

The values of t' were used in the computations for decay of impact and for the other patterns, as described below.

The second step was to construct ranges of true scores (t) that were most consistent with observed scores (x), that is, $\sigma_{t.x}$. According to equation 7-4 of Nunnally & Bernstein [86, page 259],

$$\sigma_{t.x} = \text{standard deviation} \times \sqrt{\text{reliability} \times (1 - \text{reliability})}$$

Continuing with the self-efficacy scale as an example, the standard deviations at the times of the baseline, best, and last scores were 12.17, 12.08, and 12.64, respectively. Thus,

$\sigma_{t.x}$ at baseline = $12.17 \times \sqrt{0.89 \times (1 - 0.89)} = 3.81$

$\sigma_{t.x}$ at the time of the best score = $12.08 \times \sqrt{0.89 \times (1 - 0.89)} = 3.78$

and

$\sigma_{t.x}$ at the time of the last score = $12.64 \times \sqrt{0.89 \times (1 - 0.89)} = 3.95$

The computed values of $\sigma_{t.x}$ were then used to make a confidence interval (CI) for each estimated true score (each t'): CI = $t' \pm \sigma_{t.x}$

In the present example

$CI_{baseline} = t' \pm 3.81 = 23.12 \pm 3.81$ = from 19.31 to 26.93

$CI_{best} = t' \pm 3.78 = 43.94 \pm 3.78$ = from 40.16 to 47.72

$CI_{last} = t' \pm 3.95 = 25.27 \pm 3.95$ = from 21.23 to 29.22

Scores were considered to be different only if their CIs did not overlap. In the present example, the CIs for the baseline score and the best score do not overlap, so those two values would be considered to be different. In this example, the best score was better than the baseline score, which fulfills criterion #1 for decay of impact. The CIs for the best score and the last score also do not overlap, so those two values also would be considered to be different. In this example, the last score was worse than the best score, which fulfills criterion #2 for decay of impact. Because those two criteria were met, the data in this example would be classified as an instance of decay of impact.

2.5.3 Definitions of other patterns

If the data did not meet those two criteria, that is, if there was no decay of impact in a given participant's data on a given measure, then the pattern was categorized as improvement, deterioration, or no change, as follows (using the estimated true scores and confidence intervals, as described above in section 2.5.2): "improvement" if the last measured value was better than the baseline value; "deterioration" if the last measured value was worse than the baseline value, and "no change" for all others. Because the reliabilities of

the measurements have been taken into account, the "no change" subgroup should be interpreted as "no reliable change." Examples of all four patterns are shown in Figures 3a through 3d.

2.5.4 Timing of decay of impact

For each outcome measure, the number and percentage of participants in whom decay of impact started at 3 months, and the number in whom it started at 6 months, were computed. The data used for those computations came from all of the participants who had decay of impact and also returned all 4 questionnaires.

2.5.5 Prevalence of decay of impact

The numbers of participants with each pattern were counted. Also, for each outcome measure, the percentage of the participants whose results on that measure were classified as "decay of impact" was computed. Those percentages for each of the three other patterns of change were also computed.

2.5.6 Magnitude of decay of impact

For each person with decay of impact and each outcome with decay of impact, the magnitude (amount, size) of the decay was measured. For each instance in which the definition of decay of impact was met, the magnitude of the decay was defined as the difference between the best value and the last value. In the example above, decay of impact = 43.94 – 25.27 = 18.67.

For comparisons between outcomes measured on different scales the value used was the actual decay as a percent of the maximum possible decay, that is, as a percent of the full-scale value. Two examples follow:

a. Possible scores on the scale measuring self-efficacy ranged from 0 to 60, so, continuing the example from above, a decay of 18.67 would be 31.1% of the full-scale value (18.67/60 = 0.311).

b. Possible scores on the scale measuring depression ranged from 0 to 21, so a decay of 6 would be 28.6% of the full-scale value (6/21 = 0.286).

2.5.7 Number of outcomes with decay of impact

For each participant the number of outcome measures with decay of impact was computed (minimum = 0, maximum = 8). A histogram and basic descriptive statistics were used to describe the distribution of the number of outcomes with decay.

2.5.8 Predictors of having decay of impact

2.5.8.1 Choosing the comparison subgroup

To predict which participants would have decay of impact, a decision about the comparison subgroup had to be made. That is, should the decay-of-impact subgroup be compared with the improvement subgroup, with the deterioration subgroup, with the no-change subgroup, or with a subgroup comprising all participants who do not have decay? That question was answered on the basis of the purpose of this health-education program, and also on practical grounds.

The purpose of health-education programs for people with chronic diseases is to prevent deterioration or, better still, to cause improvement. Decay of impact is, by definition, the reversal of improvement. Therefore, to understand what causes decay it is necessary to understand the difference between improvement that is continuous or maintained and improvement

that is transient. On that basis, the decay subgroup should be compared with the improvement subgroup.

Other considerations lead to the same conclusion. First, the decay subgroup was not compared with a subgroup comprising all other participants. That would have required ignoring all differences between, for example, the improvement subgroup and the deterioration subgroup, which is clearly counterintuitive and for which no justification could be found.

Second, simple inspection of the medians and interquartile ranges at baseline in the four pattern-defined subgroups revealed that for all outcomes the decay-of-impact subgroup had "worse" baseline values than the deterioration subgroup (for examples, see the results for health distress in Figure 4f and for depression in Figure 4h). That is, it was common for those who had decay of impact to have been worse at baseline than later, and it was common for those who deteriorated to have been better at baseline than later. This suggests that some of the apparent difference between those two subgroups may reflect regression to the mean, so comparisons between those two subgroups would be difficult to interpret. (See also Appendix 5.)

Third, the differences in the baseline medians between the improvement subgroup and the decay-of-impact subgroup were generally small, and for some outcomes they were nearly zero. Thus, any regression-to-the-mean artifact would be small or nonexistent for comparisons between the decay subgroup and the improvement subgroup.

For those reasons, in the analyses aimed at identifying predictors of decay of impact, the decay subgroup was compared with the improvement subgroup.

2.5.8.2 CART models to identify predictors

To identify good predictors of having decay of impact, the method of "classification and regression trees" was used [90] (the method is referred to as C&RT, CART, and CRT). In the present study the dependent variable was binary (decay vs. improvement), so the trees used were classification trees, not regression trees.

CART modeling is a type of recursive partitioning. For making prediction models, it has various potential advantages over other methods: It can be used when there are many possible predictors, it automatically separates relevant from irrelevant predictors, it is not adversely affected by outliers or by missing values, it has been found in many cases to be more accurate than logistic regression, and it is particularly well-suited to exploratory analyses (such as in the present study) because there is no need to pre-specify a statistical model. Basic introductions to CART with examples can be found in [91-93]. Detailed explanations can be found in [90, 92, 93].

When making a prediction model, it is reasonable to include independent variables on the basis of theory, and also on the basis of the results of previous research. However, there is no theory of the decay of impact in self-management education for people with chronic diseases. Nonetheless, the self-management program was designed on the basis of self-efficacy theory [18], so it was hypothesized that people with higher self-efficacy after the end of the program would be less likely to have decay of impact. Thus, self-efficacy measured at the 3-month follow-up was included as one of the independent variables. According to the theory of relapse-prevention after treatment of addiction [94], social support, coping skills, and self-efficacy are among the factors that maintain abstinence. Thus, measures

of both coping and self-efficacy are among the independent variables. In total, for the present analyses, 28 independent variables were entered: 4 socio-demographic variables, 11 clinical variables, 11 baseline values, and 2 values measured at the 3-month follow-up (which are self-efficacy at 3 months and perceived positive change).

As with linear regression models, logistic regression models, and other statistical models, CART models can "overfit" the data if there are too many independent variables, which results in loss of generalizability. With CART models, the method used to prevent overfitting is called "pruning." Following the procedure suggested by Hastie at al. [95, page 308], in this study the trees were first allowed to grow to be very large (maximum depth: 5; minimum of 5 cases per parent node and 2 per child node), and then the trees were pruned in order to prevent overfitting. In this context, "pruning" means removing some nodes from the bottom up. Nodes were removed to make the smallest tree satisfying the "1 SE rule" that was suggested by Breiman et al. [90, pages 79-80]. According to that rule, 1 standard error is the maximum difference between the misclassification risk of the pruned tree and the misclassification risk of the tree with the smallest risk of misclassification.

To evaluate the CART models, three indices were used: the risk of misclassification, the percentage of participants who were correctly classified as having decay of impact, and the area under the receiver operating characteristic (ROC) curve [93, 95, 96]. Better models have lower risks of misclassification, higher percentages of participants correctly classified as having decay of impact, and larger ROC areas.

2.5.9 Statistical software

The data were analyzed with IBM SPSS version 20.

3. Results

3.1 Participants and data collected

Baseline data were obtained from 479 participants. Decay of impact can be detected only if a variable is measured at three or more times: once at baseline and at least twice thereafter. Therefore, data from people who returned the baseline questionnaire and at least two of the follow-up questionnaires were used (n = 364; 77.1% of 479, Table 1). Basic information about the scores at baseline are shown in Table 2.

As shown in Table 3, many of the participants were middle-aged, and almost 80% of them were women. About two thirds of the participants had finished college, and about half were married and living with a spouse. The length of time since the diagnosis of their chronic disease varied widely, from less than one year to more than sixty years, with a mean of 14 years. More than 40% of the participants had more than one diagnosis, and more than 15% of them had more than two diagnoses. The variety of diagnoses was very wide. The most common diagnosis was allergic disease, while cardiovascular disease, connective tissue disease, diabetes, and rheumatic disease were also relatively common. The 10 most common diagnoses are shown in Table 3. A complete list of the other diagnoses is in Appendix 6.

3.2 Basic findings regarding the four patterns of change

For each outcome measure, Figure 4 shows the whole-group-summary and subgroup-summary changes over time in boxplots. The differences between the patterns in different subgroups can be seen clearly. All four patterns can be seen and they can be easily distinguished from each other.

For some outcomes, there appear to be no changes at all at the whole-group level, but the pattern-defined subgroups with improvement, decay of impact, no change, and deterioration are easy to identify. This is particularly noticeable with regard to pain (Figure 4b).

3.3 Prevalence of decay of impact

The percentages of participants who had each pattern of change are shown in Figure 5 for each of the 8 measures. Depending on the outcome measure, from 7% to 26% of the participants had decay of impact. The highest percentage was on the scale measuring self-rated health status, and the lowest percentage was on the scale measuring pain.

3.4 Magnitude of decay of impact

Distributions of the individual magnitudes of decay of impact as percentages of each measure's full-scale value are shown as box-and-whisker plots in Figure 6. Overall, decay of impact was greater on the measures of general health status than on the measures of self-management behavior or psychological health. The median magnitudes of the decay ranged from 16.4% of full scale for depression to 39.5% of full scale for pain. However, all of the distributions were right-skewed: some people had more than 50% decay, and some had more than 60% decay, on some measures. The coefficient of skewness ranged from 0.76 (standard error, 0.32) for communication to 1.60 (standard error, 0.32) for health distress (Table 4).

3.5 Number of outcomes with decay of impact

Among the 364 participants, 121 (36%) did not have decay on any of the 8 measures, while 117 participants (35%) had decay on at least 2 measures.

The distribution of the number of measures with decay of impact was right-skewed, with a median of 1 (Figure 7).

3.6 Timing of decay of impact

The individual starting times for decay of impact did not cluster strongly either at 3 months or at 6 months. Specifically, the percentage in whom the decay began at 3 months ranged from 26.1% to 61.4% and, correspondingly, the percentage in whom it began at 6 months ranged from 38.6% to 73.9%. On all outcome measures except self-rated health, in more than half of the participants who had decay of impact that decay began 6 months after the baseline measurement (Table 5).

3.7 Patterns of change, by diagnosis

There is no consistent, clearly discernible pattern among the results shown in Tables 6a through 6e. That is, none of the 4 subgroups was clearly associated with diagnosis, and none was clearly associated with the number of diagnoses. For all 8 outcomes, the magnitude of decay of impact was also not clearly associated either with diagnosis or with the number of diagnoses.

3.8 Predictors of having decay of impact

Results regarding the CART models as predictors of having decay of impact are summarized in Tables 7a and 7b (see also Appendix 7), and the classification trees are shown in Figures 8a through 8h. The CART models were either good or very good as classifiers. In general, the risks of misclassification were low, the percentages of participants who were correctly classified with decay of impact were high, and the areas under the ROC curves were high. The best CART models were those for predicting decay on coping, on anxiety, and on self-rated health.

In the final CART models, diagnoses were generally not included as predictors. The exceptions were fibromyalgia syndrome and Parkinson's disease. People with fibromyalgia syndrome were more likely to have decay on the self-rated health scale, and people with Parkinson's disease were more likely to have decay on the pain scale.

The most consistent single predictor was the number of years since diagnosis. That predictor was included in six of the eight CART models. In general, participants with longer disease histories were predicted to have decay of impact (Tables 7a and 7b and Figures 8a through 8h). The only exception was communication with medical doctors (Figure 8c). Participants with longer disease histories were predicted have improvement rather than decay of impact in their communication with medical doctors.

4. Discussion

4.1 Summary of the main findings

Answering the two questions at the end of section 1.10 (above), this study provides information about the prevalence, magnitude, and timing of the decay of impact. It also shows one way of predicting who will have decay, and thus who will need reinforcement.

First, on all outcomes except pain, more than 10% of the participants had decay of impact. Decay was most prevalent on self-rated health (26%), coping (20%), and communication (15%). The magnitude of the decay varied among outcomes, with medians of about 16% to 40% of the full-scale values, and the inter-individual variation was large. Regarding when the reinforcement is needed, in about 30% to 60% of the participants the decay began 3 months after the program started. That is, reinforcements are needed approximately 6 weeks after the program ends (Table 5).

Second, the best overall predictor of the need for reinforcement was the number of years since diagnosis. In general, diagnoses were not good predictors of having decay of impact. For self-rated health, coping, and anxiety, the CART models were good predictors of having decay of impact. Those models correctly classified 71%, 82%, and 77%, respectively, of the participants who needed reinforcement.

4.2 Patterns of change, in the context of previous work

<u>Patterns of change and the need for reinforcement:</u>

At least since 1977, when LW Green included the concept of decay of impact in his discussion of the problems encountered when evaluating health education [55], it has been recognized as a possible cause of both over-

estimation and under-estimation of a program's effects. Nonetheless, as a research topic it has received very little attention.

Caplin, et al. [73] studied 53 patients who had completed a self-management program, and they classified 20 of those patients as "relapsers." That study differed from the present one in at least four important ways: (1) all of its participants had asthma, (2) the self-management program was disease-specific, (3) the study had one follow-up measurement six years after the end of the program (rather than three follow-up measurements during one year after the program), and (4) relapse was defined partly on the basis of self-report (rather than by changes on psychometric scales). One of the aims of the present study (to determine when decay begins) is similar to the aim of the study by Hennessy, et al. [75], but in that study a different method was used, and the population and intervention were also different. A search of the literature revealed no previous systematic analysis of decay after self-management education for people with various chronic diseases. In the present study, the analysis of decay at the level of individuals also made it possible to measure the magnitude of decay and to compare that magnitude among various outcomes (psychological and behavioral).

This is the first study to propose a single answer to two important questions: Why are the effects of these programs generally small, and restricted to only a few of the outcomes measured? [22] And, why have reinforcement programs not succeeded? [29, 30, 68, 79, 80] The proposed answer is that using whole-group summary statistics only, i.e. not analyzing pattern-defined subgroups, results in *dilution*. Specifically, in answer to the first question, previous studies have focused on whole-group analyses, in which the benefits to some participants are diluted by the lack of benefits to

others. Regarding the second question, the present results show that studies of the effects of reinforcements may have included many people who did not need reinforcements, so the benefits of reinforcements to the people who need them were diluted by the lack of benefit to people who do not need them. Any success of reinforcement programs among people with decay of impact would be diluted by the lack of effect of such programs among people who maintain their benefit or continue to improve during the follow-up period even without reinforcement. The solution, therefore, is to focus more attention on pattern-defined subgroups.

As mentioned above in section 1.9, only a few reinforcement programs have been studied, and the results have generally not been positive [29, 30, 68, 79, 80]. The present study provides a first step toward a clear and accurate understanding of the phenomenon that reinforcement programs ("booster sessions") are intended to prevent or reverse. The present findings also may help to explain why some of the reinforcement programs that have already been studied have been only slightly effective or ineffective. Specifically, one of the main findings of this study is that decay of impact can occur in up to 25% of the participants. However, equally important is the finding that many participants did *not* have decay. Because they did not have decay, those people would be expected to benefit only little, or not at all, from reinforcement. Therefore, when a reinforcement is found to have no benefits, or to have benefits that are small, or that are limited to only a few of the measured outcomes, the reason may be that only a subgroup in fact needed the reinforcement. Reinforcements may appear to be ineffective because they are actually needed by fewer than half of the people who receive them. In

whole-group statistics, the benefits of reinforcements to the people who need them are masked by their lack of benefit to people who do not.

Previous studies of subgroups:

This is not the first study of changes after a self-management program in subgroups. Lorig, et al. [49] analyzed the effects of such a program by diagnosis-defined subgroup, and found only small differences. More recently, Franks, et al. [65] found that a self-management program was more effective in subgroups defined by certain personality characteristics, and Smeulders et al. [64] found that a self-management program was more effective in a subgroup with a low educational background. Reeves et al. [63] found differences in the effects of a self-management program between subgroups defined by age and by baseline self-efficacy. Ritter et al. [54] found no important differences in the program's effects between subgroups defined by baseline status on demographics and health-related variables. However, in those studies none of the subgroups were defined by their pattern of change.

There have been two studies in which subgroups were defined by their pattern of change. First, Caplin, et al. [73] studied 53 people with asthma 6 years after they had completed a self-management program, and they classified 20 of them as "relapsers". As noted above, that study was very different from the present one. Still, it is interesting that one of the factors associated with relapse was a lack of "commitment to strengthen their self-efficacy." Second, Nolte et al. [62] found that improvement was more common among younger participants, and particularly among younger women. In Nolte et al.'s study there was only one measurement after the baseline measurement, and therefore people with decay of impact could not be identified.

Other patterns:

There might be important patterns of change over time other than the four defined in this study. For example, people whose scores start high and stay high might differ in important ways from people whose scores start low and stay low [97]. That is beyond the scope of this study, but it is a worthwhile topic for future work.

4.3 Timing of decay, in the context of previous work

Because no previous studies have focused on the decay of impact after this type of educational program, it is not possible to directly compare the timing of decay found here with previous findings. However, some studies have been done in related areas.

Caplin et al. [73] identified a subgroup that they called "relapsers." However, their measurements were not made at multiple times after the baseline data were collected, so from their results it is not clear when the relapse began.

Krebs, et al. [74] meta-analyzed a total of 88 studies of computer-tailored interventions intended to influence various health-related behaviors (smoking cessation, dietary fat reduction, increasing fruit and vegetable intake, physical activity, and mammography screening). They found large variation between studies, but overall there was a decay of impact that began 4 to 6 months after the baseline measurements [Figure 1 in 74], which is generally consistent with the present results.

In a study of an intervention to increase self-efficacy for condom use to prevent HIV infection, the aim of Hennessy et al. [75] was to determine when boosters should be given. They found that the decay of impact began less than 3 months after the end of the initial intervention. That finding is also

consistent with the present findings, and it strengthens the conclusion that reinforcements should start early.

4.4 Predictors of decay, in the context of previous work

To predict who will have decay of impact and will therefore need reinforcement, at least three approaches are possible: theory-based, empirical, and mixed. The approach used in the present study could be classified as mixed. As mentioned above (in section 2.5.8.2) the use of self-efficacy and of coping as independent variables has theoretical justification. Nonetheless, much of this study is exploratory, observational, and empirical. That is, because there was no pre-existing framework to explain the decay of impact in this context, many independent variables were analyzed to derive the CART models. In general, this was successful. The classification trees worked well as predictors even after they had been pruned to avoid overfitting. The risk of misclassification was less than 0.25 for 5 of the 8 outcomes. Still, more work is needed to test those prediction trees in other settings and with other populations.

A conceptual framework that acknowledges the importance of self-efficacy, such as the theory of relapse prevention after treatment of addiction [77, 94] or the model proposed by LW Green and MW Kreuter [98, pages 160-161], might be adapted to fit the decay of impact after education for chronic disease self-management. Because this educational program was founded and developed on the basis of self-efficacy theory [18], it would be reasonable to expect scores on the self-efficacy scale at 3 months to be good predictors of having decay of impact. In fact, as shown in Tables 7a and 7b those scores were included in the prediction models for only 4 of the 8 outcome measures (self-rated health, coping, health distress, and anxiety). These results seem to

indicate that a high level of self-efficacy after the program is not necessary for all of the benefits of this program to be maintained. That finding need not contradict the idea that self-efficacy is important as a determinant of changes in health-related behavior, but it does raise the possibility that change and maintenance are mediated by different factors, which is consistent with the general theory proposed by Rothman [99]. According to Rothman, "Decisions regarding behavioral initiation are predicted to depend on favorable expectations regarding future outcomes, whereas decisions regarding behavioral maintenance are predicted to depend on perceived satisfaction with received outcomes." That theory should be operationalized in the context of the patterns of change after self-management education for people with chronic diseases, and it should be tested in that context.

One noteworthy finding regarding predictors is that, on 5 of the 8 outcome measures, participants with longer diseases histories were more likely to have decay of impact. The reasons for this finding are not clear, but it does imply that people with longer diseases histories are more likely to need reinforcements.

On the scale measuring communication with medical doctors, participants with longer diseases histories were less likely to have decay of impact. In this context it may be worth noting that older participants were also less likely to be completely lost to follow-up after this program [100]. It is possible that people with more experience as patients might be more motivated to maintain their new self-management skills.

In the CART models, people with fibromyalgia syndrome were predicted to have decay of impact on self-rated health, and people with Parkinson's disease were predicted to have decay on pain. However, none of

the other diagnoses were included in the prediction models for any of the outcomes. Multimorbidity was also not included in the prediction models for any of the outcomes. One possible explanation is that some of the diagnosis groups were small, which would limit the ability to detect their effects. Also, multimorbidity was common, which would make the effects of any single diagnosis more difficult to detect. Another possibility is that the causes of decay of impact actually have little or no relationship with diagnoses. The latter interpretation is consistent with previous studies showing that the benefits of the program are not related to diagnoses [49, 54].

4.5 Limitations

a. In this study the assignment to a diagnosis category, and the numbers of diagnoses among participants with multimorbidity, were based on self-reported diagnoses. A few studies have found inaccuracy of self-reported diagnosis [101-103]. Many others have found that the accuracy of self-reported diagnosis varies by disease, and that, for many chronic conditions, self-reported diagnosis is accurate enough to be used in survey research in Japan [104] and in other countries [105-119]. Nonetheless, when participants are categorized by diagnosis those categorizations should be done with reference to medical records whenever possible, which might improve the predictive value of the classification-tree models.
b. One possible explanation of any apparent improvement or deterioration is response shift [120, 121]. The concept of response shift was developed in the context of quality-of-life measurement, and it has been applied to outcomes of programs such as this one [122, 123]. However, for response shift to account for the apparent decays of

impact in the present study it would have to occur in opposite directions sequentially, and it would have to do so not only in measures of psychological health but also in reports of frequencies of health-related behaviors.

c. Regarding the timing of events, having measurements at only four time points over one year limits the ability to identify the time at which decay of impact began. If more points had been available on the time axis, then the time of the start of the decay could have been identified more precisely. For example, if the interval before the first measurement is shorter than 3 months, then some people might be found to have decay starting earlier than 3 months after the baseline measurement. If data were available not only at 3 months, but also 2 months after the start of the program, and if some participants had decay starting at 2 months, then the overall percentage of participants classified as having decay of impact would be larger. Those participants who meet the criteria for decay of impact (section 2.5.1 above) at 3 months, 6 months, or both, would still meet those criteria, and to their number would be added the participants with decay starting at 2 months. Therefore, the results reported in Figure 5 (and elsewhere) might be biased toward low apparent prevalences of decay of impact. That is, the prevalences of decay of impact reported here should be considered to be lower boundaries, and the actual prevalences are likely to be higher. The existence of decay of impact starting earlier than 3 months after the start of the program would also have practical implications. Specifically, it would mean that reinforcement programs should start almost immediately after the

main program ends. Furthermore, the CDSMP was designed to last 6 weeks, but in fact the ideal length of the program is not known. If decay of impact starts less than 3 months after the start of the main program, then one option would be to extend the main program beyond 6 weeks. As another practical point, these considerations also show that more frequent monitoring of outcomes is important. If a small number of particularly important outcome indices can be identified, then even daily monitoring would be practical [124]. Also, of course, at least some of the participants whose patterns were categorized as "improving" might have been categorized as having decay if the study had lasted for more than 12 months. In this study, "long-term" follow up is follow-up lasting for 12 months after the baseline measurement. Follow-up beyond 12 months might reveal other patterns. For example, some of the people who had decay might "reverse" again and begin improving.

d. This study is, in many ways, descriptive and exploratory, because the topic is new. In fields with essentially no previous research it is natural to start with exploratory work.
e. Another possible limitation is the fact that many more women than men were in this study. Although the present results might not apply to a program with a much smaller percentage of women, such programs seem to be rare, while programs with many more women than men are typical. In 17 studies of programs such as this one (i.e., focusing on self-management of chronic illness [57, 20, 21, 23, 27, 28, 60, 125-133], the percentage of women participants ranged from 61.1%

[129] to 88.9% [130] and the mean was 75.7%. In the present study it was 79.1%.

f. One limitation concerns the operational definitions of the four patterns of change. Those definitions depend on each measure's coefficient of test-retest reliability. Those coefficients of test-retest reliability are based on measurements made in a previous study that was not done in Japan, and the measurements were separated by 10 days [page 18 and Table 2.4 of reference 82]. That time period (10 days) is typical for assessments of test-retest reliability. However, the measurements in the present study were separated not by 10 days but by 3-12 months. As noted by DeVellis [134, page 52], test-retest reliability coefficients "tell us about the measure only when we are highly confident that the phenomenon has remained stable. Such confidence is not often warranted." For example, even if we assume that the actual self-efficacy remained stable over 10 days, we cannot be confident that it would have remained stable over 3-12 months. A 12-month study to measure test-retest reliability would definitely be impractical, and it might even be impossible because it would depend on the true value of the measured variable not changing over those 12 months. Limitations like this are unavoidable in almost all long-term follow-up studies.

5. Conclusions and recommendations

This study found that decay of impact occurred in 7% to 26% of the participants. Among those who do have decay of impact, it can start as early as 3 months after the baseline measurement, i.e. 6 weeks after the end of the program. The magnitude of decay ranged from 16.4% to 39.5% of the full-scale value. Decay occurred in all outcomes, and it was most common in self-rated health.

The CART models for self-rated health, coping, and anxiety were different from each other, but each one gave good predictions of who would have decay, and therefore who would need reinforcement. In general, diagnoses were not included in the CART models, which means that in most cases diagnoses could not be used to predict the need for reinforcement. Regarding self-rated health, coping, self-efficacy, health distress, and depression, people with longer disease histories were predicted to have decay, and so they need reinforcement.

The findings reported from many previous studies, including randomized controlled trials, are likely to be wrong. Specifically, the existence of subgroups that can be distinguished by their patterns of change can explain why randomized controlled trials have found that these programs have only small effects, and why reinforcements generally seem to be unsuccessful. The effects in whole groups appear to be small and isolated because of dilution. More attention must be given to pattern-defined subgroups.

As part of that extra attention to subgroups, patterns other than decay of impact should be studied. Although four patterns of change after the CDSMP were identified, the analyses of predictors focused on the differences between participants who had decay of impact and those who had

improvement only. In future studies, predictors of the other two patterns should also be identified.

There is a need for external validation and replication, to determine the extent to which these findings can be generalized to other CDSMP participants. There is also a need for qualitative studies, including in-depth interviews with participants who have decay of impact and with those who do not. The interviews could include questions about social support, self-efficacy, and satisfaction with changes brought on by the program. In those interviews it will be important to find out about how people avoid decay or overcome factors that could cause decay. That is, it will be important to investigate resilience among people with chronic illness.

Future studies might identify other good predictors of decay if those studies were based on relevant theories. For example, Hendershot's conceptual model of relapse prevention [94] implies that social support is important, so it should be measured in future studies. Also, in Rothman's theory of behavioral maintenance "perceived satisfaction with received outcomes" [99] is important, so that type of satisfaction should also be measured. In this context it is also important to note that future studies should include measures of satisfaction with the program, in addition to measures of response shift.

Another recommendation is to use patterns of change to evaluate programs. For example, in future studies a successful program might be defined as one in which very few of the participants eventually have decay of impact, or as one in which that decay is small.

A recommendation can also be made regarding studies of educational programs in general. The results presented here show how classification-tree

models might also be applied in other contexts. The software needed for generating and evaluating tree models is now easy to obtain and use, and health-education researchers often need to study many possible predictors with little or no theory to guide the construction of a linear or logistic model. It should be remembered that the predictors actually identified are likely to differ from those found in the present study, but nonetheless classification and regression trees may be useful in those situations.

Finally, some recommendations can also be made regarding practice. Specifically, based on the results of this study (and consistent with the results of Hennessy [75]) it is clear that many people who have decay need reinforcement very soon after the main program ends. In the case of the CDSMP, "very soon" means as early as 3 months after the start of the main program. Also, if the resources needed to implement reinforcement programs are limited, then the classification trees (especially the trees for self-rated health, coping, and anxiety) can be used to predict which of the participants will have decay of impact. The program administrators can then focus their limited resources on the participants who are likely to need them most.

6. Acknowledgements

Funding for this study was provided by the Japanese Ministry of Health, Labor and Welfare aids for Scientific Research (health disease prevention, treatment and research projects for Allergy and Immunology).

I am very grateful to all of the participants for cooperating in this study, to Yoshihiko Yamazaki and all the members of the self-management research team at the University of Tokyo for their advice and collaboration, to the Japan Chronic Disease Self-Management Association for implementing the program and for allowing the data to be collected and used, to Noriko Okamoto for administrative and technical support, and to Joseph Green for advice and for language editing. I am also very grateful to all of the members of the dissertation committee for their extremely helpful comments and suggestions, and to Professors Takahiro Kiuchi and Hirono Ishikawa for their valuable advice and academic support.

7. References

1. Yach D, Hawkes C, Gould L, Hofman KJ: The global burden of chronic diseases: Overcoming impediments to prevention and control. Journal of the American Medical Association. 2004;**291**:2616-2622.

2. Hughes BB, Kuhn R, Peterson CM, Rothman DS, Solórzano JR, Mathers CD, Dickson JR. Projections of global health outcomes from 2005 to 2060 using the International Futures integrated forecasting model. Bull World Health Organization. 2011;**89**:478–486.

3. Beaglehole, R. & Bonita, R. Global public health: A scorecard. The Lancet. 2008;**372(9654):**1988-1996.

4. World Health Organization. Noncommunicable diseases country profiles 2011. http://whqlibdoc.who.int/publications/2011/9789241502283_eng.pdf

5. Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. Journal of the American Medical Association. 2002;**288(19)**:2469-2475.

6. World Health Organization. Global status report on noncommunicable diseases 2010; 2011. ISBN 978 92 4 068645 8 (PDF).

7. Van Olmen J, Ku GM, Bermejo R, Kegels G, Hermann K, Van Damme W. The growing caseload of chronic life-long conditions calls for a move towards full self-management in low-income countries. Global Health. 2011;**7(1):**38.

8. Huber M, Knottnerus JA, Green L, van der Horst H, Jadad AR, Kromhout D, Leonard B, Lorig K, Loureiro MI, van der Meer JW, Schnabel P, Smith R, van Weel C, Smid H. How should we define health? The British Medical Journal. 2011;**343**:d4163.

9. Coleman MT, Newton KS. Supporting Self-management in Patients with Chronic Illness. American Family Physician. 2005;**72**:1503-1510.

10. Lorig K, Lubeck D, Kraines RG, Seleznick M, Holman HR. Outcomes of self-help education for patients with arthritis. Arthritis & Rheumatism. 1985;**28(6)**:680-685.

11. Arthritis Self-Management Program. <http://patienteducation.stanford.edu/programs/asmp.html> Accessed on 2012.02.11.

12. Lorig K, Seleznick M, Lubeck D, Ung E, Chastain RL, Holman HR. The beneficial outcomes of the arthritis self-management course are not adequately explained by behavior change. Arthritis & Rheumatism. 1989;**32(1)**:91-95.

13. Lenker SL, Lorig K, Gallagher D. Reasons for the lack of association between changes in health behavior and improved health status: an exploratory study. Patient Education and Counseling. 1984;**6(2)**:69-72.

14. Bandura, A. Self-efficacy: Toward a unifying theory of behavioral change. Psychological Review. 1977;**84(2)**:191-215.

15. Bandura A (Editor). Self-Efficacy in Changing Societies. Cambridge University Press; 1997.

16. Lorig K, Chastain RL, Ung E, Shoor S, Holman HR. Development and evaluation of a scale to measure perceived self-efficacy in people with arthritis. Arthritis & Rheumatism. 1989;**32(1)**:37-44.

17. Lorig K, González V. The integration of theory with practice: a 12-year case study. Health Education Quarterly. 1992;**19(3)**:355-368.

18. Rosenstock IM, Strecher VJ, Becker MH. Social Learning Theory and the Health Belief Model. Health Education & Behavior (Health Education Quarterly). 1988;**15(2)**:175-183.

19. Japan Chronic Disease Self-Management Association. <http://www.j-cdsm.org/program/cContents.html> Accessed on 2012.02.12.

20. Lorig K, Sobel D, Ritter P, Laurent DD, Hobbs M. Effect of a Self-Management Program on Patients with Chronic Disease. Effective Clinical Practice. 2001;**4(6)**:256-262.

21. Damush TM, Weinberger M, Perkins SM, Rao JK, Tierney WM, Qi R, Clark D. The Long-term Effects of a Self-management Program for Inner-city Primary Care Patients with Acute Low Back Pain. Archives of Internal Medicine. 2003;**163**: 2632-2638.

22. Foster G, Taylor SJC, Eldridge S, Ramsay J, Griffiths CJ. Self-management education programmes by lay leaders for people with chronic conditions. Cochrane Database of Systematic Reviews.
2007;Issue 4. Art. No.: CD005108.

23. Fu Dongbo, Fu Hua, McGowan, P., Shen Yi-e, Zhu Lizhen, Yang Huiqin, Mao Jianguo, Zhu Shitai, Ding Yongming, Wei Zhihua. Implementation

and quantitative evaluation of chronic disease self-management programme in Shanghai, China: randomized controlled trial. Bulletin of the World Health Organization. 2003;**81(3)**:174-182.

24. Lorig K, Ritter P, Stewart A, Sobel DS, Brown BW, Andura A, Gonzalex VM, Laurend DD, Homan HR. Chronic Disease Self-Management Program: 2-Year Health Status and Health Care Utilization Outcomes. Medical Care. 2001;**39(11)**:1217-1223.

25. Chan SC, Siu AM, Poon PK, Chan CC. Chronic disease self-management program for Chinese patients: a preliminary multi-baseline study. International Journal of Rehabilitation Research. 2005;**28**:351-354.

26. Smeulders E, Haastregt J, Dijkman-Domanska B, Hoef E, Eijk J, Kempen G. Nurse-and peer-led self-management programme for patients with an implantable cardioverter defibrillator; a feasibility study. BioMed Central Nursing. 2007;**6(6)**:1-8.

27. Lorig K, Ritter P, Jacquez A. Outcomes of Border Health Spanish/English Chronic Disease Self-management Programs. The Diabetes Educator. 2005;**31**:401-409.

28. Swerissen H, Belfrage J, Weeks A, Jordan L, Walker C, Furler J, McAvoy B, Carter M, Peterson C. A randomised control trial of a self-management program for people with a chronic illness from Vietnamese, Chinese, Italian and Greek backgrounds. Patient Education & Counseling. 2006;**64(1-3)**:360-368.

29. Lorig K, Ritter P, Villa F, Piette JD. Spanish diabetes self-management with

and without automated telephone reinforcement: two randomized trials. Diabetes Care. 2008;**31(3)**:408-414.

30. Lorig K, Ritter P, Laurent DD, Plant K, Green M, Jernigan VB, Case S. Online diabetes self-management program: a randomized study. Diabetes Care. 2010;**33(6)**:1275-1281.

31. Gibson PG, Powell H, Coughlan J, Wilson AJ, Abramson M, Haywood P, Bauman A, Hensley MJ, Walters EH. Self-management education and regular practitioner review for adults with asthma. Cochrane Database of Systematic Reviews. 2003;**(1)**:CD001117.

32. Effing T, Monninkhof EM, Valk PD, Palen J, Herwaarden CL, Partidge MR, Walters EH, Zielhuis GA. Self-management education for patients with chronic obstructive pulmonary disease. Cochrane Database of Systematic Reviews. 2007;**(4)**:CD002990.

33. Barlow C, Cooke D, Mulligan K, Beck E, Newman S. A critical review of self-management and educational interventions in inflammatory bowel disease. Gastroenterology Nursing. 2010;**33(1)**:11-18.

34. Boren SA, Wakefield BJ, Gunlock TL, Wakefield DS. Heart failure self-management education: a systematic review of the evidence. International Journal of Evidence-Based Healthcare. 2009;**7(3)**:159-168.

35. Otsu H, Moriyama M. Effectiveness of an educational self-management program for outpatients with chronic heart failure. Japan Journal of Nursing Science. 2011;**8(2)**:140-152.

36. Brody BL, Roch-Levecq AC, Kaplan RM, Moutier CY, Brown SI. Age-related macular degeneration: self-management and reduction of depressive symptoms in a randomized, controlled study. Journal of the American Geriatric Society. 2006;**54(10)**:1557-1562.

37. Rees G, Keeffe JE, Hassell J, Larizza M, Lamoureux E. A self-management program for low vision: program overview and pilot evaluation. Disability and Rehabilitation. 2010;**32(10)**:808-815.

38. Gifford AL, Laurent DD, Gonzales VM, Chesney MA, Lorig KR. Pilot randomized trial of education to improve self-management skills of men with symptomatic HIV/AIDS. Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology. 1998;**18(2)**:136-144.

39. LeFort SM, Gray-Donald K, Rowat KM, Jeans ME. Randomized controlled trial of a community-based psychoeducation program for the self-management of chronic pain. Pain. 1998;**74(2-3)**:297-306.

40. Francis KL, Matthews BL, Van Mechelen W, Bennell KL, Osborne RH. Effectiveness of a community-based osteoporosis education and self-management course: a wait list controlled trial. Osteoporosis International. 2009;**20(9)**:1563-1570.

41. McGillion MH, Watt-Watson J, Stevens B, Lefort SM, Coyte P, Graham A. Randomized controlled trial of a psychoeducation program for the self-management of chronic cardiac pain. Journal of Pain and Symptom Management. 2008;**36(2)**:126-140.

42. Vogeli C, Shields AE, Lee TA, Gibson TB, Marder WD, Weiss KB, Blumenthal D. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. Journal of General Internal Medicine. 2007;**22(3)**:391-395.

43. International Research Community on Multimorbidity <http://pages.usherbrooke.ca/crmcspl-blog/> Accessed on 2012.02.12.

44. Boyd CM, Fortin M. Future of Multimorbidity research: How Should Understanding of Multimorbidity Inform Health System Design? Public Health Reviews. 2010;**32(2)**:451-474.

45. Berger V. W. Generalizability. Encyclopedia of Statistics in Behavioral Science. John Wiley & Sons, Ltd.; 2005, Volume 2, p 702.

46. Ellenberg JH. Selection bias in observational and experimental studies. Statistics in Medicine. 1994;**13(5-7)**:557-567.

47. Clark NM, Becker MH, Janz NK, Lorig K, Rakowski W, Anderson L. Self-management of chronic disease by older adults. A review and questions for older adults. Journal of Aging and Health. 1991;**3**:3-27.

48. Barlow JH, Sturt J, Hearnshaw H. Self-management interventions for people with chronic conditions in primary care: examples from arthritis, asthma and diabetes. Health Education Journal. 2002;**61(4)**:365–378.

49. Lorig KR, Sobel DS, Stewart AL, Brown BW, Bandura A, Ritter P, Gonzalez VM, Laurent DD, Holman HR: Evidence Suggesting That a Chronic Disease Self-Management Program Can Improve Health Status

While Reducing Hospitalization: A Randomized Trial. Medical Care. 1999;**37**:5-14.

50. Internet homepage of the Chronic Disease Self-Management Program (accessed on 2011.10.04): <http://patienteducation.stanford.edu/programs/cdsmp.html>.

51. ケイト・ローリッグ他　著（近藤 房恵　訳）病気とともに生きる ― 慢性疾患のセルフマネジメント　**2008**　（株）日本看護協会出版社

52. Lorig K, Ritter PL, Plant K. A disease-specific self-help program compared with a generalized chronic disease self-help program for arthritis patients. Arthritis & Rheumatism. 2005;**53(6)**:950-957.

53. Goeppinger J, Armstrong B, Schwartz T, Ensley D, Brady TJ. Self-management education for persons with arthritis: Managing comorbidity and eliminating health disparities. Arthritis & Rheumatism. 2007;**57(6)**:1081-1088.

54. Ritter PL, Lee J, Lorig K. Moderators of chronic disease self-management programs: who benefits? Chronic Illness. 2011;**7(2)**:162-172.

55. Green LW. Evaluation and measurement: some dilemmas for health education. American Journal of Public Health. 1977;**67(2)**:155-161.

56. Cohen J. Statistical power analysis for the behavioral sciences. 2nd edition. Lawrence Erlbaum Associates; 1988.

57. Gitlin LN, Chernett NL, Harris LF, Palmer D, Hopkins P, Dennis MP. Harvest health: translation of the chronic disease self-management

program for older African Americans in a senior setting. Gerontologist. 2008;**48(5)**:698-705.

58. Siu AM, Chan CC, Poon PK, Chui DY, Chan SC. Evaluation of the chronic disease self-management program in a Chinese population. Patient Education and Counseling. 2007;**65(1)**:42-50.

59. Barlow J, Turner A, Edwards R, Gilchrist M. A randomised controlled trial of lay-led self-management for people with multiple sclerosis. Patient Education and Counseling. 2009;**77(1)**:81-89.

60. Lorig KR, Ritter P, Laurent DD, Plant K: Internet-Based Chronic Disease Self-Management: A Randomized Trial. Medical Care. 2006;**44**: 964-971.

61. Jerant A, Moore-Hill M, Franks P. Home-based, peer-led chronic illness self-management training: findings from a 1-year randomized controlled trial. The Annals of Family Medicine. 2009;**7(4)**:319-327.

62. Nolte S, Elsworth GR, Sinclair AJ, Osborne RH. The extent and breadth of benefits from participating in chronic disease self-management courses: a national patient-reported outcomes survey. Patient Education and Counseling. 2007;**65(3)**:351-360.

63. Reeves D, Kennedy A, Fullwood C, Bower P, Gardner C, Gately C, Lee V, Richardson G, Rogers A. Predicting who will benefit from an Expert Patients Programme self-management course. British Journal of General Practice. 2008;**58(548)**:198-203.

64. Smeulders ES, Haastregt JC, Ambergen T, Stoffers HE, Janssen-Boyne JJ, Uszko-Lencer NH, Gorgels AP, Lodewijks Bolt CL, Eijk JT, Kempen GI.

Heart failure patients with a lower educational level and better cognitive status benefit most from a self-management group programme. Patient Education and Counseling. 2010;**81(2)**:214-221.

65. Franks P, Chapman B, Duberstein P, Jerant A. Five factor model personality factors moderated the effects of an intervention to enhance chronic disease management self-efficacy. British Journal of Health Psychology. 2009;**14(3)**:473-487.

66. Swerissen H, Belfrage J, Weeks A, Jordan L, Walker C, Furler J, McAvoy B, Carter M, Peterson C. A randomised control trial of a self-management program for people with a chronic illness from Vietnamese, Chinese, Italian and Greek backgrounds. Patient Education and Counseling. 2006;**64(1-3)**:360-368.

67. Barlow J, Turner A, Swaby L, Gilchrist M, Wright C, Doherty M. An 8-yr follow-up of arthritis self-management programme participants. Rheumatology. 2009;**48(2)**:128-133.

68. Riemsma RP, Taal E, Rasker JJ. Group education for patients with rheumatoid arthritis and their partners. Arthritis & Rheumatism. 2003;**49(4)**:556-566.

69. Mulligan K, Newman S. Self-management interventions. In: Ayers S, Baum A, McManus C, Newman S, Wallston K, Weinman J, West R, editors. Cambridge handbook of psychology, health and medicine. 2nd edition. Cambridge University Press; 2007. p 393.

70. Barlow JH, Wright CC, Turner AP, Bancroft GV. A 12-month follow-up

study of self-management training for people with chronic disease: Are changes maintained over time? British Journal of Health Psychology. 2005;**10**:589–599.

71. Lorig KR, Ritter PL, Laurent DD, Fries JF. Long-term randomized controlled trials of tailored-print and small-group arthritis self-management interventions. Medical Care. 2004;**42(4)**:346-354.

72. Clark NM. Management of chronic disease by patients. Annual Review Public Health. 2003;**24**:289-313.

73. Caplin DL, Creer TL. A self-management program for adult asthma. III. Maintenance and relapse of skills. Journal of Asthma. 2001;**38(4)**:343-356.

74. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. Preventive Medicine. 2010;**51(3-4)**:214-221.

75. Hennessy M, Bolan GA, Hoxworth T, Iatesta M, Rhodes F, Zenilman JM. Using growth curves to determine the timing of booster sessions. Structural Equation Modeling. 1999;**6(4):**322-342.

76. Marlatt GA, George WH. Relapse prevention: introduction and overview of the model. British Journal Addiction. 1984;**79(3)**:261-273.

77. Newman S, Steed E, Mulligan K. Chronic Physical Illness: Self Management and Behavioural Interventions. Open University Press; 2009. p 72.

78. Lorig K, Holman HR. Long-term outcomes of an arthritis self-management

study: effects of reinforcement efforts. Social Science & Medicine. 1989;**29(2)**:221-224.

79. Glasgow RE, Toobert DJ, Hampson SE, Strycker LA. Implementation, generalization and long-term results of the "choosing well" diabetes self-management intervention. Patient Education and Counseling. 2002;**48(2)**:115-122.

80. Nguyen HQ, Carrieri-Kohlman V, Rankin SH, Slaughter R, Stulbarg MS. Is Internet-based support for dyspnea self-management in patients with chronic obstructive pulmonary disease possible? Results of a pilot study. Heart Lung. 2005;**34(1)**:51-62.

81. Internet homepage of the Japan Chronic Disease Self-Management Association (accessed on 2011.10.04): <http://www.j-cdsm.org/>.

82. Lorig K, Stewart A, Ritter P, Gonzalez VM, Laurent D, Lynch J. Outcome Measures for Health Education and Other Health Care Interventions. Sage Publications; 1996.

83. Yukawa K, Yamazaki Y, Yonekura Y, Togari T, Abbott F, Homma M, Park MJ, Kagawa Y. Effectiveness of chronic disease self-management program in Japan: Preliminary report of a longitudinal study. Nursing and Health Sicences. 2010;**12**:456-463.

84. Hatta H, Higashi A, Yashiro H, Ozasa K, Hayashi K, Kiyota K, Inokuchi H, Ikeda J, Fujita K, Watanabe Y, Kawai K. A validation of the hospital anxiety and depression scale. Japanese Journal of Psychosomatic Medicine. 1998;**38**:309-315.

85. Oxford Advanced Learner's Dictionary of Current English. Seventh edition. Oxford University Press; 2005.

86. Furr RM, Bacharach VR. Psychometrics: An Introduction. SAGE Publications, Inc; First edition, 2008.

87. Stein, C. (1956), "Inadmissibility of the usual estimator for the mean of a multivariate distribution", Proc. Third Berkeley Symp. Math. Statist. Prob. 1, pp. 197–206.

88. Efron, B.; Morris, C. (1977). "Stein's paradox in statistics". Scientific American 236 (5): 119–127.

89. Nunnally J, Bernstein I. Psychometric Theory. Third edition. McGraw-Hill; 1994.

90. Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and Regression Trees. 1st edition. Chapman and Hall/CRC; 1984.

91. Lewis RJ. An Introduction to Classification and Regression Tree (CART) Analysis. Presented at the 2000 Annual Meeting of the Society for Academic Emergency Medicine in San Francisco, California. Available at: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.95.4103>.

92. Lemon SC, Roy J, Clark MA, Friedmann PD, Rakowski W. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. Annals of Behavioral Medicine. 2003;**26(3)**:172-181.

93. Goel R, Misra A, Kondal D, Pandey RM, Vikram NK, Wasir JS, Dhingra V, Luthra K. Identification of insulin resistance in Asian Indian adolescents: classification and regression tree (CART) and logistic regression based classification rules. Clinical Endocrinology. 2009;**70(5)**:717-724.

94. Hendershot CS, Witkiewitz K, George WH, Marlatt GA. Relapse prevention for addictive behaviors. Substance Abuse Treatment, Prevention, and Policy. 2011;6:17. <http://www.substanceabusepolicy.com/content/6/1/17>.

95. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Second edition. Springer; 2009.

96. Zhang H, Singer BH. Recursive Partitioning and Applications (Springer Series in Statistics). 2nd edition. Springer; 2010.

97. Togari T, Yamazaki Y, Nakayama K, Shimizu J. Development of a short version of the sense of coherence scale for population survey. Journal of Epidemiology & Community Health. 2007;**61**:921-922.

98. Green L, Kreuter M. Health Program Planning: An Educational and Ecological Approach. Fourth edition. McGraw-Hill; 2005.

99. Rothman AJ. Toward a theory-based analysis of behavioral maintenance. Health Psychology. 2000;**19(1)**:64-69.

100. Park MJ, Yamazaki Y, Yonekura Y, Yukawa K, Ishikawa H, Kiuchi T, Green J. Predicting complete loss to follow-up after a health-education program: number of absences and face-to-face contact with a researcher. BioMed Central Medical Research Methodology. 2011;**11(1)**:145.

101. Kvien TK, Glennås A, Knudsrød OG, Smedstad LM: The validity of self-reported diagnosis of rheumatoid arthritis: results from a population survey followed by clinical examinations. Journal of Rheumatology. 1996;**23(11)**:1866-1871.

102. Foltynie T, Matthews FE, Ishihara L, Brayne C: The frequency and validity of self-reported diagnosis of Parkinson's Disease in the UK elderly: MRC CFAS cohort. BMC Neurology. 2006;6:29.

103. Singh JA: Discordance between self-report of physician diagnosis and administrative database diagnosis of arthritis and its predictors. Journal of Rheumatology. 2009;**36(9)**:2000-2008.

104. Wada K, Yatsuya H, Ouyang P, Otsuka R, Mitsuhashi H, Takefuji S, Matsushita K, Sugiura K, Hotta Y, Toyoshima H, Tamakoshi K: Self-reported medical history was generally accurate among Japanese workplace population. Journal of Clinical Epidemiology. 2009;**62**:306-313.

105. Wetzler S, Kahn R, Strauman TJ, Dubro A: Diagnosis of major depression by self-report. Journal of Personality Assessment. 1989;**53(1)**:22-30.

106. Linet MS, Harlow SD, Mclaughlin JK, Mccaffrey LD: A comparison of interview data and medical records for previous medical conditions and surgery. Journal of Clinical Epidemiology. 1989;**42(12)**:1207-1213.

107. Bush TL, Miller SR, Golden AL, Hale WA: Self-Report and Medical Record Report Agreement of Selected Medical Conditions in the Elderly. American Journal of Public Health. 1989;**79(11)**:1554-1556.

108. Heliövaara M, Aromaa A, Klaukka T, Knekt P, Joukamaa M: Reliability and validity of interview data on chronic diseases : The mini-Finland health survey. Journal of Clinical Epidemiology. 1993;**46(2)**:181-191.

109. Rasooly I, Papageorgiou AC, Badley EM: Comparison of clinical and self-reported diagnosis for rheumatology outpatients. Annals of the Rheumatic Diseases. 1995;**54**:850-852.

110. Muhujarine N, Mustard C, Roos LL, Young TK, Gelskey DE: Comparison of Survey and Physician Claims Data for Detecting Hypertension. Journal of Clinical Epidemiology. 1997;**50(6)**:711-718.

111. Bergmann MM, Byers T, Freedman DS, Mokdad A: Validity of Self-reported Diagnoses Leading to Hospitalization: A Comparison of Self-reports with Hospital Records in a Prospective Study of American Adults. American Journal of Epidemiology. 1998;**147(10)**:969-977.

112. Karlson EW, Lee IM, Cook NR, Manson JE, Buring JE, Hennekens CH: Comparison of Self-Reported Diagnosis of Connective Tissue Disease with Medical Records in Female Health Professionals: The Women's Health Cohort Study. American Journal of Epidemiology. 1999;**150(6)**:652-660.

113. Lampe FC, Walker M, Lennon LT, Whincup PH, Ebrahim S: Validity of a self-reported history of doctor-diagnosed angina. Journal of Clinical Epidemiology. 1999;**52(1)**:73-81.

114. Engstad T, Bønaa KH, Viitanen M: Validity of Self-Reported Stroke: The Tromsø Study. Stroke. 2000;**31**:1602-1607.

115. Bobadilla A, Guerra S, Sherrill D, Barbee R: How accurate is the self-reported diagnosis of chronic bronchitis? CHEST 2002;**122**:1234–1239.

116. Alonso A, Beunza JJ, Delgado-Rodríguez M, Martínez-González MA: Validation of self reported diagnosis of hypertension in a cohort of university graduates in Spain. BMC Public Health 2005;**5**:94

117. Merkin SS, Cavanaugh K, Longenecker JC, Fink NE, Levey AS, Powe NR: Agreement of self-reported comorbid conditions with medical and physician reports varied by disease among end-stage renal disease patients. Journal of Clinical Epidemiology. 2007;**60**:634-642

118. Parimi N, Lane NE, Bauer D, Hochberg MC, Nevitt MC: Accuracy of Self Reported Diagnosis of Hip Replacement. Arthritis Care & Research. 2010;**62(5)**: 719–724.

119. Malik AS, Giamouzis G, Georgiopoulou VV, Fike LV, Kalogeropoulos AP, Norton CR, Sorescu D, Azim S, Laskar SR, Smith AL, Dunbar SB, Butler J: Patient Perception Versus Medical Record Entry of Health-

Related Conditions Among Patients With Heart Failure. The American Journal of Cardiology 2011;**107**: 569 –572.

120. Schwartz CE, Sprangers MAG (editors). Adaptation to Changing Health: Response Shift in Quality-Of-Life Research. American Psychological Association; 2000.

121. Schwartz CE, Bode R, Repucci N, Becker J, Sprangers MA, Fayers PM. The clinical significance of adaptation to changing health: a meta-analysis of response shift. Quality of Life Research. 2006;**15(9)**:1533-1550.

122. Osborne RH, Hawkins M, Sprangers MA. Change of perspective: a measurable and desired outcome of chronic disease self-management intervention programs that violates the premise of preintervention/postintervention assessment. Arthritis & Rheumatism. 2006;**55(3)**:458-465.

123. Osborne RH, Elsworth GR, Whitfield K. The Health Education Impact Questionnaire (heiQ): an outcomes and evaluation measure for patient education and self-management interventions for people with chronic conditions. Patient Education and Counseling. 2007;**66(2)**:192-201.

124. Elzen H, Slaets JP, Snijders T, Steverink N: Evaluation of the chronic disease self-management program (CDSMP) among chronically ill older people in the Netherlands. Social Science & Medicine. 2007;**64**:1832-1841.

125. Gwaltney CJ, Shiffman S, Balabanis MH, Paty JA: Dynamic self-efficacy and outcome expectancies: prediction of smoking lapse and relapse. Journal of Abnormal Psychology. 2005;114(4):661-675.

126. Jordan JE, Haynes K, Livingston JA, Osborne RH: Comparison of the pre-post and transition question assessments in a health education setting. Journal of Clinical Epidemiology. 2009;**62**: 642-649.

127. Barlow JH, Turner AP, Wright CC: Long-term outcomes of an arthritis self-management programme. British Journal of Rheumatology. 1998;**37**:1315-1319.

128. Hass M, Groupp E, Muench J, Kraemer D, Mrummel-Smith K, Shama R, Ganger B, Attwood M, Fairweather A: Chronic Disease Self-Management Program for Low Back Pain in the Elderly. Journal of Manipulative and Physiological Therapeutics. 2005;**28**:228-237.

129. Smith L, Bosnic-Anticevich SZ, Mitchell B, Saini B, Krass I, Armour C: Treating asthma with a self-management model of illness behavior in an Australian community pharmacy setting. Social Science & Medicine. 2007;**64**:1501-1511.

130. Yip YB, Sit JW, Wong D, Chong S, Chung LH: A 1-year follow-up of an experimental study of a self-management arthritis programme with an added exercise component of clients with osteoarthritis of the knee. Psychology, Health & Medicine. 2008;**13**:402-414.

131. Chan SC, Siu AMH, Poon PKK, Chan CCH: Chronic disease self-management program for Chinese patients: a preliminary multi-baseline study. International Journal of Rehabilitation Research. 2005;**28**:351-354.

132. Goeppinger J, Armstrong B, Schwartz T, Ensley D, Brady TJ: Self-Management Education for Persons With Arthritis: Managing

Comorbidity and Eliminating Health Disparities. Arthritis & Rheumatism. 2007;**57**:1081-1088.

133. Jerant A, Kravitz R, Moore-Hill M, Franks P: Depressive Symptoms Moderated the Effect of Chronic Illness Self-Management Training on Self-Efficacy. Medical Care. 2008;**46**:523-531.
134. DeVellis RF. Scale Development: Theory and Applications. Third Edition. Sage Publications, Inc.; 201

Table 1. Patterns of questionnaire return

	Baseline	3 months	6 months	12 months	Number of participants	Total
Usable*						
	○	○	○	○	270	
	○	○	○	×	41	
	○	○	×	○	27	
	○	×	○	○	26	364
Not usable						
	○	○	×	×	28	
	○	×	×	○	16	
	○	×	○	×	12	
	○	×	×	×	59	115

○: questionnaire returned. ×: questionnaire not returned.

Table 2. Basic information about the measurements (n = 364)

Dependent variables[a]	Number of items	Possible range	Actual range	Mean	Standard deviation	Skewness (standard error)	Kurtosis (standard error)	n and % at floor	n and % at ceiling	n and % missing	Cronbach's alpha
Health status											
Self-rated health	1	1-5	1-5	3.41	0.93	-0.49 (0.13)	-0.01 (0.26)	12 (3.3%)	32 (8.8%)	4 (1.1%)	b
Pain	1	0-10	0-10	3.52	2.98	0.42 (0.13)	-1.02 (0.26)	82 (22.5%)	8 (2.2%)	2 (0.5%)	b
Self-management behaviors											
Communication with MDs	3	0-15	0-15	6.29	3.71	0.59 (0.13)	-0.25 (0.25)	7 (1.9%)	16 (4.4%)	1 (0.3%)	0.78
Coping with symptoms	6	0-30	0-30	7.28	4.79	1.19 (0.13)	2.52 (0.26)	18 (4.9%)	1 (0.3%)	0	0.72
Psychological health											
Self-efficacy	6	0-60	0-60	32.21	12.18	-0.12 (0.13)	-0.38 (0.25)	2 (0.5%)	2 (0.5%)	0	0.92
Health distress	4	0-20	0-20	8.56	5.25	0.46 (0.13)	-0.70 (0.25)	11 (3.0%)	12 (3.3%)	1 (0.3%)	0.92
Anxiety	7	0-21	0-18	6.80	4.16	0.54 (0.13)	-0.40 (0.25)	9 (2.5%)	1 (0.3%)	1 (0.3%)	0.83
Depression	7	0-21	0-18	7.05	3.67	0.46 (0.13)	-0.32 (0.25)	2 (0.5%)	2 (0.5%)	1 (0.3%)	0.72
Others (measures at 3 months)											
Self-efficacy at 3 months	6	0-60	0-60	34.81	11.83	-0.44 (0.13)	-0.33 (0.27)	1 (0.3%)	1 (0.3%)	35 (9.6%)	0.93
Perceived Positive Change	7	0-28	3-28	21.15	4.50	-0.53 (0.13)	0.34 (0.27)	1 (0.3%)	29 (8.0%)	34 (9.3%)	0.88

[a] The results shown are for the baseline data (except for self-efficacy and Perceived Positive Change, which was also measured 3 months after the baseline data were collected).

[b] Cronbach's alpha cannot be computed for one-item measures.

Table 3. Demographic and clinical characteristics of the group as a whole (n = 364)

		Number (%)		
Age (years)	mean ± SD (range)	48.6 ± 14.1 (18-83)		
Sex	Male	79 (21.7)		
	Female	285 (78.3)		
Schooling	High school or less	121 (33.2)		
	College or more	241 (66.2)		
Civil status	Living together	190 (52.2)		
	Others	174 (47.8)		
Years since diagnosis	mean ± SD (range)	14.2 ± 10.0 (0.4-63)		
	median (25%, 75%)	10.0 (5.0, 20.0)		
Number of diagnoses	median (25%, 75%)	1.0 (1.0, 2.0)		
	min-max	1-7		
	1	210 (57.7%)		
	2	91 (25.0%)		
	3	41 (11.3%)		
	≥ 4	22 (6.0%)		
		Only 1 diagnosis (number)	> 1 diagnosis (number)	Total (n, %)
Diagnoses	Allergic disease	28	73	101 (27.7%)
	Cardiovascular disease	14	66	80 (20.0%)
	Connective tissue disease	32	35	67 (18.4%)
	Diabetes	29	36	65 (17.9%)
	Rheumatoid arthritis	20	19	39 (10.7%)
	Fibromyalgia syndrome	12	17	29 (8.0%)
	Asthma	0	18	18 (4.9%)
	Inflammatory bowel disease	10	6	16 (4.4%)
	Parkinson's disease	10	2	12 (3.3%)
	Depression	2	8	10 (2.7%)
	Others	51	87	138 (37.9%)
Number of absences*	median (25%, 75%)	0 (0, 1.0)		
	min-max	0-5		

*Number of absences from program sessions; minimum possible number = 0, maximum possible number = 6.

Table 4. Magnitude of the decay of impact for each outcome measure.

Outcome	N	Skewness (standard error)	% of full scale Median	% of full scale 25%-75%	Standardized effect size Median	Standardized effect size 25%-75%
Health status						
Self-rated health	94	1.40 (0.25)	23.20	23.13-46.13	0.98	0.97-1.94
Pain	25	1.02 (0.46)	39.54	31.96-51.55	1.70	1.38-2.06
Self-management behaviors						
Communication	54	0.76 (0.32)	26.25	17.89-35.69	0.99	0.68-1.35
Coping	71	1.74 (0.28)	16.70	13.94-24.44	0.97	0.81-1.36
Psychological health						
Self-efficacy	52	1.31 (0.33)	21.37	14.80-31.12	1.41	0.79-1.66
Health distress	56	1.60 (0.32)	21.80	17.42-34.90	1.07	0.86-1.71
Anxiety	44	1.02 (0.36)	22.20	16.30-32.10	1.17	0.86-1.69
Depression	41	1.43 (0.37)	16.40	16.40-26.38	1.09	1.09-1.63

Table 5. Time of the start of decay of impact

Variable	n*	Decay-of-impact starting time	
		3 months, n (%)	6 months, n (%)
Health status			
Self-rated health	70	43 (61.4%)	27 (38.6%)
Pain	19	9 (47.4%)	10 (52.6%)
Self-management behaviors			
Communication	46	12 (26.1%)	34 (73.9%)
Coping	61	28 (45.9%)	33 (54.1%)
Psychological health			
Self-efficacy	46	20 (43.5%)	26 (56.5%)
Health distress	43	14 (32.6%)	29 (67.4%)
Anxiety	37	10 (27.0%)	27 (73.0%)
Depression	37	14 (37.8%)	23 (62.2%)

* Participants with decay of impact and with data at all follow-up times.

Table 6a. Pattern-defined subgroups for each outcome, by diagnosis

	Self-rated health				Pain			
	Improvement	Decay Mag.*	No change	Deterioration	Improvement	Decay Mag.*	No change	Deterioration
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Diagnoses								
Allergic disease (n = 101)	30 (30.6%)	**46.1%** 17 (17.3%)	28 (28.6%)	23 (23.5%)	13 (13.0%)	**39.9%** 5 (5.0%)	67 (67.0%)	15 (15.0%)
Cardiovascular disease (n = 80)	23 (30.3%)	**23.2%** 21 (27.6%)	20 (26.3%)	12 (15.8%)	10 (12.7%)	**39.9%** 7 (8.9%)	54 (68.4%)	8 (10.1%)
Connective tissue disease (n = 67)	23 (35.9%)	**23.2%** 12 (18.8%)	18 (28.1%)	11 (17.2%)	8 (11.9%)	**31.8%** 2 (3.0%)	50 (74.6%)	7 (10.4%)
Diabetes (n = 65)	26 (44.1%)	**23.2%** 17 (28.8%)	11 (18.6%)	5 (8.5%)	8 (12.3%)	**39.9%** 6 (9.2%)	42 (64.6%)	9 (13.8%)
Rheumatoid arthritis (n = 39)	9 (23.1%)	**23.2%** 12 (30.8%)	11 (28.2%)	7 (17.9%)	4 (10.3%)	**0** 0	30 (76.9%)	5 (12.8%)
Fibromyalgia syndrome (n = 29)	10 (34.5%)	**23.2%** 11 (37.9%)	6 (20.7%)	2 (6.9%)	5 (17.2%)	**35.6%** 2 (6.9%)	21 (72.4%)	1 (3.4%)
Asthma (n = 18)	3 (16.7%)	**34.7%** 4 (22.2%)	3 (16.7%)	8 (44.4%)	3 (16.7%)	**32.0%** 1 (5.6%)	11 (61.1%)	3 (16.7%)
Inflammatory bowel disease (n = 16)	6 (37.5%)	**46.1%** 4 (25.0%)	2 (12.5%)	4 (25.0%)	1 (6.3%)	**32.0%** 1 (6.3%)	12 (75.0%)	2 (12.5%)
Parkinson's disease (n = 12)	2 (18.2%)	**46.2%** 4 (36.4%)	3 (27.3%)	2 (18.2%)	0	**59.4%** 2 (16.7%)	6 (50.0%)	4 (33.3%)
Depression (n = 10)	5 (50.0%)	**23.2%** 3 (30.0%)	1 (10.0%)	1 (10.0%)	1 (10.0%)	**31.8%** 2 (20.0%)	4 (44.4%)	2 (22.2%)
Multimorbidity								
> 1 diagnosis (n = 154)	51 (34.5%)	**23.2%** 35 (23.6%)	34 (23.0%)	28 (18.9%)	24 (15.9%)	**35.9%** 10 (6.6%)	98 (64.9%)	19 (12.6%)
1 diagnosis (n = 210)	62 (30.4%)	**23.2%** 58 (28.4%)	51 (25.0%)	33 (16.2%)	19 (9.0%)	**39.5%** 15 (7.1%)	152 (72.5%)	24 (11.4%)
2 diagnoses (n = 91)	30 (34.5%)	**23.2%** 24 (27.6%)	19 (21.8%)	14 (16.1%)	17 (19.3%)	**32.0%** 5 (5.7%)	58 (65.9%)	8 (9.1%)
3 diagnoses (n = 41)	14 (35.0%)	**23.2%** 7 (17.5%)	9 (22.5%)	10 (25.0%)	5 (12.2%)	**39.9%** 4 (9.7%)	25 (61.0%)	7 (17.1%)
4 or > 4 diagnoses (n = 22)	7 (33.4%)	**23.2%** 4 (19.0%)	6 (28.6%)	4 (19.0%)	2 (9.1%)	**39.9%** 1 (4.5%)	15 (68.2%)	4 (18.2%)

* Percentages in bold indicate the magnitude of the decay of impact as a percent of the full-scale value.

Table 6b. Pattern-defined subgroups for each outcome, by diagnosis

	Communication				Coping			
	Improvement	Decay Mag.*	No change	Deterioration	Improvement	Decay Mag.*	No change	Deterioration
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Diagnoses								
Allergic disease (n = 101)	24 (24.2%)	**41.3%** 14 (14.1%)	45 (45.5%)	16 (16.2%)	29 (28.7%)	**21.4%** 21 (20.8%)	43 (42.6%)	8 (7.9%)
Cardiovascular disease (n = 80)	19 (24.1%)	**23.6%** 10 (12.7%)	34 (43.0%)	16 (20.0%)	15 (19.5%)	**16.7%** 19 (24.7%)	39 (50.6%)	4 (5.2%)
Connective tissue disease (n = 67)	16 (24.2%)	**23.5%** 7 (10.6%)	31 (47.0%)	12 (18.2%)	15 (22.7%)	**16.7%** 11 (16.7%)	35 (53.0%)	5 (7.6%)
Diabetes (n = 65)	17 (26.2%)	**35.7%** 9 (13.8%)	27 (41.5%)	12 (18.5%)	15 (23.4%)	**22.2%** 9 (13.8%)	35 (54.7%)	5 (7.8%)
Rheumatoid arthritis (n = 39)	10 (25.6%)	**23.8%** 5 (12.8%)	20 (51.3%)	4 (10.3%)	14 (35.9%)	**22.2%** 6 (15.4%)	16 (41.0%)	3 (7.7%)
Fibromyalgia syndrome (n = 29)	7 (24.1%)	**23.7%** 6 (20.7%)	14 (48.3%)	2 (6.9%)	10 (35.7%)	**19.4%** 6 (21.4%)	10 (35.7%)	2 (7.1%)
Asthma (n = 18)	4 (22.2%)	**35.4%** 5 (27.8%)	6 (33.3%)	3 (16.7%)	7 (38.9%)	**13.8%** 2 (11.1%)	7 (38.9%)	2 (11.1%)
Inflammatory bowel disease (n = 16)	3 (18.8%)	**20.9%** 2 (12.5%)	9 (56.3%)	2 (12.5%)	2 (12.5%)	**16.6%** 4 (25.0%)	10 (62.5%)	0
Parkinson's disease (n = 12)	1 (8.3%)	**26.6%** 2 (16.7%)	7 (58.3%)	2 (16.7%)	2 (16.7%)	**15.6%** 1 (8.3%)	6 (50.0%)	3 (25.0%)
Depression (n = 10)	5 (50.0%)	**29.5%** 1 (10.0%)	3 (30.0%)	1 (10.0%)	5 (50.0%)	**16.7%** 1 (10.0%)	4 (40.0%)	0
Multimorbidity								
> 1 diagnosis (n = 154)	41 (27.0%)	**23.8%** 24 (15.7%)	65 (42.8%)	22 (14.5%)	41 (27.2%)	**19.5%** 31 (20.5%)	69 (45.7%)	10 (6.6%)
1 diagnosis (n = 210)	43 (20.7%)	**29.1%** 30 (14.4%)	99 (47.6%)	36 (17.3%)	44 (21.2%)	**16.7%** 40 (19.2%)	107 (51.4%)	17 (8.2%)
2 diagnoses (n = 91)	28 (31.1%)	**23.8%** 14 (15.6%)	40 (44.4%)	8 (8.9%)	22 (25.0%)	**21.8%** 20 (22.7%)	39 (44.3%)	7 (8.0%)
3 diagnoses (n = 41)	9 (22.5%)	**23.8%** 7 (17.5%)	17 (42.5%)	7 (17.5%)	13 (31.7%)	**16.7%** 9 (22.0%)	17 (41.5%)	2 (4.8%)
4 or > 4 diagnoses (n = 22)	4 (18.2%)	**41.2%** 3 (13.6%)	8 (36.4%)	7 (31.8%)	6 (27.3%)	**13.9%** 2 (9.1%)	13 (59.1%)	1 (4.5%)

* Percentages in bold indicate the magnitude of the decay of impact as a percent of the full-scale value.

Table 6c. Pattern defined subgroups for each outcome, by diagnosis

	Self-efficacy				Health distress			
	Improvement	Decay Mag.*	No change	Deterioration	Improvement	Decay Mag.*	No change	Deterioration
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Diagnoses								
Allergic disease (n = 101)	32 (31.7%)	**19.9%** 10 (9.9%)	35 (34.7%)	24 (23.8%)	28 (27.7%)	**19.6%** 20 (19.8%)	44 (43.6%)	9 (8.9%)
Cardiovascular disease (n = 80)	23 (29.1%)	**28.9%** 10 (12.7%)	36 (45.6%)	10 (12.7%)	24 (30.0%)	**21.7%** 11 (13.8%)	35 (43.8%)	10 (12.5%)
Connective tissue disease (n = 67)	18 (27.3%)	**16.2%** 7 (10.6%)	34 (51.5%)	7 (10.6%)	20 (29.9%)	**21.7%** 11 (16.4%)	31 (46.3%)	5 (7.5%)
Diabetes (n = 65)	22 (33.8%)	**17.8%** 9 (13.8%)	25 (38.5%)	9 (13.8%)	21 (32.3%)	**30.5%** 9 (13.8%)	29 (44.6%)	6 (9.2%)
Rheumatoid arthritis (n = 39)	11 (28.2%)	**26.7%** 5 (12.8%)	18 (46.2%)	5 (12.8%)	11 (28.2%)	**36.9%** 2 (5.1%)	24 (61.5%)	2 (5.1%)
Fibromyalgia syndrome (n = 29)	11 (37.9%)	**31.1%** 3 (10.3%)	8 (28.6%)	6 (21.4%)	11 (37.9%)	**21.8%** 7 (24.1%)	9 (31.0%)	2 (6.9%)
Asthma (n = 18)	3 (16.7%)	**13.2%** 2 (11.1%)	7 (38.9%)	6 (33.3%)	7 (38.9%)	**30.5%** 3 (16.7%)	7 (38.9%)	1 (5.6%)
Inflammatory bowel disease (n = 16)	6 (37.5%)	**22.2%** 1 (6.3%)	6 (37.5%)	3 (18.8%)	6 (37.5%)	**23.9%** 4 (25.0%)	5 (31.3%)	1 (6.3%)
Parkinson's disease (n = 12)	0	**36.2%** 2 (16.7%)	7 (58.3%)	3 (25.0%)	2 (16.7%)	**30.5%** 2 (16.7%)	4 (36.4%)	3 (27.3%)
Depression (n = 10)	5 (50.0%)	**17.0%** 4 (40.0%)	0	1 (10.0%)	6 (60.0%)	**39.1%** 3 (30.0%)	1 (10.0%)	0
Multimorbidity								
> 1 diagnosis (n = 154)	51 (33.3%)	**17.7%** 19 (12.4%)	60 (39.2%)	23 (15.1%)	55 (35.7%)	**21.7%** 24 (15.6%)	61 (39.6%)	14 (9.1%)
1 diagnosis (n = 210)	57 (27.4%)	**23.7%** 33 (15.9%)	83 (39.9%)	35 (16.8%)	60 (28.7%)	**21.8%** 30 (14.4%)	104 (49.8%)	15 (7.1%)
2 diagnoses (n = 91)	36 (40.0%)	**17.7%** 9 (10.0%)	36 (40.0%)	9 (10.0%)	33 (36.3%)	**21.7%** 11 (12.1%)	40 (44.0%)	7 (7.6%)
3 diagnoses (n = 41)	9 (22.0%)	**17.0%** 8 (19.5%)	15 (36.5%)	9 (22.0%)	14 (34.1%)	**17.5%** 10 (24.4%)	14 (34.1%)	3 (7.4%)
4 or > 4 diagnoses (n = 22)	6 (27.3%)	**17.7%** 2 (9.1%)	9 (40.9%)	5 (22.7%)	8 (36.4%)	**39.2%** 3 (13.6%)	7 (31.8%)	4 (18.2%)

* Percentages in bold indicate the magnitude of the decay of impact as a percent of the full-scale value.

Table 6d. Pattern-defined subgroups for each outcome, by diagnosis

	Anxiety				Depression			
	Improvement	Decay Mag.*	No change	Deterioration	Improvement	Decay Mag.*	No change	Deterioration
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Diagnoses								
Allergic disease (n = 101)	19 (19.4%)	**22.2%** 10 (10.2%)	52 (53.1%)	17(17.3%)	16 (16.3%)	**16.4%** 13 (13.3%)	52 (53.1%)	17 (17.3%)
Cardiovascular disease (n = 80)	16 (20.0%)	**26.9%** 7 (8.8%)	43 (53.8%)	14 (17.5%)	12 (15.0%)	**16.3%** 8 (10.0%)	47 (58.8%)	13 (16.3%)
Connective tissue disease (n = 67)	11 (16.7%)	**18.3%** 6 (9.1%)	42 (63.6%)	7 (10.6%)	14 (21.2%)	**16.4%** 5 (7.6%)	40 (60.6%)	7 (10.6%)
Diabetes (n = 65)	12 (18.5%)	**20.3%** 7 (10.8%)	40 (61.5%)	6 (9.2%)	12 (18.5%)	**28.4%** 4 (6.2%)	38 (58.5%)	11 (16.9%)
Rheumatoid arthritis (n = 39)	6 (15.8%)	**39.6%** 2 (5.3%)	29 (76.3%)	1 (2.6%)	7 (18.4%)	**16.1%** 5 (13.2%)	24 (63.2%)	2 (5.3%)
Fibromyalgia syndrome (n = 29)	9 (31.0%)	**27.8%** 7 (24.1%)	9 (31.0%)	4 (13.8%)	7 (24.1%)	**40.4%** 3 (10.3%)	12 (41.4%)	7 (24.1%)
Asthma (n = 18)	2 (11.1%)	**18.0%** 2 (11.1%)	10 (55.6%)	4 (22.2%)	3 (16.7%)	**16.4%** 3 (16.7%)	8 (44.4%)	4 (22.2%)
Inflammatory bowel disease (n = 16)	5 (31.3%)	**44.3%** 1 (6.3%)	7 (43.8%)	3 (18.8%)	3 (18.8%)	**20.2%** 3 (18.8%)	7 (43.8%)	3 (18.8%)
Parkinson's disease (n = 12)	1 (8.3%)	**16.1%** 3 (25.0%)	6 (50.0%)	2 (16.7%)	3 (25.0%)	**17.3%** 2 (16.7%)	5 (41.7%)	2 (16.7%)
Depression (n = 10)	3 (30.0%)	**16.3%** 3 (30.0)	3 (30.0%)	1 (10.0%)	3 (30.0%)	**30.3%** 2 (20.0%)	4 (40.0%)	1 (10.0%)
Multimorbidity								
> 1 diagnosis (n = 154)	35 (23.0%)	**24.3%** 19 (12.5%)	77 (50.7%)	21 (13.8%)	32 (21.1%)	**16.4%** 14 (9.2%)	87 (57.2%)	19 (12.5%)
1 diagnosis (n = 210)	32 (15.3%)	**20.3%** 25 (12.0%)	124 (59.3%)	28 (13.4%)	41 (19.6%)	**20.2%** 27 (12.9%)	107 (51.2%)	34 (16.3%)
2 diagnoses (n = 91)	26 (28.9%)	**20.3%** 10 (11.1%)	40 (44.4%)	14 (15.6%)	27 (30.0%)	**16.4%** 6 (6.7%)	46 (51.1%)	11 (12.2%)
3 diagnoses (n = 41)	7 (17.5%)	**23.6%** 4 (10.0%)	24 (60.0%)	5 (12.5%)	2 (5.0%)	**16.4%** 7 (17.5%)	28 (70.0%)	3 (7.5%)
4 or > 4 diagnoses (n = 22)	2 (9.1%)	**28.3%** 5 (22.7%)	13 (59.1%)	2 (9.1%)	3 (13.6%)	**24.2%** 1 (4.5%)	13 (59.2%)	5 (22.7%)

* Percentages in bold indicate the magnitude of the decay of impact as a percent of the full-scale value.

Table 6e. The results of Fisher Exact tests of Table 6a-6d.

	Self-rated health	Pain	Communication	Coping	Self-efficacy	Health distress	Anxiety	Depression
	Fisher Exact p-value							
Allergic*CVD	0.29	0.49	> .999	0.27	0.6	0.37	> .999	0.78
Allergic*CTD	> .999	> .999	0.78	> .999	0.77	0.64	> .999	0.23
Allergic*DM	0.74	0.46	> .999	0.80	0.79	0.34	> .999	0.22
Allergic*RA	0.12	0.53	> .999	0.42	0.74	0.12	> .999	> .999
Allergic*FMS	0.29	> .999	0.74	> .999	> .999	> .999	0.75	0.48
Allergic*Asthma	0.41	> .999	0.45	0.46	0.59	0.72	0.61	> .999
Allergtic*IBD	> .999	0.52	> .999	0.39	> .999	> .999	0.64	> .999
Allergic*Parkinson	0.20	0.11	0.55	> .999	0.10	> .999	0.28	> .999
Allergic*Depression	> .999	0.25	0.65	0.39	0.24	0.73	0.65	> .999
CVD*CTD	0.26	0.41	> .999	0.43	> .999	0.79	> .999	0.50
CVD*DM	0.52	> .999	> .999	0.19	> .999	> .999	0.75	0.48
CVD*RA	0.60	0.26	> .999	0.10	> .999	0.46	> .999	> .999
CVD*FMS	0.79	0.67	0.51	0.36	0.73	0.76	0.50	0.70
CVD*Asthma	0.70	> .999	0.44	0.13	0.64	> .999	0.58	> .999
CVD*IBD	0.74	> .999	> .999	> .999	0.65	0.71	0.65	> .999
CVD*Parkinson	0.67	0.21	0.54	0.58	0.11	0.59	0.13	> .999
CVD*Depression	0.71	0.57	0.64	0.18	0.45	> .999	0.63	> .999
CTD*DM	0.65	0.39	> .999	0.78	> .999	0.79	> .999	> .999
CTD*RA	0.11	> .999	> .999	0.54	> .999	0.28	> .999	0.45
CTD*FMS	0.26	> .999	0.47	> .999	0.72	> .999	0.73	> .999
CTD*Asthma	0.40	> .999	0.24	0.43	0.62	> .999	0.62	0.34
CTD*IBD	0.73	0.45	> .999	0.38	0.65	> .999	0.62	0.34
CTD*Parkinson	0.19	0.10	0.27	> .999	0.10	0.62	0.27	0.61
CTD*Depression	> .999	0.20	0.65	0.37	0.42	> .999	0.64	0.61
DM*RA	0.28	0.24	> .999	0.75	> .999	0.46	0.68	0.43
DM*FMS	0.42	0.66	0.51	> .999	0.72	0.54	0.74	> .999
DM*Asthma	0.43	> .999	0.43	0.68	0.63	> .999	> .999	0.33
DM*IBD	> .999	> .999	> .999	0.36	0.65	0.70	0.62	0.33
RA*FMS	> .999	0.49	0.70	0.73	0.69	0.24	0.66	0.67
RA*Asthma	> .999	> .999	0.40	> .999	> .999	0.62	0.55	> .999
RA*IBD	0.46	0.33	> .999	0.16	0.62	0.34	> .999	> .999
RA*Parkinson	> .999	0.67	0.53	> .999	0.14	0.22	0.22	> .999
RA*Depression	0.43	0.14	0.62	> .999	0.67	0.61	0.58	> .999
FMS*Asthma	> .999	> .999	> .999	0.66	0.57	0.70	> .999	0.61
FMS*IBD	0.70	> .999	> .999	0.35	> .999	> .999	0.35	0.61
FMS*Parkinson	0.66	0.17	> .999	> .999	0.48	> .999	0.58	> .999
FMS*Depression	0.68	0.50	0.33	0.62	0.36	> .999	> .999	> .999
Asthma*IBD	0.64	> .999	> .999	0.14	0.52	> .999	0.50	> .999
Asthma*Parkinson	> .999	0.40	> .999	> .999	0.43	0.58	> .999	> .999
Asthma*Depression	0.62	0.49	0.29	> .999	> .999	> .999	> .999	> .999
IBD*Parkinson	0.61	> .999	> .999	0.52	0.26	> .999	0.19	> .999
IBD*Depression	> .999	> .999	0.54	0.24	0.31	> .999	0.54	> .999
Parkinson*Depression	0.59	> .999	0.23	> .999	0.45	> .999	0.57	> .999
> 1 diagnosis*1 diagnosis	0.32	0.31	0.73	0.63	0.23	0.74	0.44	0.33

Table 7a. CART models for predicting decay of impact: indices of classification tree performance and relative importance of independent variables

	Self-rated health	Pain	Communication	Coping
Number: decay, improvement	93, 113	25, 43	54, 84	71, 85
Indices of classification-tree performance				
Misclassification risk ± SE (0-1; Lower is better.)	0.25 ± 0.03	0.25 ± 0.05	0.24 ± 0.04	0.19 ± 0.03
% correctly classified as decay (0-100%; Higher is better.)	71.0%	40.0%	46.3%	81.7%
Area under ROC curve (0.5-1.0; Higher is better.)	0.789	0.680	0.775	0.863
Relative importance of independent variables. For each outcome, maximum importance = 100%. Percentages in **bold** are for indepnedent variables that were included in the final models. Other percentages are for independent variables that were not included in the final models. Asterisks (*) are for independent variables that were not included in any model.				
Socio-demographic				
Age	20.0%	**87.8%**	45.7%	**74.9%**
Sex	5.1%	*	*	40.5%
Schooling	3.5%	*	*	42.3%
Civil status	1.5%	*	14.1%	1.7%
Clinical				
Years since diagnosis	**42.6%**	*	**100.0%**	**37.6%**
Number of diagnoses	15.2%	*	6.8%	1.7%
Allergic disease	14.0%	*	14.1%	20.9%
Cardiovascular disease	1.0%	*	*	11.7%
Connective tissue disease	5.7%	*	*	13.8%
Diabetes	9.5%	*	11.9%	17.6%
Rheumatoid arthritis	12.2%	*	*	26.7%
Fibromyalgia syndrome	**25.1%**	*	*	11.4%
Asthma	1.4%	*	13.0%	2.5%
Inflammatory bowel disease	1.6%	*	0.0%	*
Parkinson's disease	1.6%	**64.0%**	*	2.5%
Depression	*	*	13.8%	5.2%
1 diagnosis	10.4%	*	*	1.7%
Health status at baseline				
Self-rated health	14.8%	28.0%	**47.4%**	27.0%
Pain	**44.5%**	**83.1%**	28.1%	47.0%
Self-management behaviors at baseline				
Communication	**79.1%**	*	25.9%	**62.4%**
Coping	*	**100.0%**	18.1%	**70.8%**
Psychological health at baseline				
Self-efficacy	*	*	38.5%	**62.7%**
Health distress	**28.3%**	60.6%	**51.6%**	40.5%
Anxiety	**100.0%**	*	**55.5%**	**100.0%**
Depression	*	18.6%	55.7%	32.6%
Others				
Self-efficacy at 3 months	**48.2%**	*	40.6%	**68.3%**
Number of absences	7.4%	*	**67.3%**	**27.3%**
Perceived positive change	20.3%	*	**99.5%**	**60.6%**

Table 7b. CART models for predicting decay of impact: indices of classification tree performance and relative importance of independent variables

	Self-efficacy	Health distress	Anxiety	Depression
Number: decay, improvement	52, 108	54, 115	44, 67	41, 73
Indices of classification-tree performance				
Misclassification risk ± SE (0-1; Lower is better.)	0.23 ± 0.03	0.24 ± 0.03	0.20 ± 0.04	0.28 ± 0.04
% correctly classified as decay (0-100%; Higher is better.)	30.8%	35.2%	77.3%	24.4%
Area under ROC curve (0.5-1.0; Higher is better.)	0.732	0.696	0.832	0.683
Relative importance of independent variables. For each outcome, maximum importance = 100%.				
Percentages in **bold** are for independent variables that were included in the final models.				
Other Percentages are for independent variables that were not included in the final models.				
Asterisks (*) are for independent variables that were not included in any model.				
Socio-demographic				
Age	**52.1%**	8.6%	**65.8%**	*
Sex	*	*	*	*
Schooling	*	**43.7%**	1.2%	*
Civil status	*	*	*	*
Clinical				
Years since diagnosis	**100.0%**	**71.6%**	12.6%	**91.2%**
Number of diagnoses	5.2%	7.0%	11.5%	*
Allergic disease	*	*	31.9%	*
Cardiovascular disease	*	*	8.8%	*
Connective tissue disease	*	*	4.8%	*
Diabetes	*	*	4.8%	*
Rheumatoid arthritis	*	*	5.0%	*
Fibromyalgia syndrome	*	*	*	*
Asthma	*	*	*	*
Inflammatory bowel disease	*	9.2%	24.6%	*
Parkinson's disease	*	*	*	*
Depression	*	*	3.6%	*
1 diagnosis	*	*	6.3%	*
Health status at baseline				
Self-rated health	34.1%	7.4%	3.1%	*
Pain	*	**62.9%**	18.2%	*
Self-management behaviors at baseline				
Communication	*	**53.6%**	**50.8%**	*
Coping	**37.9%**	**64.0%**	26.8%	**88.6%**
Psychological health at baseline				
Self-efficacy	**58.8%**	*	43.7%	*
Health distress	*	*	**68.6%**	**100.0%**
Anxiety	**42.9%**	8.2%	22.0%	27.8%
Depression	3.7%	*	44.7%	6.8%
Others				
Self-efficacy at 3 months	26.3%	**100.0%**	**100.0%**	2.9%
Number of absences	**38.7%**	9.2%	**61.7%**	8.8%
Perceived positive change	12.9%	7.9%	30.3%	22.6%

From: “Evaluation and measurement: some dilemmas for health education”

(L.W. Green, Am J Pub Hlth, 1977)

Decay of impact
Attenuation
Backsliding
Deterioration
Relapse

Therefore, “**reinforcement** is as important to education as booster shots are to sustained immunization”.

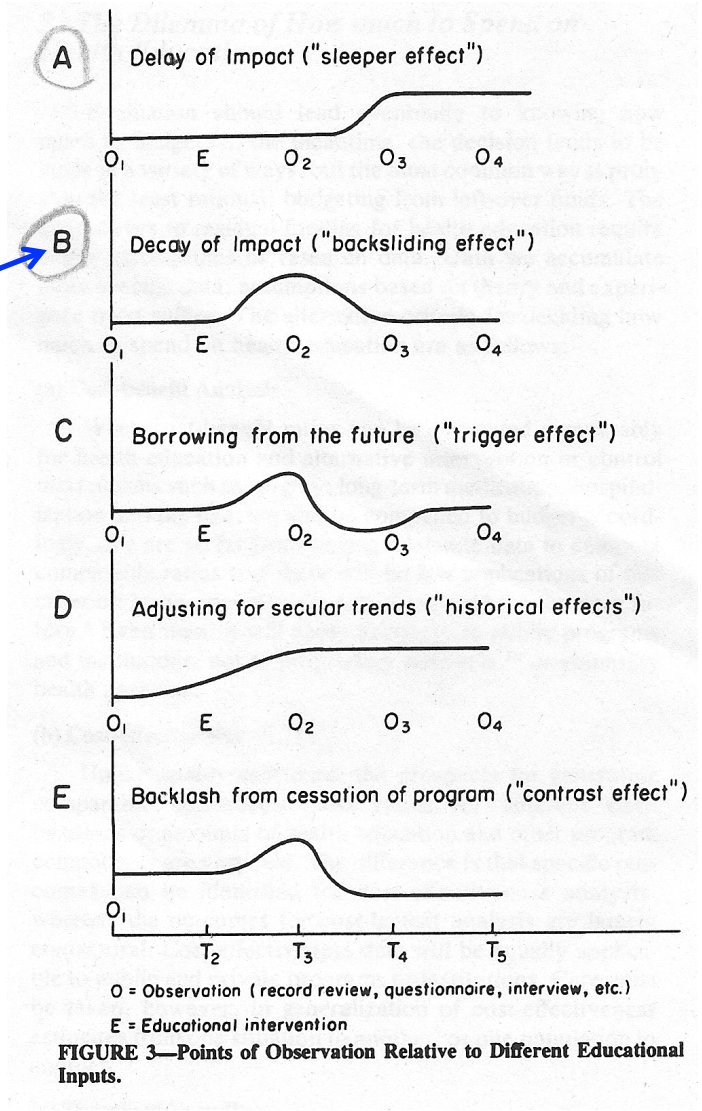


FIGURE 3—Points of Observation Relative to Different Educational Inputs.

Figure 1.
Decay of impact and the need for reinforcement, as defined by LW Green in 1977.

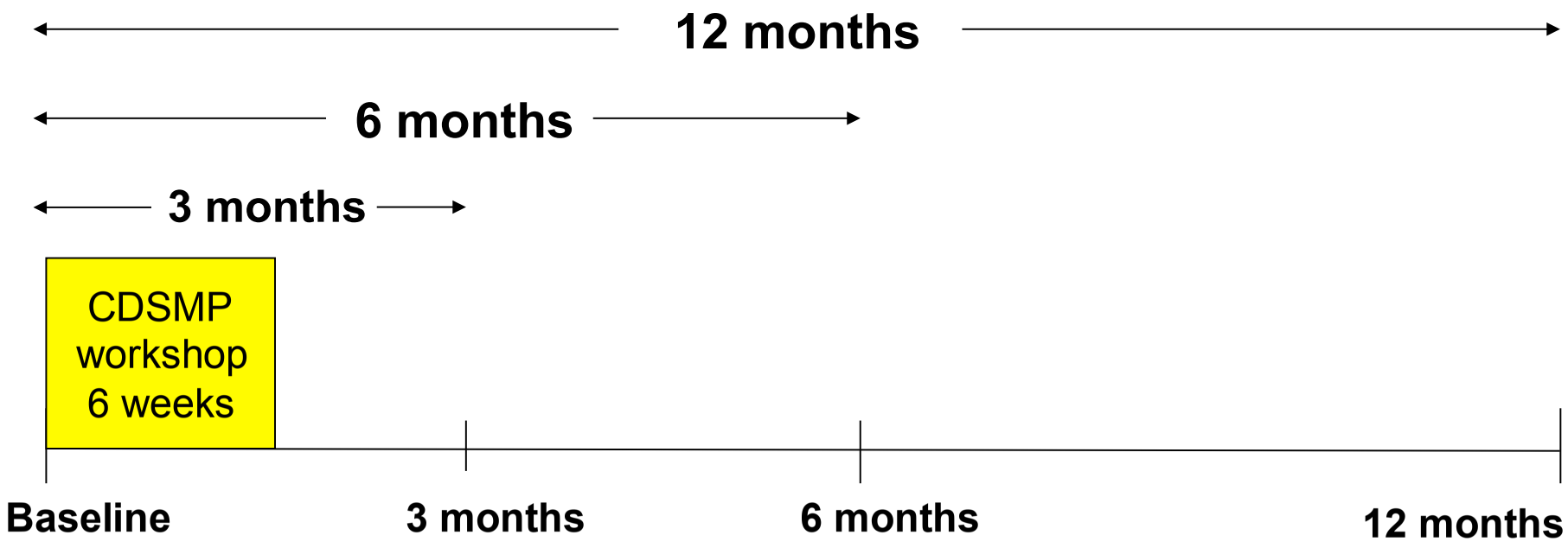
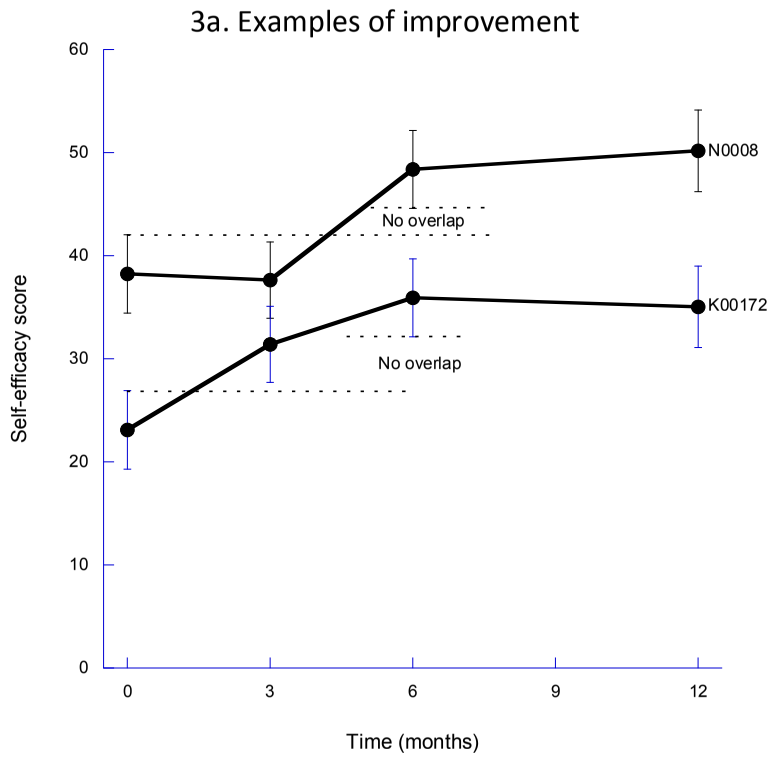
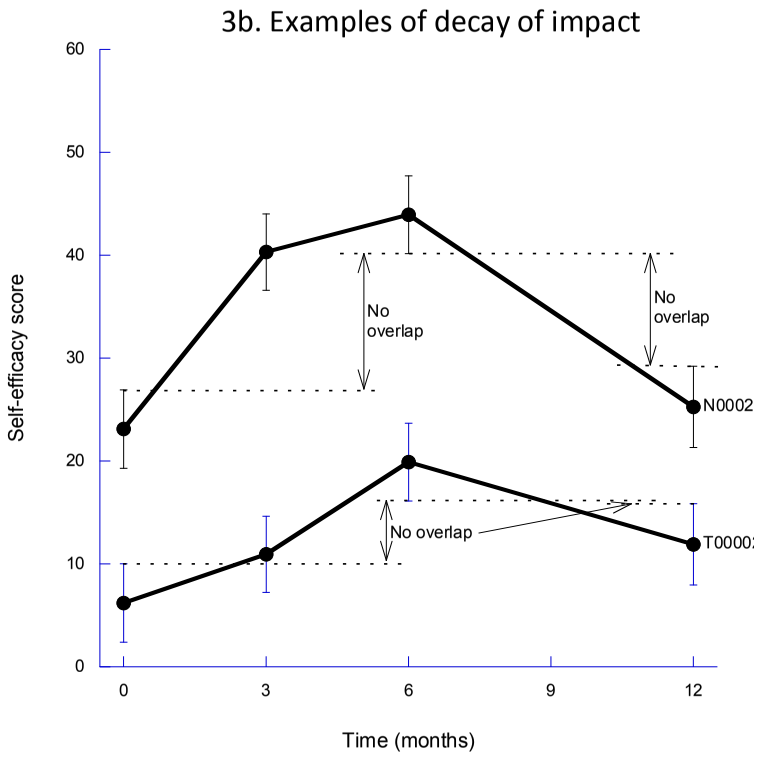
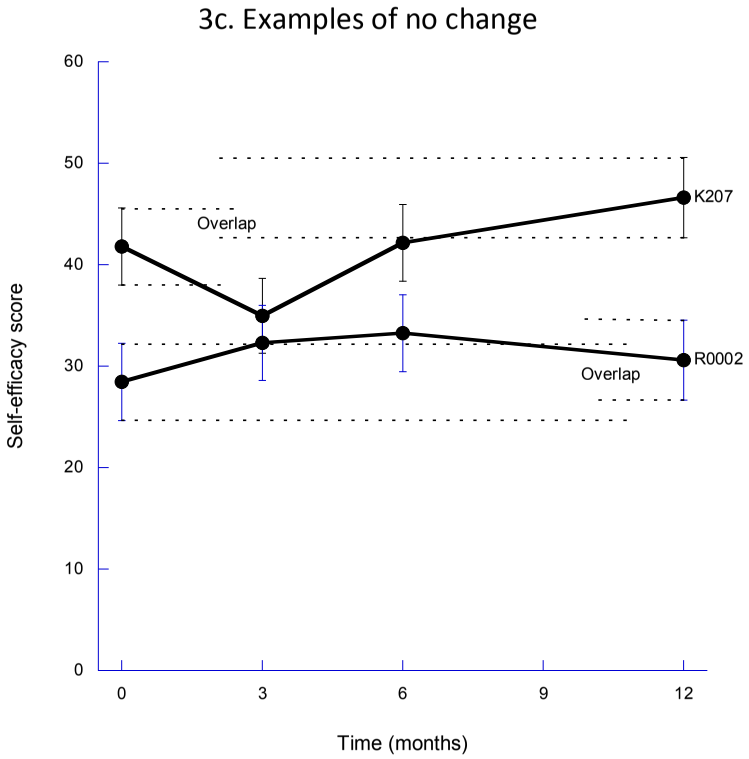


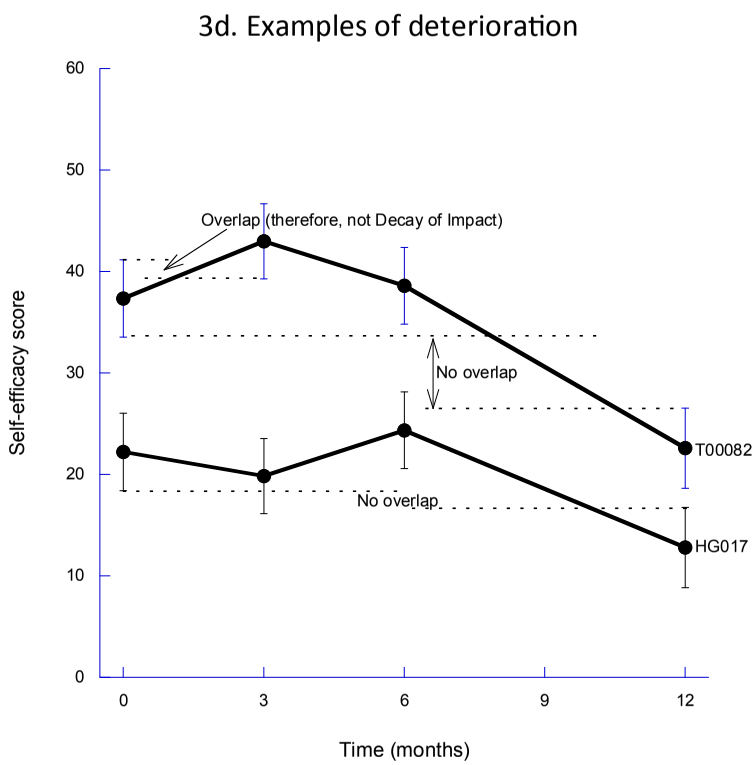
Figure 2. Data collection for longitudinal cohort study





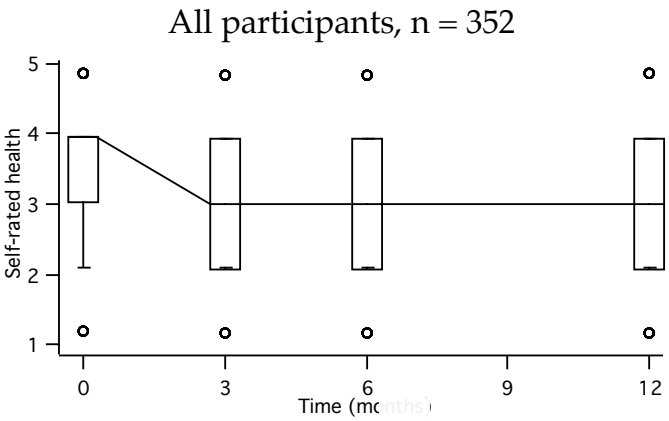
Figures 3a and 3b.
Examples of two patterns of change: Improvement and decay of impact.

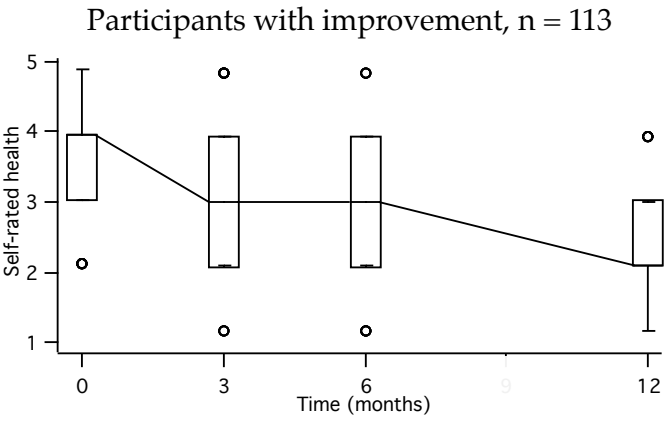


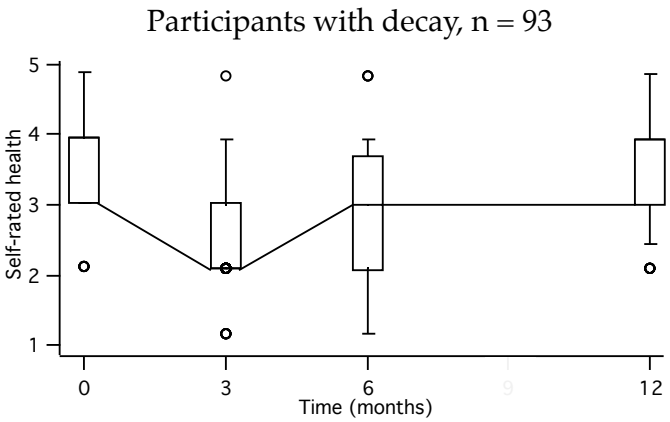


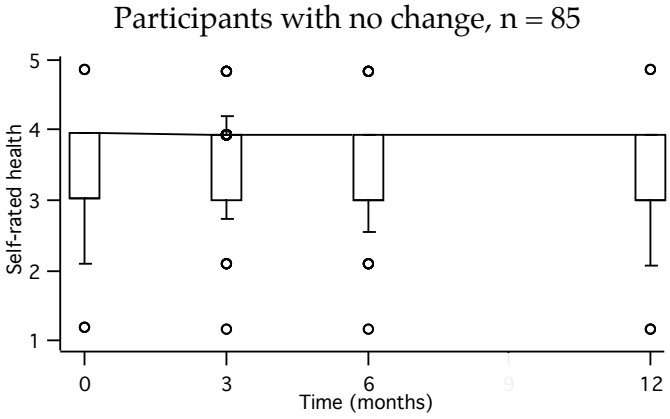
Figures 3c and 3d.
Examples of two patterns of change: No change and deterioration.

Figure 4a.
Individual-level changes over time in self-rated health
Range: 1-5
Lower scores are better.









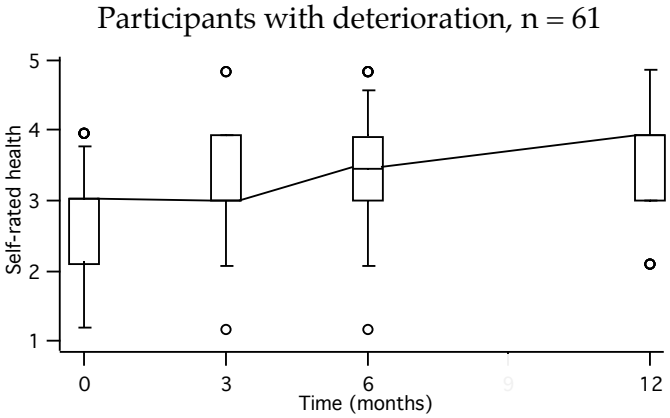
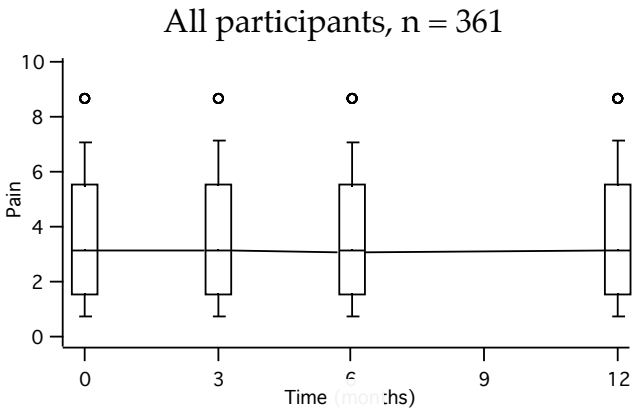
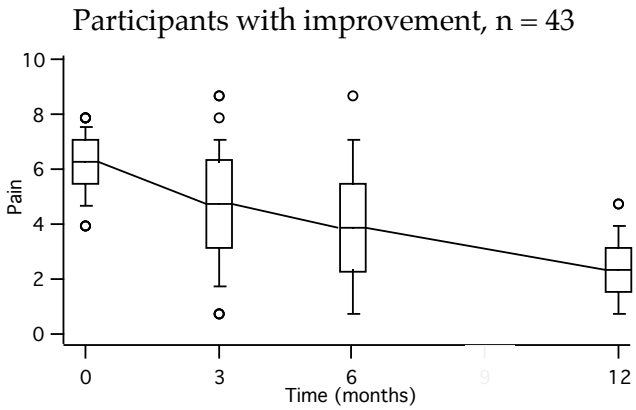
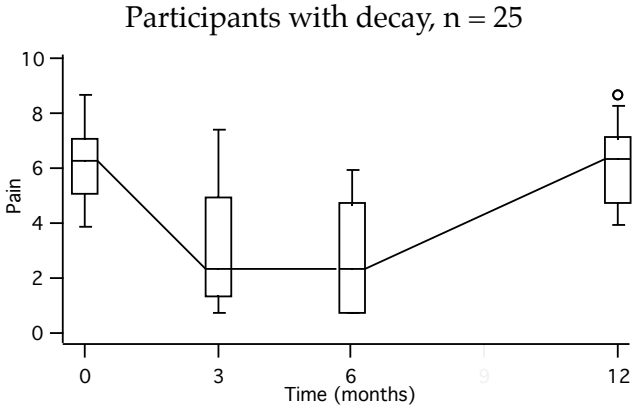
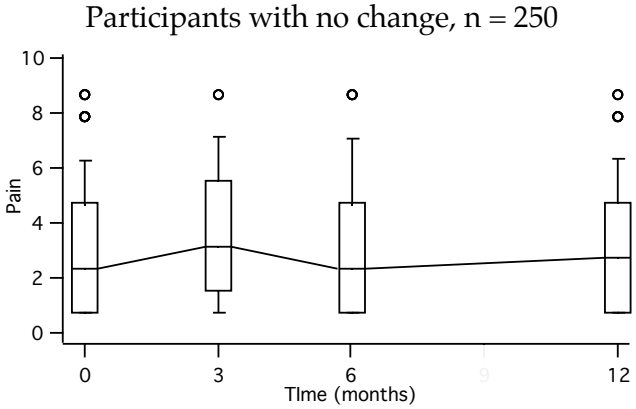


Figure 4b.
Individual-level changes over time in pain
Range: 0-10
Lower scores are better.









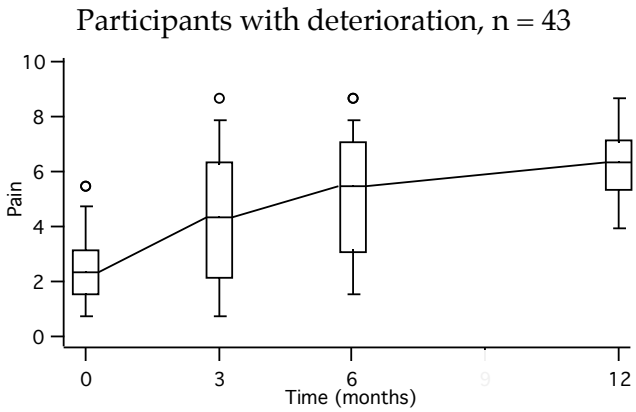
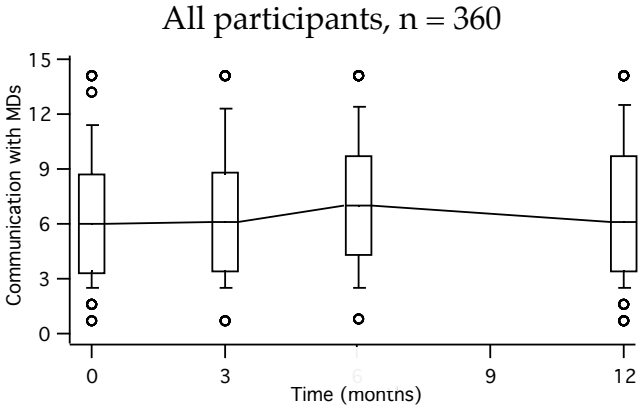
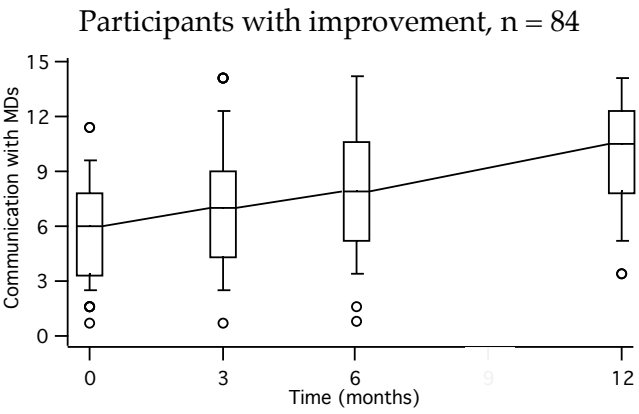
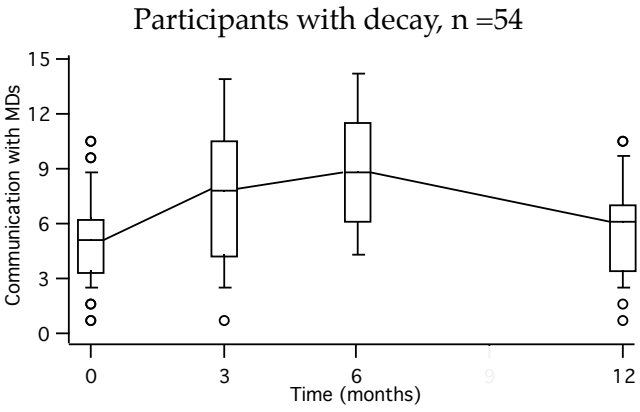
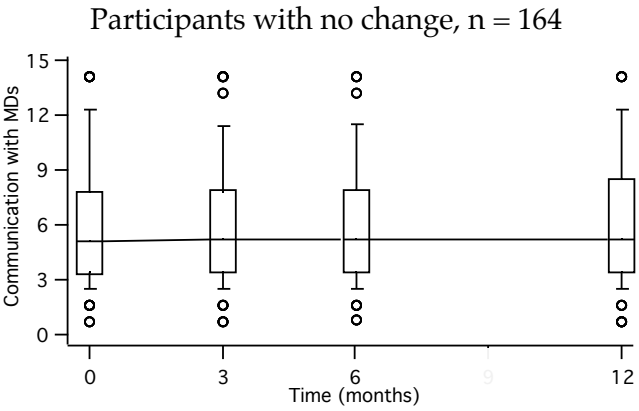


Figure 4c.
Individual-level changes over time in communication with MDs
Range: 0-15
Higher scores are better.









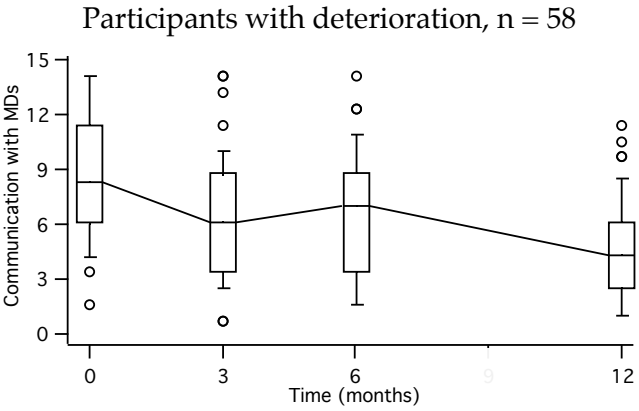
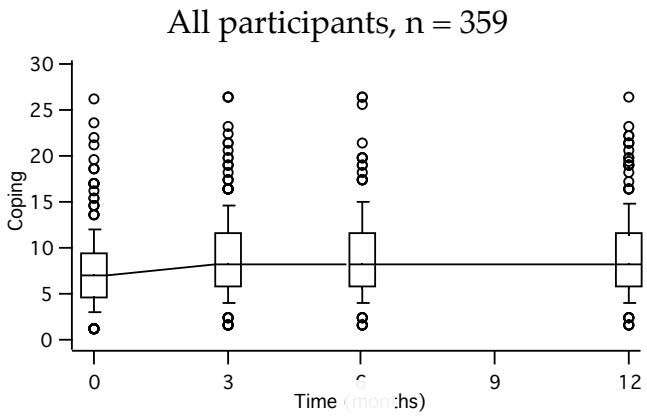
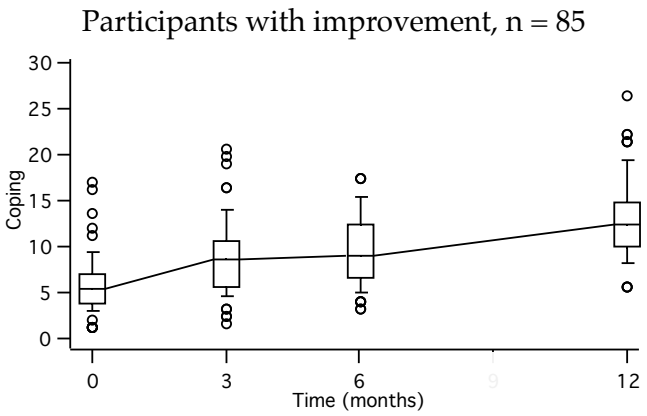
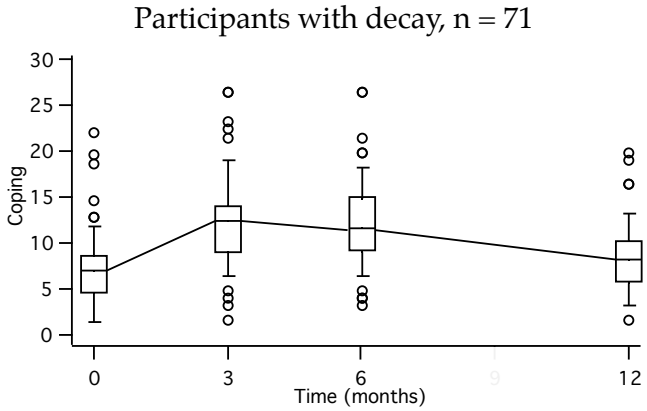
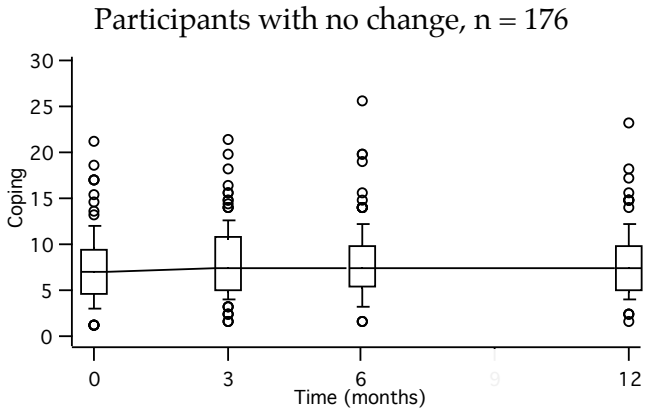


Figure 4d.
Individual-level changes over time in coping
Range: 0-30
Higher scores are better.









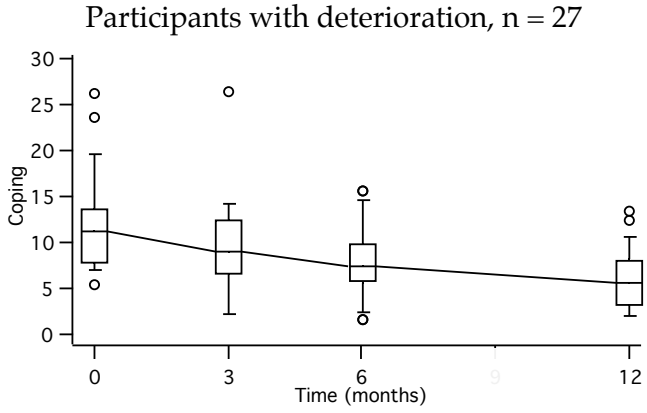
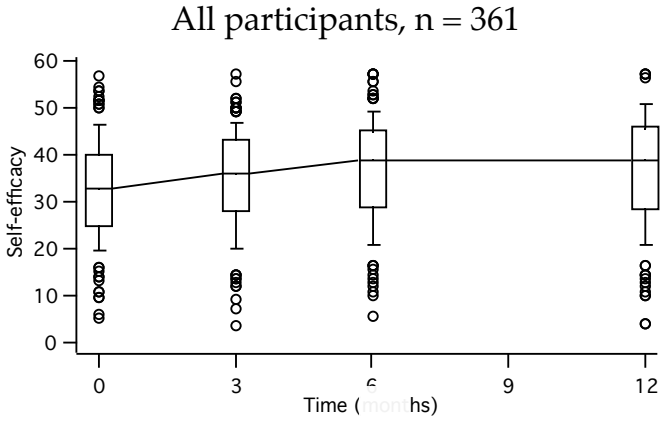
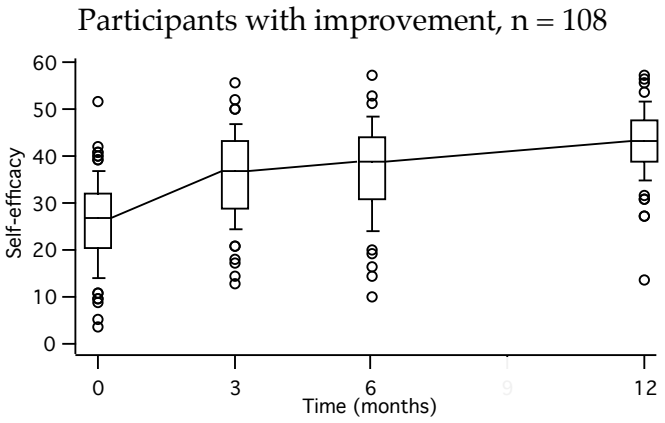
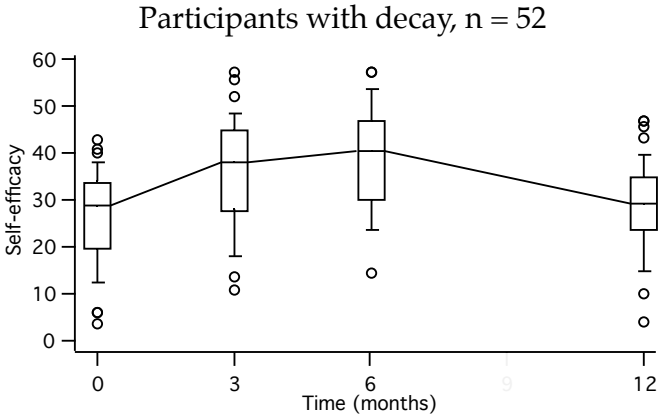
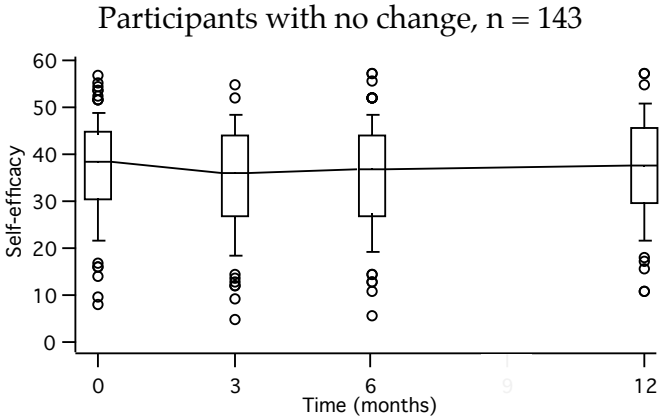


Figure 4e.
Individual-level changes over time in self-efficacy
Range: 0-60
Higher scores are better.









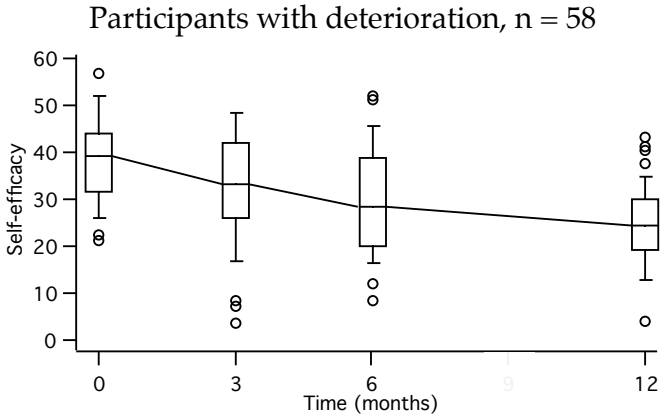
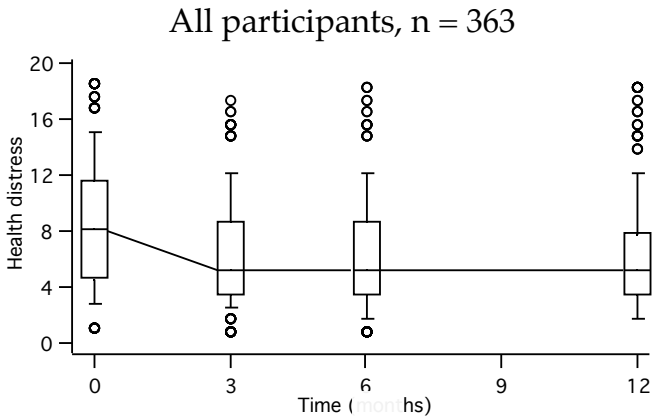
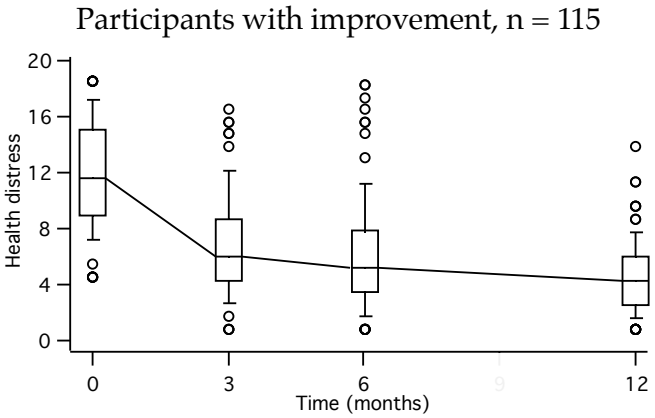
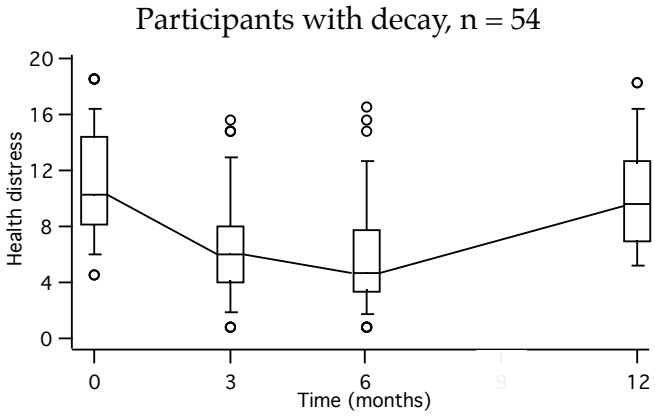
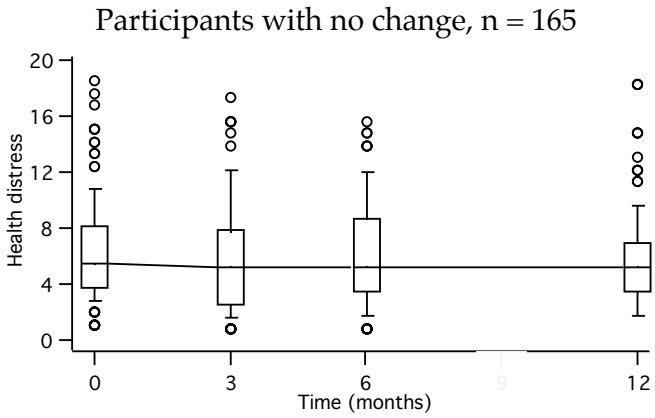


Figure 4f.
Individual-level changes over time in health distress
Range: 0-20
Lower scores are better.









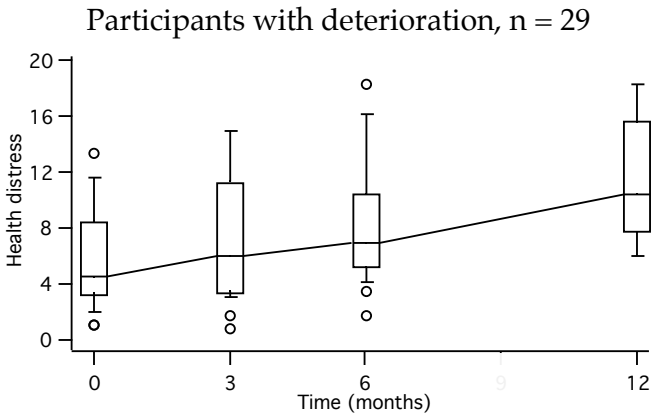
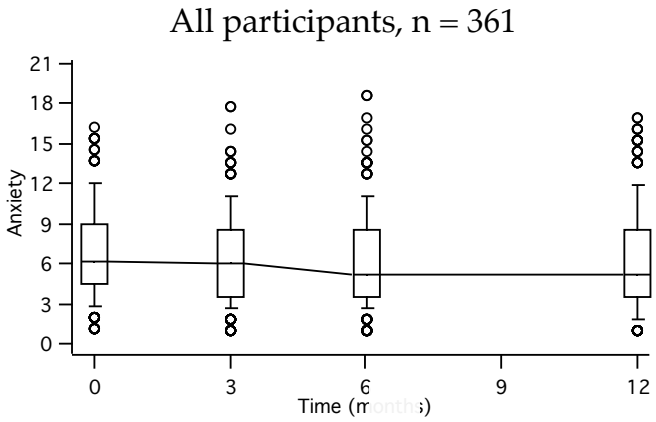
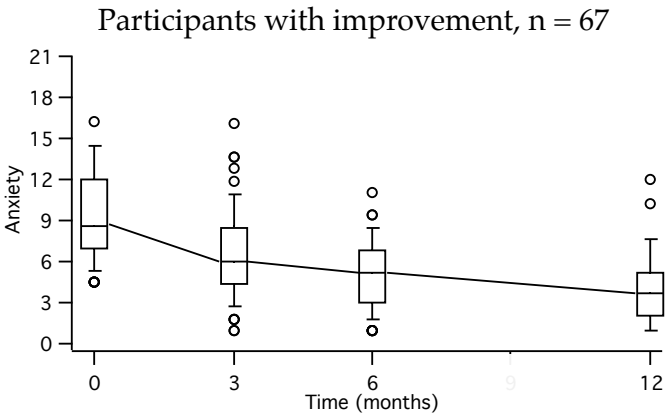
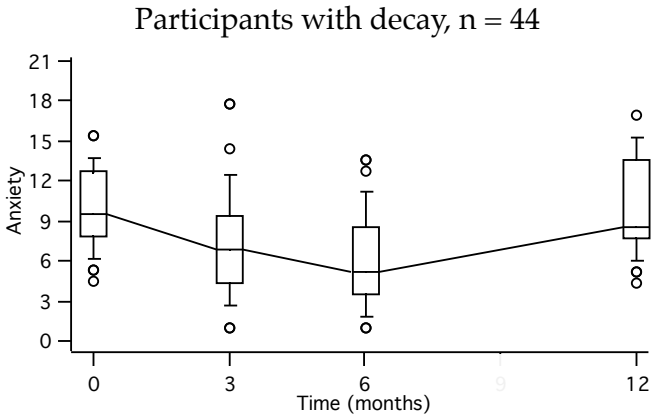
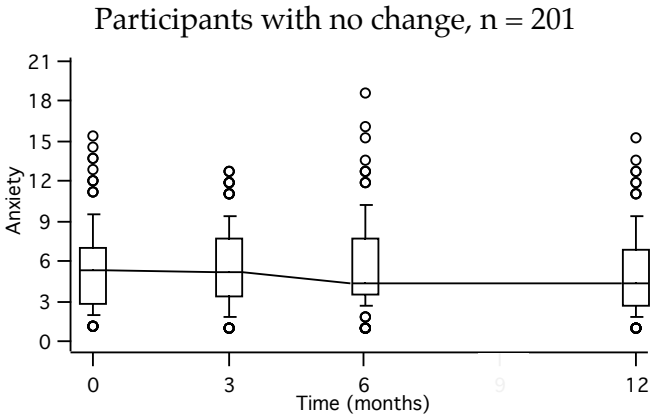


Figure 4g.
Individual-level changes over time in anxiety
Range: 0-21
Lower scores are better.









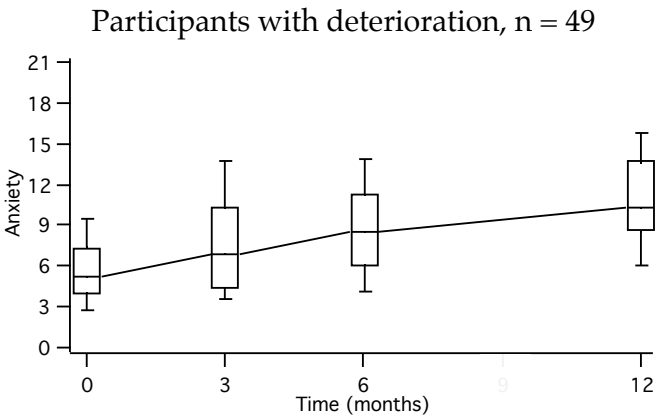
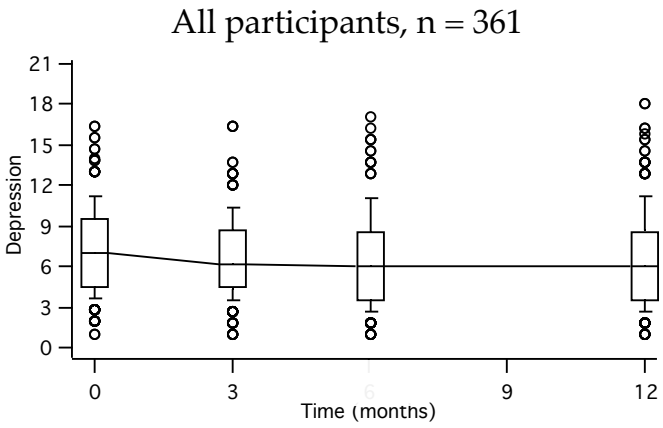
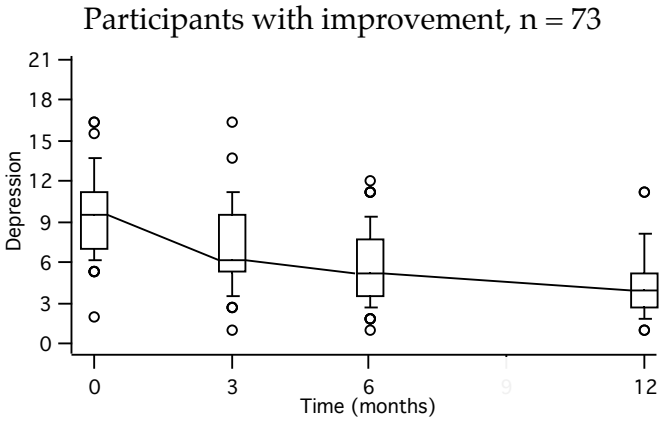
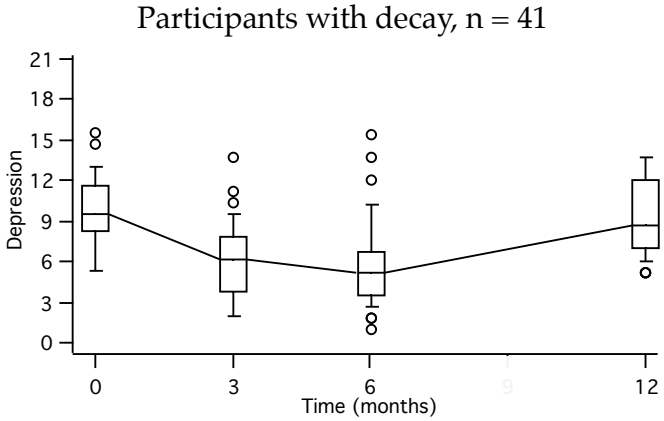
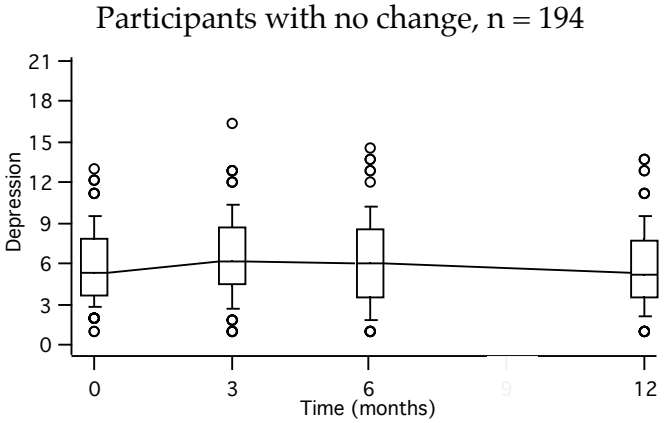


Figure 4h.
Individual-level changes over time in depression
Range: 0-21
Lower scores are better.









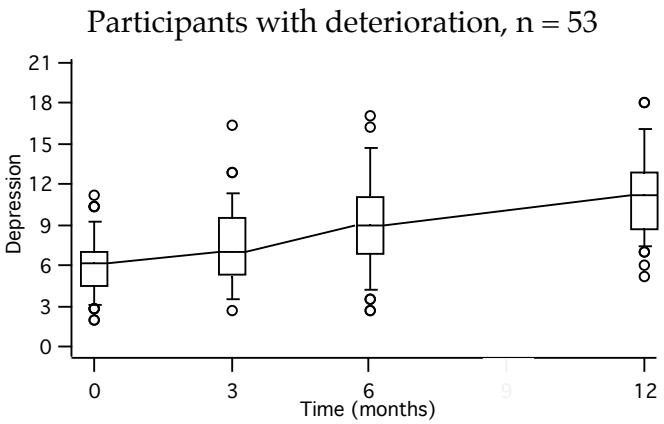


Figure 5. Percentages of participants who had each of the four patterns of change.

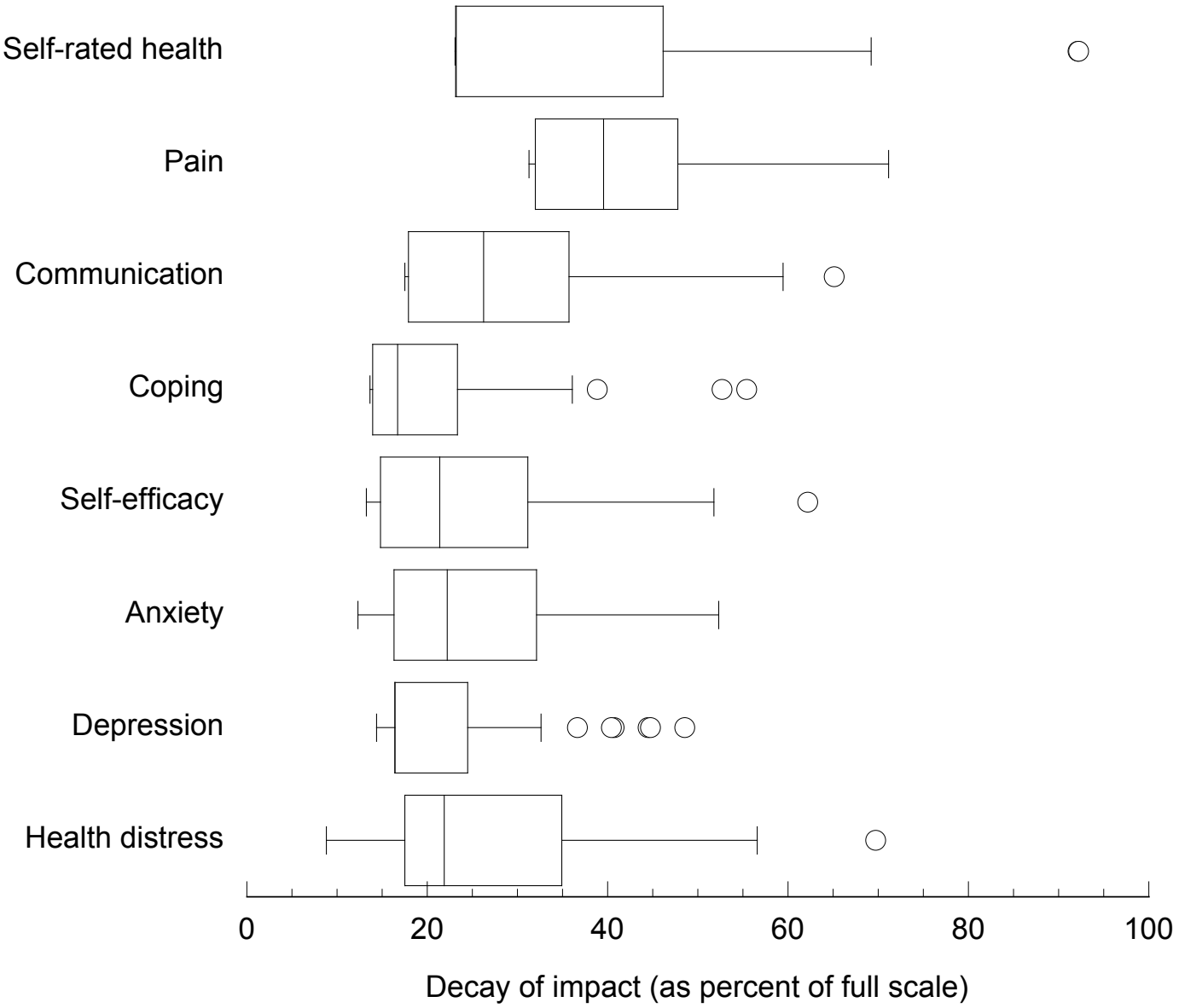


Figure 6a. Boxplots showing magnitudes of decay of impact as percentages of each measure's full-scale value.

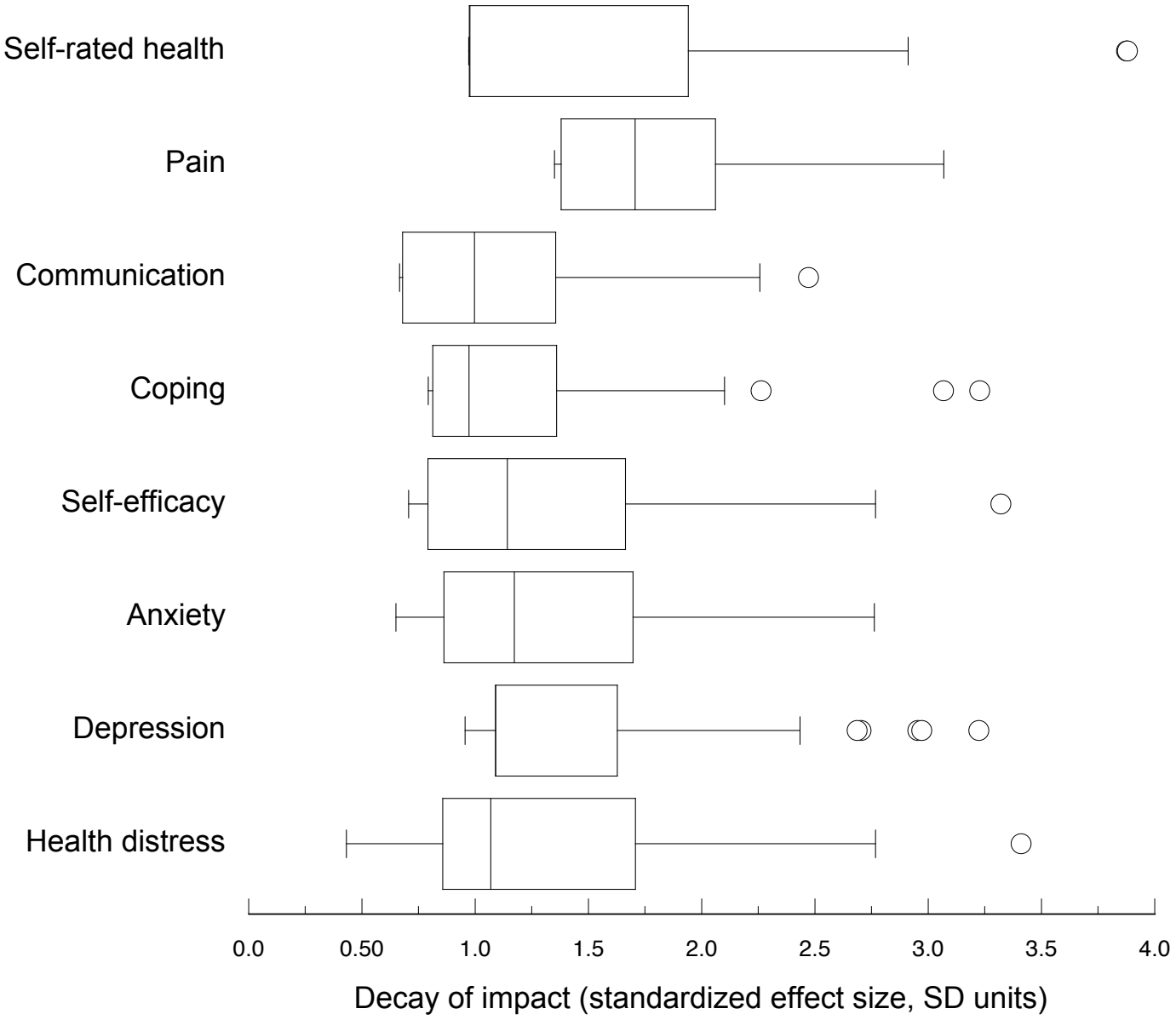


Figure 6b. Boxplots showing magnitudes of decay of impact as standardized effect sizes. The smallest median standardized effect size was 0.97 (for Coping). Standardized effect sizes greater than 0.8 are considered to be "large."

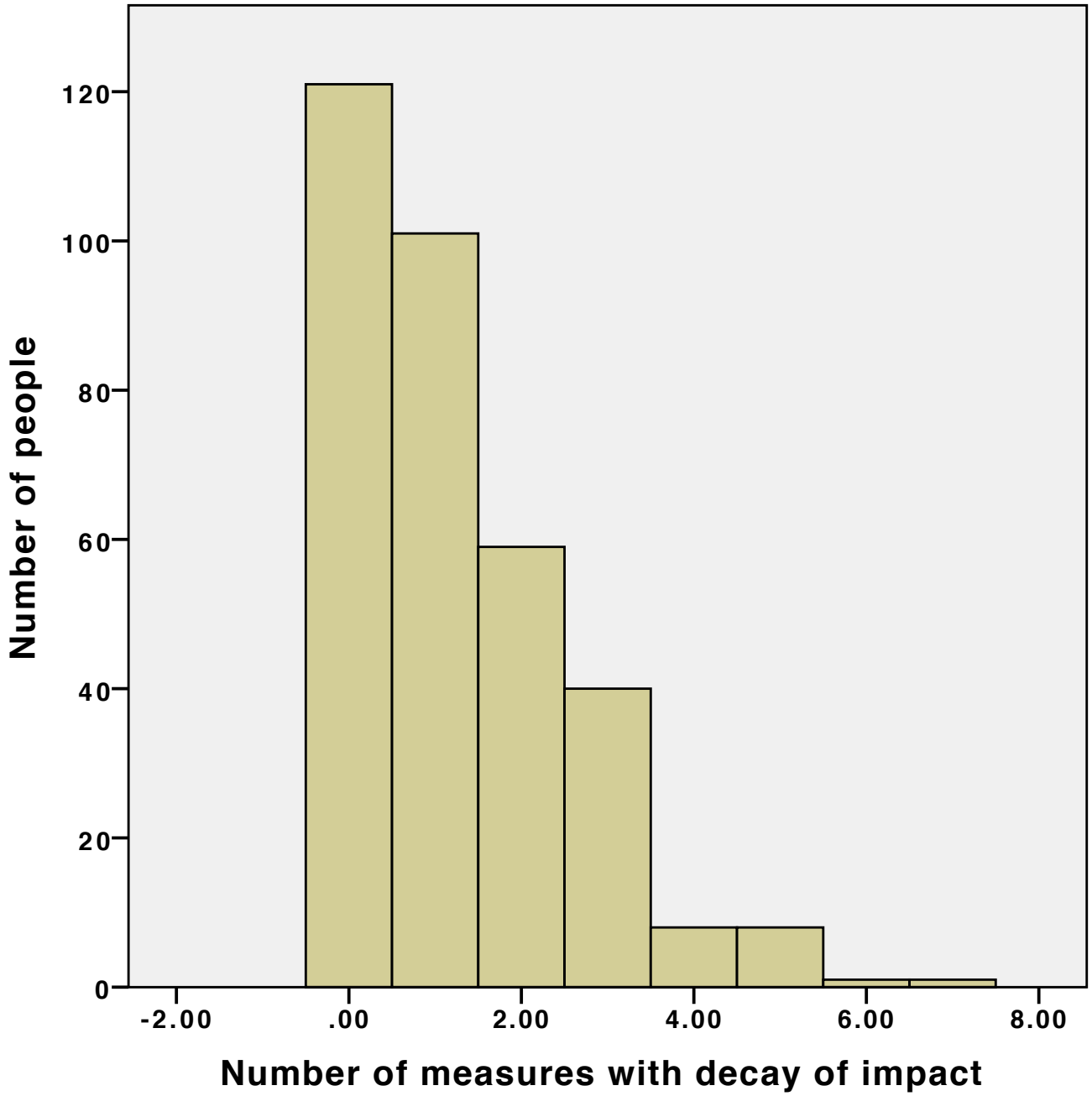


Figure 7. Frequency distribution of the number of outcomes with decay of impact.

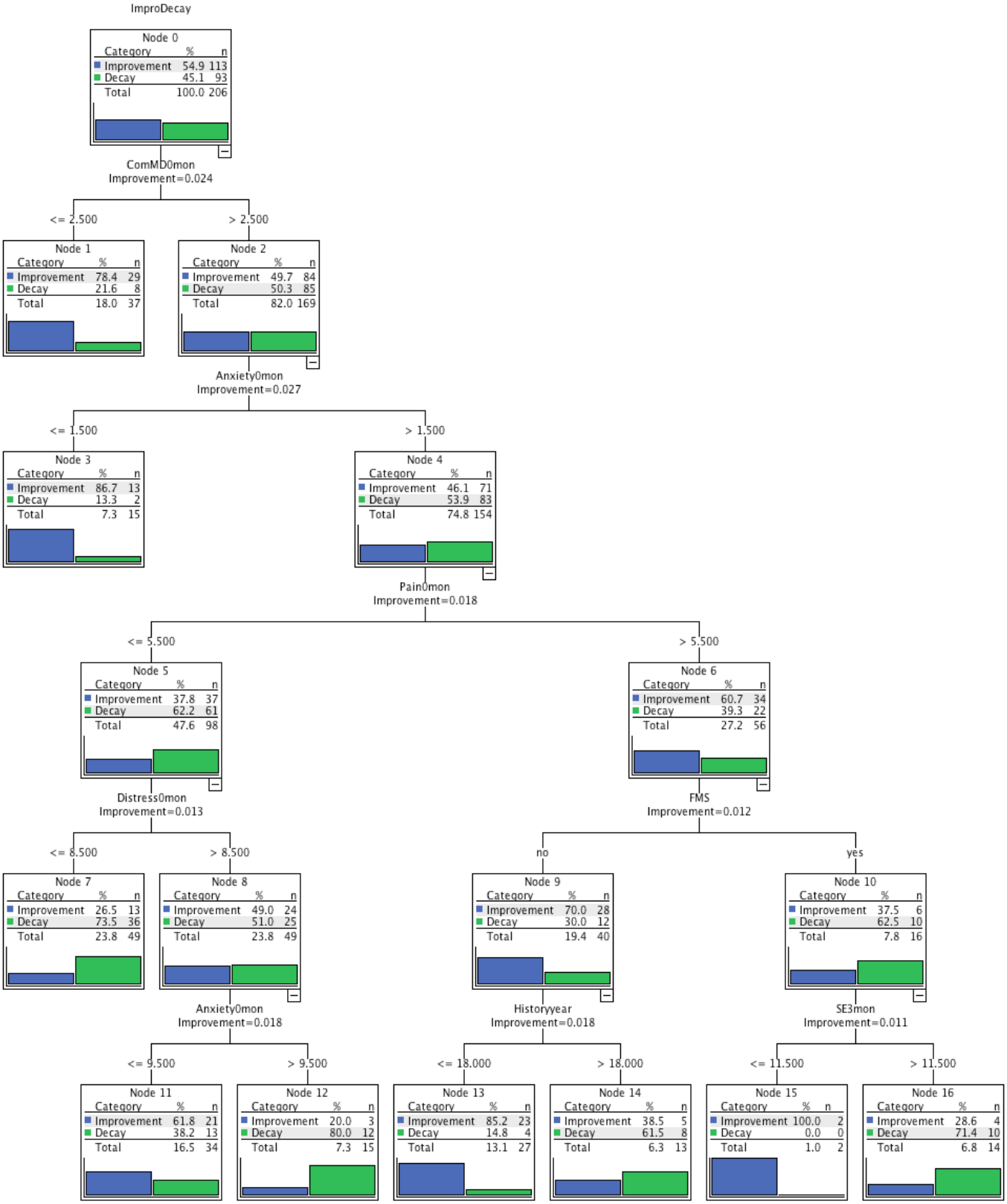


Figure 8a. Classification tree for self-rated health.

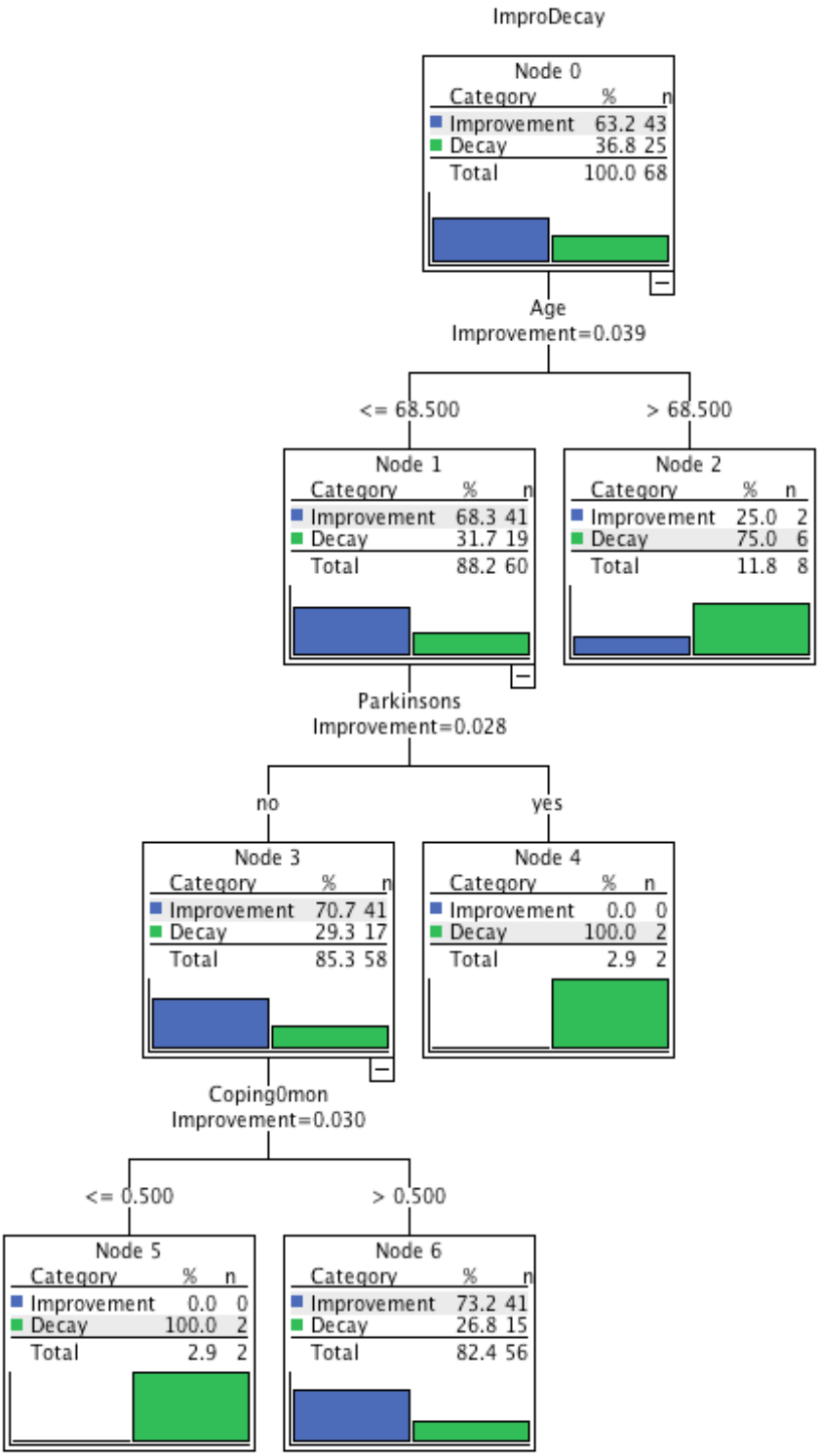


Figure 8b. Classification tree for Pain

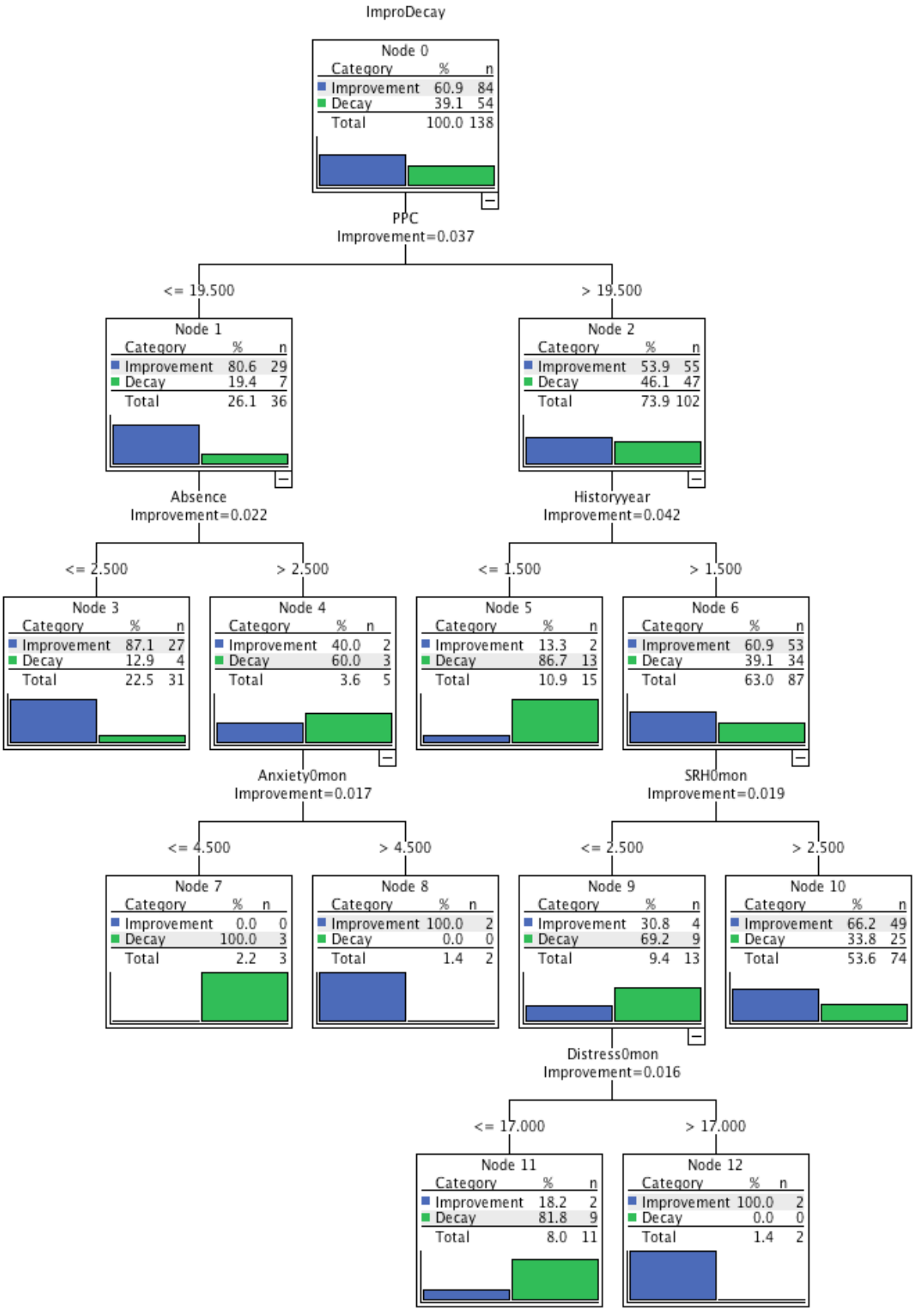


Figure 8c. Classification tree for Communication

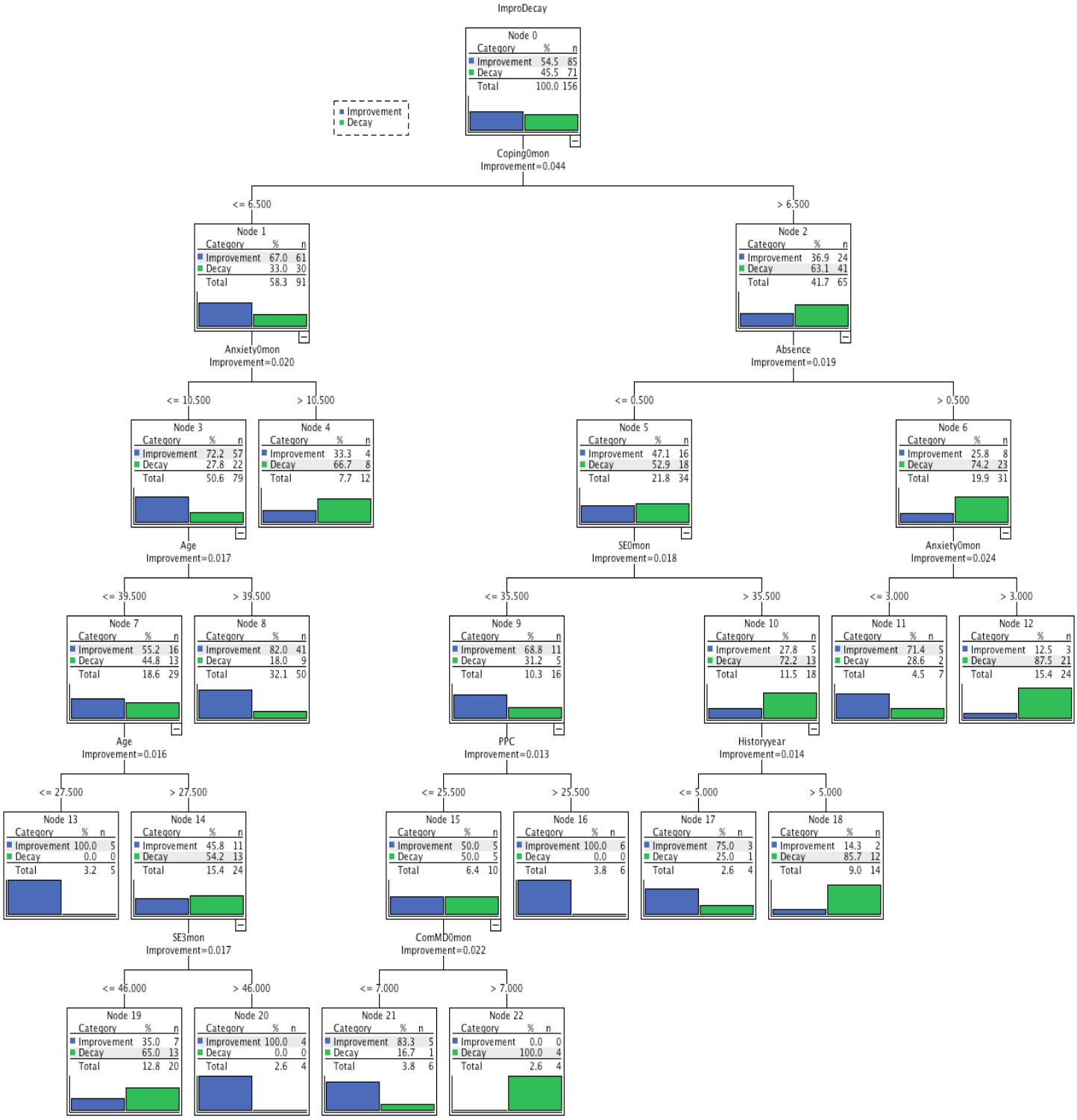


Figure 8d. Classification tree for Coping

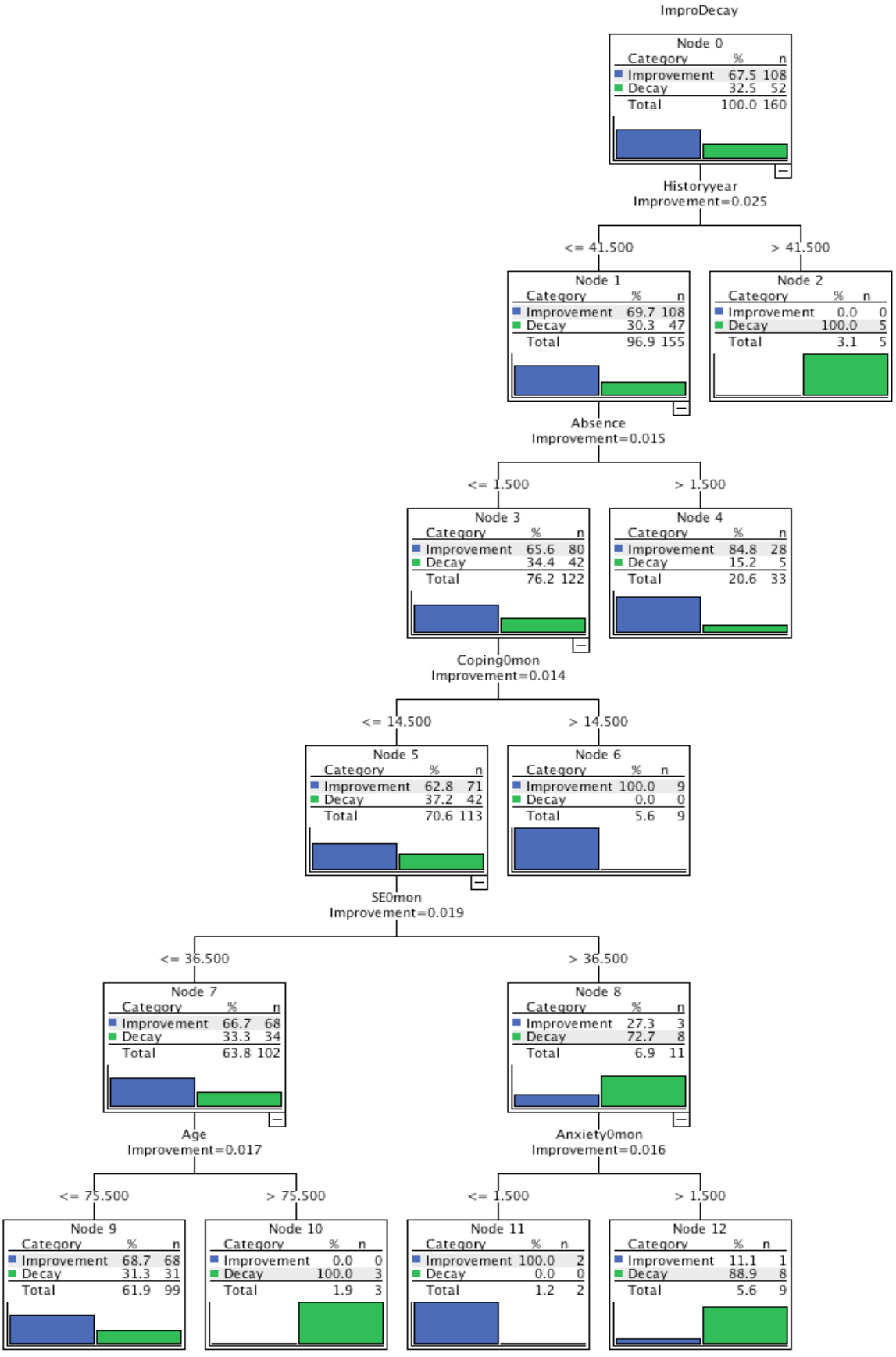


Figure 8e. Classification tree for self-efficacy

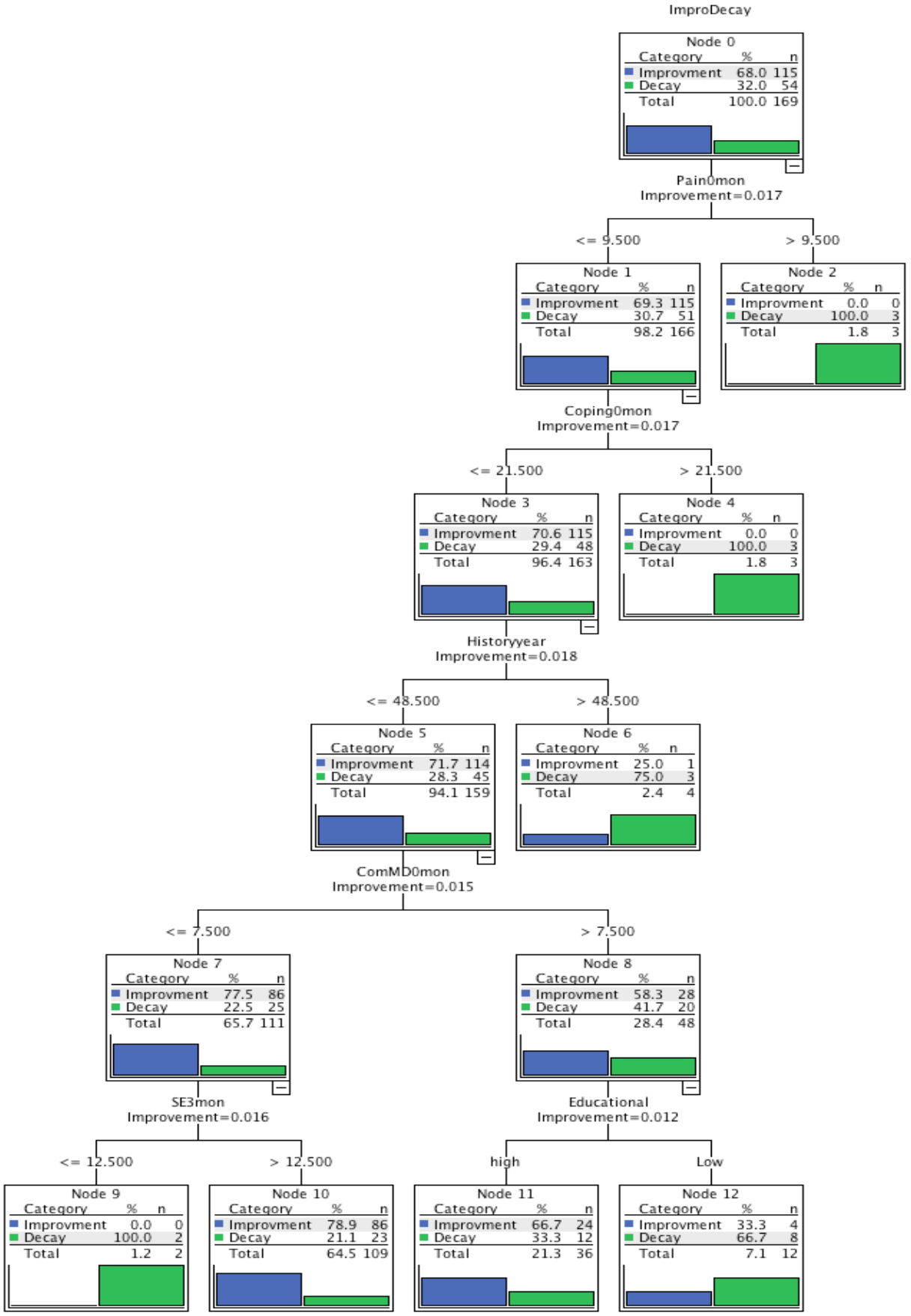


Figure 8f. Classification tree for Health distress.

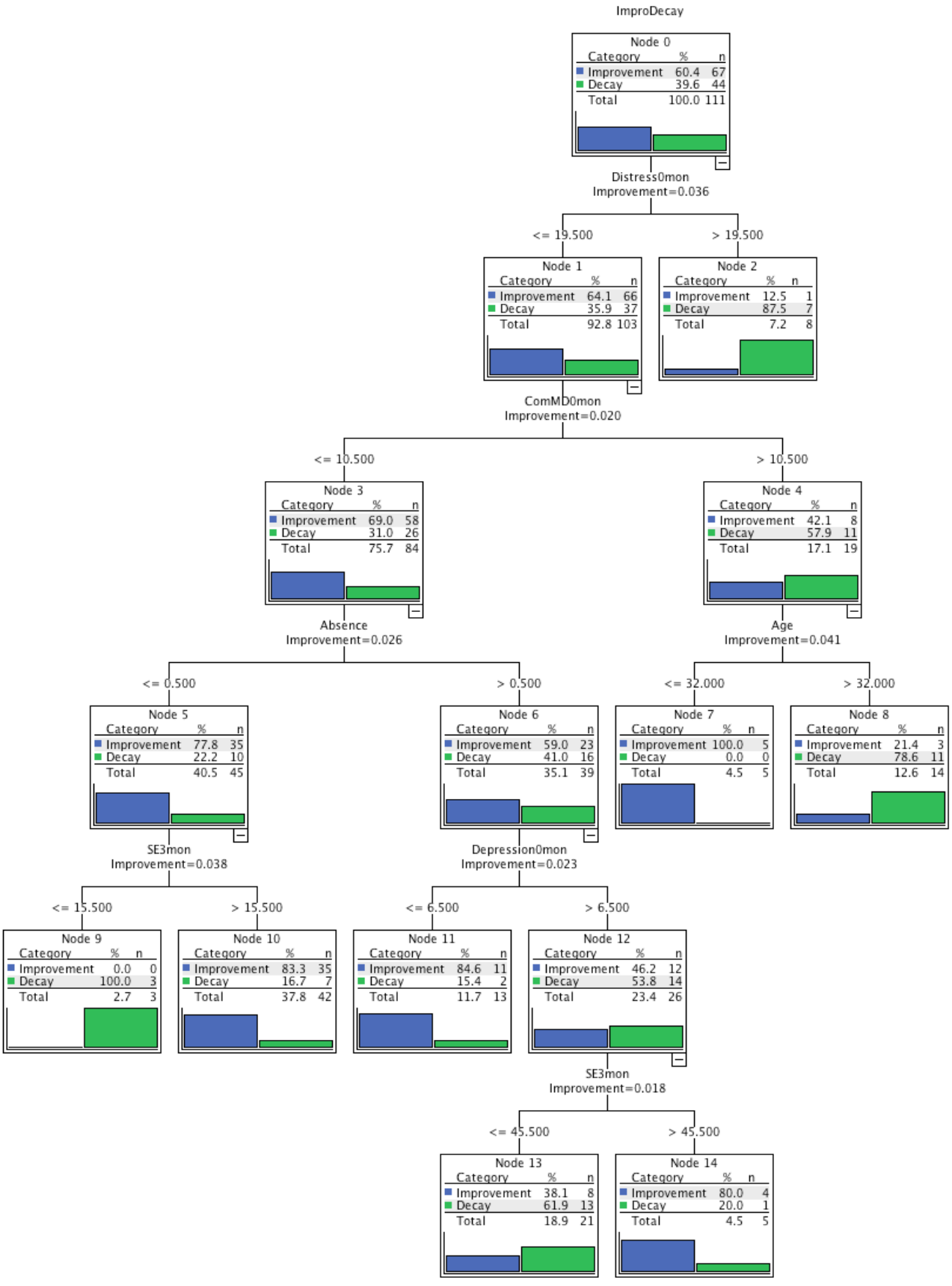


Figure 8g. Classification tree for anxiety.

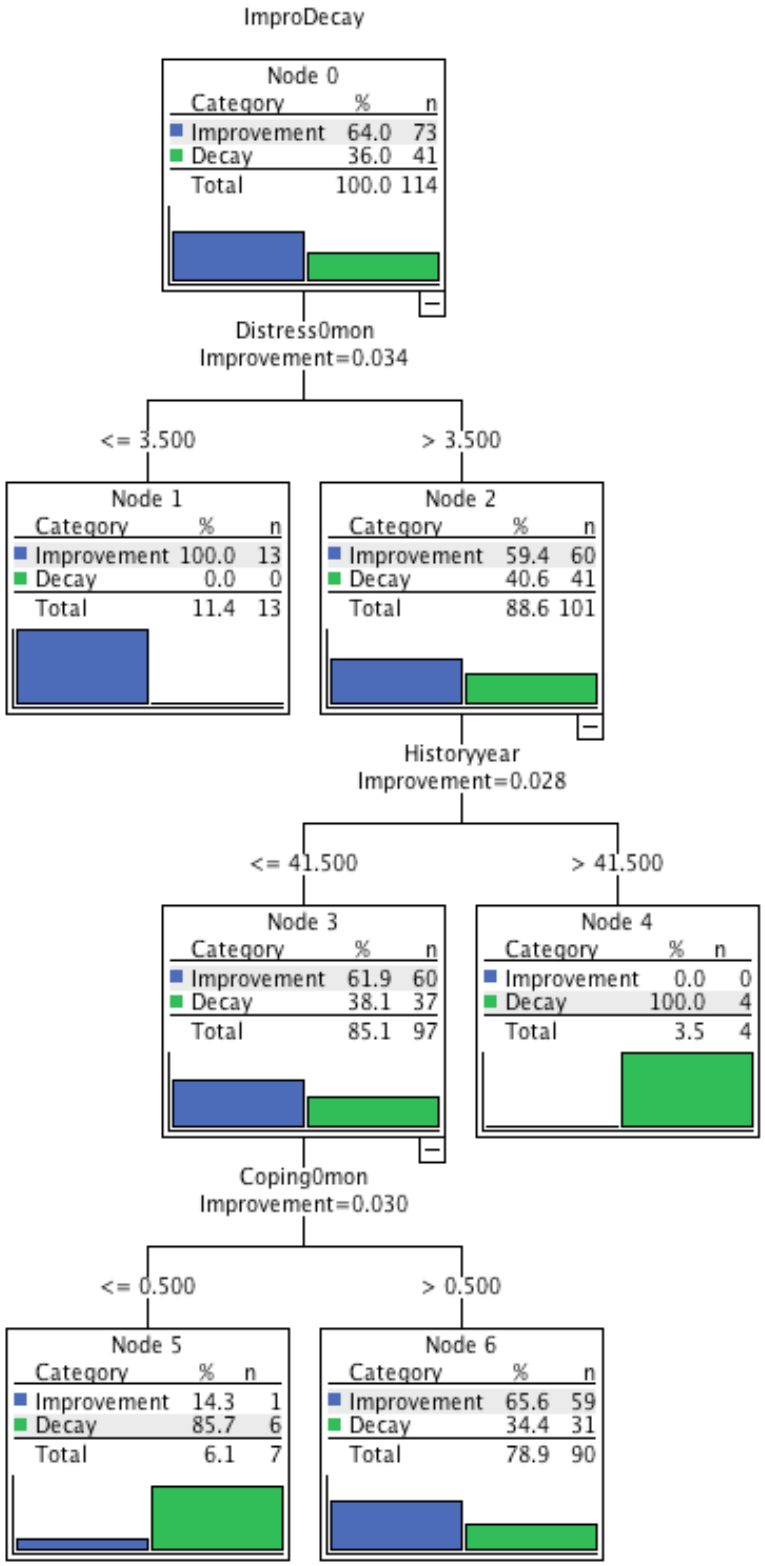


Figure 8h. Classification tree for depression.

Appendix 1. Ethics-committee approval form

様式第２号

倫　理　委　員　会
審　査　結　果　報　告　書

平成２０年７月９日

申請者　健康社会学
准教授
山崎喜比古　殿

大学院医学系研究科・医学部
倫理委員会
委員長　赤林　朗　印

受付番号　１４７２－（３）

研究課題　慢性疾患セルフマネジメントプログラムのプロセスおよび効果に関する評価研究

研究者　山崎喜比古、Fusae Kondo Abbott、米倉佑貴、湯川慶子、神内謙至、沖野露美、小野万里子、本間三恵子、朴敏延、香川由美

上記研究計画を No.１４７２－（２）の軽微な変更と認めます。
ここに通知致します。

判　定 ○承認する。　　承認しない。 条件付きで承認する。　　該当しない。 変更を勧告する。
条件あるいは変更勧告の理由（細則第３条第２項）

ID

Appendix 2. Informed-consent form.

研究参加同意書

研究者代表
東京大学大学院医学系研究科　健康科学・看護学専攻
健康社会学分野　准教授　山崎　喜比古

私は「慢性疾患セルフマネジメントプログラムに関する調査」について、
下記の項目について十分な説明を受け、理解しましたので、
調査に協力することに同意します。

1．調査の目的について

2．調査の方法について

3．予想される問題とその対応について

4．プライバシーの保護について

5．調査に参加しない場合でも不利益を受けないことについて

6．同意した後でも、随時これを撤回できることについて

【ご記入欄】

平成　　年　　月　　日

ふりがな

御名前 ＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿

ご住所　〒＿＿＿＿＿＿＿＿

都道府県 ＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿

電話番号 ＿＿＿＿＿－＿＿＿＿＿－＿＿＿＿＿

※注意　上記のご住所・電話番号に研究室より連絡をさしあげることがあります。
別の連絡先をご希望の場合には、こちらにお書き下さい。

《連絡はこちらに》

ご 住　所　〒＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿

電話番号 ＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿

Appendix 3. Baseline questionnaire

慢性疾患セルフマネジメントプログラム に関する調査

東京大学大学院医学系研究科　健康社会学教室

●●●お問合わせ先●●●

ご質問やご不明な点がございましたら、下記までお問合わせください。

東京大学大学院医学系研究科　健康社会学教室

セルフマネジメントプログラム評価研究チーム

担当：朴敏廷（パクミンジョン）・湯川慶子（ゆかわけいこ）

電話：０３－５８４１－３５１４

FAX：０３－５６８４－６０８３

E メール：mjpark-tky@umin.ac.jp（受付時間：平日 10 時～17 時）

このアンケートにお答えいただいている今日の日付：＿＿＿＿年＿＿＿＿月＿＿＿＿日

【あなたご自身のことについておうかがいします】

［　　］のなかで、あてはまるものひとつに○をつけてください。（　）には具体的にお書きください。

1．　あなたの性別・年齢を教えてください。

男性・女性　　（　　歳）

2．あなたの出身国を教えてください。

1．日本　　2．その他（　　）

3．あなたの最終学歴を教えてください。

1．小学校　2．中学校　3．高校　4．専門学校　5．短大
6．大学　7．大学院　8．その他（　　）

4．現在の婚姻状況を教えてください。

1．未婚　2．既婚同居　3．既婚別居　4．離婚　5．死別

5．あなたは、慢性疾患をお持ちですか。ご家族・医療従事者の方もお答えください。

1．慢性疾患がある　　2．慢性疾患はない

次のページへお進みください

■あなたには次の疾患がありますか？　（すべてにお答えください）

1．糖尿病　**なし・あり**　1．1型　2．2型　3．その他（　　）

2．喘息　**なし・あり**

3．肺気腫あるいは慢性閉塞性肺疾患　**なし・あり**

4．その他の肺疾患　**なし・あり**（診断名：＿＿＿＿＿＿＿＿）

5．心疾患　**なし・あり**　1．高血圧　2．高脂血症　3．その他（診断名：＿＿＿＿＿＿）

6．膠原病・ﾘｳﾏﾁ性疾患　**なし・あり**　1．膠原病　2．関節ﾘｳﾏﾁ　3．その他（診断名：＿＿＿＿＿＿）

7．がん　**なし・あり**（診断名：＿＿＿＿＿＿＿＿）

8．アレルギー性鼻炎その他の耳鼻疾患　**なし・あり**（診断名：＿＿＿＿＿＿＿＿）

9．アトピー性皮膚炎その他の皮膚疾患　**なし・あり**（診断名：＿＿＿＿＿＿＿＿）

10．1～9以外の慢性疾患　**なし・あり**（診断名：＿＿＿＿＿＿＿＿）

■あなたは慢性疾患をかかえて何年になりますか。
（複数の疾患がある方は最も長いものでお答え下さい）　［　　］**年**　［　　］**ヶ月**

【全般的な健康状態について】

あなたの健康状態は、全般的に見て、いかがですか。（あてはまる番号ひとつに○）

1．とてもよい
2．よい
3．普通
4．ややおもわしくない
5．おもわしくない

【症状について】

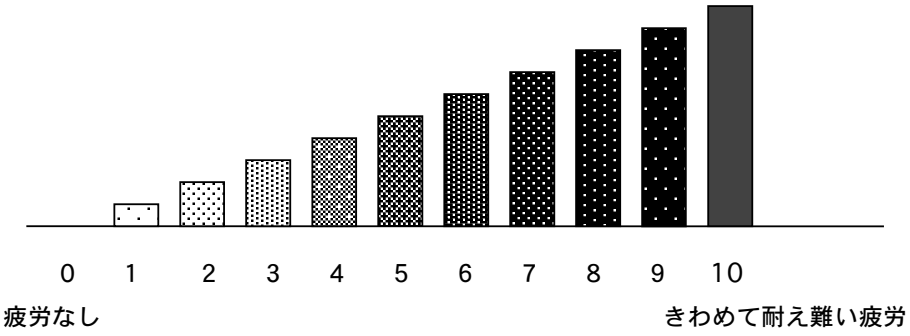
ここ１ヵ月間、次のように思ったり感じたりしたことがどのくらいの頻度でありましたか。

（それぞれあてはまる番号ひとつに○）

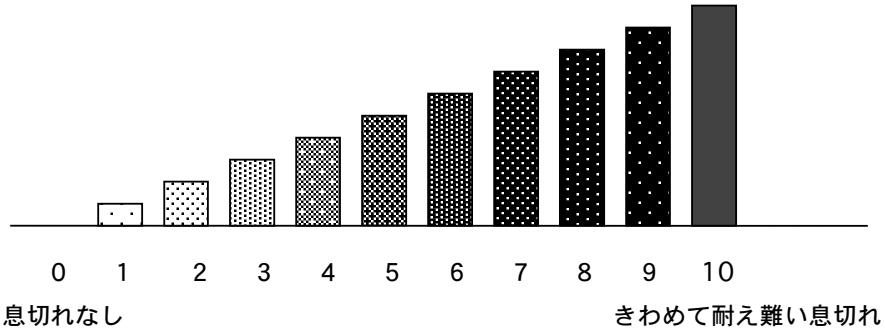
	全くなかった	たまにあった	時々あった	よくあった	ほとんどいつもあった	いつもあった
1）自分の健康上の問題で落ち込むことがあった	0	1	2	3	4	5
2）自分の将来の健康状態を考えると怖くなることがあった	0	1	2	3	4	5
3）自分の健康状態は人生における心配事のひとつだと思うことがあった	0	1	2	3	4	5
4）自分の健康状態は期待どおりにいっていないと感じることがあった	0	1	2	3	4	5

ここ２週間の状態についてうかがいます。

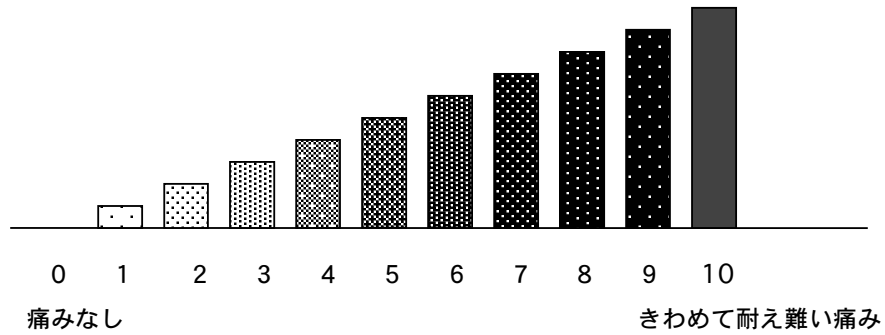
１）ここ２週間のあなたの疲労の程度について、下の図であてはまる数字ひとつに○をつけてください。



２）ここ２週間のあなたの息切れの程度について、下の図であてはまる数字ひとつに○をつけてください。



３）ここ２週間のあなたの痛みの程度について、下の図であてはまる数字ひとつに○をつけてください。



【運動について】

あなたは次の運動を、ここ１週間で合計何時間くらい行いましたか。
（その１週間がいつも通りでなかったとしても、かまわずお答え下さい。）

（それぞれあてはまる番号ひとつに○）

	なし	30 分未満/週	30～60 分/週	1～3 時間/週	3 時間以上/週
1）ストレッチまたは筋力トレーニング	0	1	2	3	4
2）ウォーキング	0	1	2	3	4
3）水泳またはアクアエクササイズ	0	1	2	3	4
4）サイクリング（エアロバイクも含む）	0	1	2	3	4
5）その他のマシンを使用した有酸素運動	0	1	2	3	4
6）その他の有酸素運動	0	1	2	3	4

（具体的に＿＿＿＿＿＿＿＿＿＿）

【症状への対処について】

気分が落ち込んだり、痛みや他の不快な症状があるとき、あなたはどのように対処していますか。

（それぞれあてはまる番号ひとつに○）

	全くしない	たまにする	時々する	よくする	ほとんどいつもする	いつもする
1）不快な症状から離れて、自分の体の一部ではないと感じるよう努める	0	1	2	3	4	5
2）不快な症状として考えずに、温かいとか無感覚だという何か他の感覚として考える	0	1	2	3	4	5
3）不快な症状から気をまぎらわすために、頭の中でゲームをしたり、歌を歌ったりする	0	1	2	3	4	5
4）部分ごとに体の筋肉をリラックスさせる	0	1	2	3	4	5
5）自分がどこか別のところにいるような想像をしたり、音声に導かれるイメージ法を行う	0	1	2	3	4	5
6）物事を前向きに考えるようにする	0	1	2	3	4	5

【日常の動作について】

現時点で、あなたは次のことをどのくらいできますか。　（それぞれあてはまる番号ひとつに○）

	何の困難もない	いくらか困難	かなり困難	できない
1）自分で身支度ができますか？（靴ひもを結ぶこと、ボタン掛けも含む）	0	1	2	3
2）就寝・起床の動作ができますか？	0	1	2	3
3）いっぱいに水が入っている茶碗やコップを口元まで運べますか？	0	1	2	3
4）屋外で平坦な道を歩くことができますか？	0	1	2	3
5）全身を洗い、タオルで拭くことができますか？	0	1	2	3
6）腰をまげ床にある衣類を拾うことができますか？	0	1	2	3
7）蛇口の開け閉めができますか？	0	1	2	3
8）車の乗り降りができますか？	0	1	2	3

【日常生活について】

ここ４週間、次の活動をする際、どのくらいの支障がありましたか。

（それぞれあてはまる番号ひとつに○）

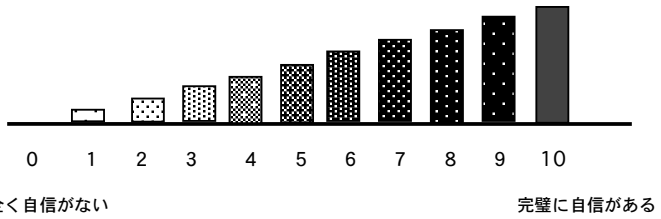
	全くなかった	少しあった	時々あった	よくあった	いつもあった
1）健康上の問題によって、家族・友人・近隣の人たちとのふだんの社会生活に支障がありましたか？	0	1	2	3	4
2）健康上の問題によって、趣味または娯楽に支障がありましたか？	0	1	2	3	4
3）健康上の問題によって、家事に支障がありましたか？	0	1	2	3	4
4）健康上の問題によって、用事や買物に出るのに支障がありましたか？	0	1	2	3	4

【実行できる自信について】

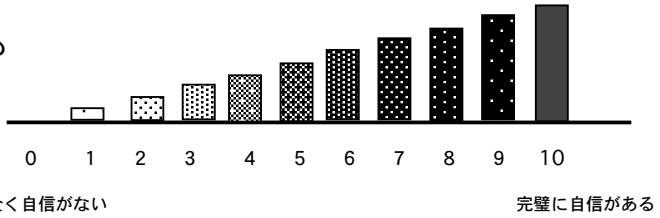
現時点で、あなたが次のことを実行できる自信はどのくらいありますか。

（それぞれあてはまる数字ひとつに○）

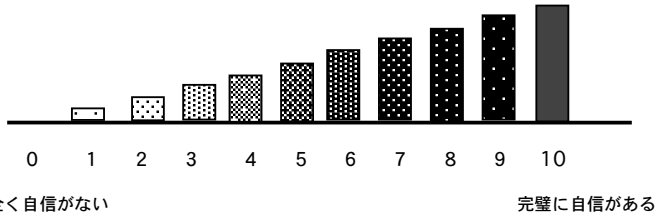
1）病気による疲労があっても
やりたいことを実行できる自信は
どのくらいありますか？



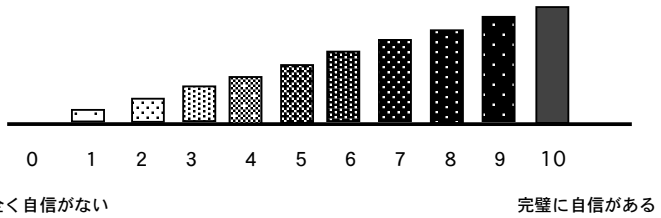
2）病気による体の不快さや痛みがあっても
やりたいことを実行できる自信は
どのくらいありますか？



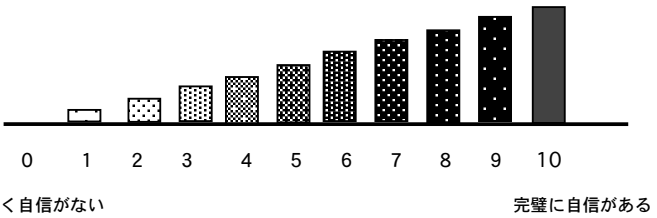
3）病気による精神的苦痛があっても
やりたいことを実行できる自信は
どのくらいありますか？



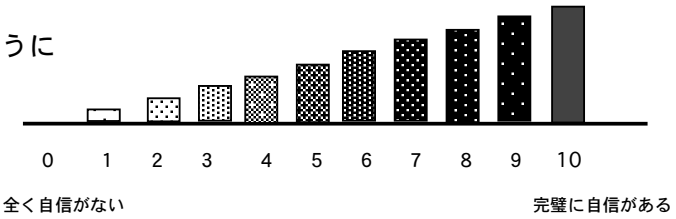
4）その他の症状や健康問題があっても
やりたいことを実行できる自信は
どのくらいありますか？



5）医師にかかる回数が減るように
あなた自身の健康管理に必要な
さまざまなことを実行できる自信は
どのくらいありますか？



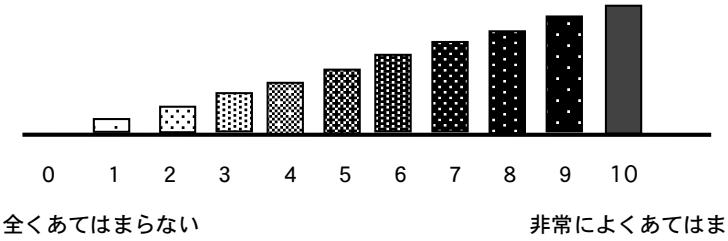
6）病気による日常生活への影響が減るように
服薬以外のことも実行できる自信は
どのくらいありますか？



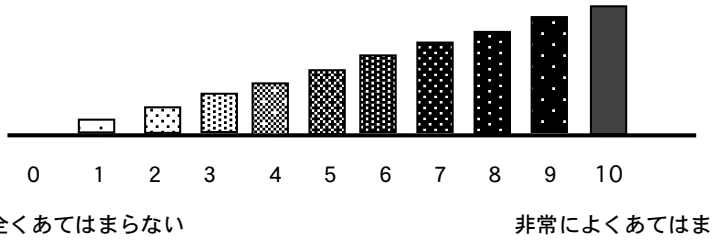
【生活の感じ方について】

あなたの人生に対する感じ方についてうかがいます。それぞれ 0 から 10 までのうち、
あなたの感じ方を最もよく表している数字ひとつに◯をつけてください。

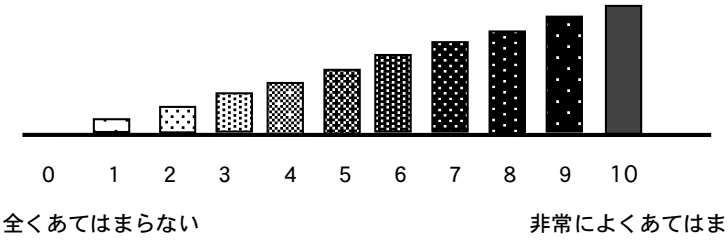
1）私は、日常生活で直面する困難や問題の
解決方法を見つけることができる



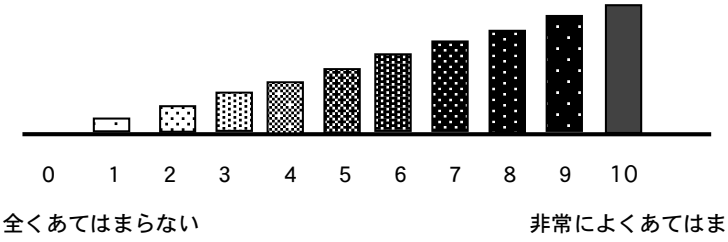
2）日常生活で直面する困難や問題の
いくつかは向き合い取り組むに値する、と
私は思える



3）私は、日常生活で生じる困難や問題を
理解したり予測したりできる



4）私の日常生活は、
喜びと満足を与えてくれる



【心の状態について】

ここ１週間のあなたのご様子についてうかがいます。
あなたはここ１週間どのように感じていますか。（それぞれあてはまる番号 ひとつに○）

１．緊張したり、気持ちが張りつめたりすることが	しょっちゅうあった	たびたびあった	ときどきあった	全くなかった
	0	1	2	3
２．むかし楽しんだことを今でも楽しいと思うことが	全く同じだけあった	かなりあった	少しだけあった	めったになかった
	0	1	2	3
３．なにか恐ろしいことが起ころうとしているという恐怖感を持つことが	しょっちゅうあって非常に気になった	たびたびあるがあまり気にならなかった	少しあるが気にならなかった	全くなかった
	0	1	2	3
４．物事の面白い面を笑ったり、理解したりすることが	いつもと同じだけできた	かなりできた	少しだけできた	全くできなかった
	0	1	2	3
５．心配事が心に浮かぶことが	しょっちゅうあった	たびたびあった	それほど多くはないが、ときどきあった	ごくたまにあった
	0	1	2	3
６．きげんの良いことが	全くなかった	たまにあった	ときどきあった	しょっちゅうあった
	0	1	2	3
７．楽に座って、くつろぐことが	かならずできた	たいていできた	たまにできた	全くできなかった
	0	1	2	3
８．仕事を怠けているように感じることが	ほとんどいつもあった	たびたびあった	ときどきあった	全くなかった
	0	1	2	3
９．不安で落ち着かないような恐怖感を持つことが	全くなかった	ときどきあった	たびたびあった	しょっちゅうあった
	0	1	2	3
１０．自分の顔、髪型、服装に関して	関心がなくなった	以前より気を配っていなかった	以前ほどは気を配っていなかったかもしれない	いつもと同じように気を配っていた
	0	1	2	3
１１．じっとしていられないほど落ち着かないことが	しょっちゅうあった	たびたびあった	少しだけあった	全くなかった
	0	1	2	3
１２．物事を楽しみにして待つことが	いつもと同じだけあった	以前ほどはなかった	以前よりも明らかに少なかった	めったになかった
	0	1	2	3
１３．突然、理由のない恐怖感（ﾊﾟﾆｯｸ）におそわれることが	しょっちゅうあった	たびたびあった	少しだけあった	全くなかった
	0	1	2	3
１４．面白い本や、ラジオまたはテレビ番組を楽しむことが	たびたびできた	ときどきできた	たまにできた	ほとんどめったにできなかった
	0	1	2	3

【医療との関わりについて】

1．あなたが医師を受診する際、次のことをどのくらい行いますか。（それぞれあてはまる番号ひとつに○）

	全くしない	たまにする	時々する	よくする	ほとんどいつもする	いつもする
1）医師に質問したいことのリストを用意する	0	1	2	3	4	5
2）治療について知りたいことや理解できていないことを質問する	0	1	2	3	4	5
3）病気にかかわるあなたの個人的な問題について話し合う	0	1	2	3	4	5

2．ここ**6ヶ月間**で、あなたは何回医師を受診しましたか。 □ 回

（入院中の医師の回診や、救急外来への受診は除く）

3．ここ**6ヶ月間**で、あなたは何回救急外来を利用しましたか。 □ 回

4．ここ**6ヶ月間**で、あなたは何回入院しましたか。 □ 回

ここ**6ヶ月間**で、あなたは何泊入院しましたか。 □ 泊

（病院で過ごした夜の数を記入してください）

４ページ　質問５で、
糖尿病、喘息、高血圧、高脂血症、
膠原病、関節リウマチ、アレルギー性鼻炎
アトピー性皮膚炎　があるとお答えの方に
うかがいます

【現在のあなたの体調について】

あなたの症状や、ここ３ヶ月内の検査結果について
差し支えない範囲で、次の質問にお答え下さい。

【糖尿病の方】

１．検査結果がある　　２．検査を受けていない／わからない

↓

ＨｂA1c　（　　％）　（検査日：　　月　　日）
空腹時血糖（　　mg／dl）（検査日：　　月　　日）

【喘息の方】

①　ピークフローを測定していますか？

１．定期的に測定している　　２．測定していない

↓

この１週間の平均値を平常値と比べると
１．よい　　２．ほぼ同じ　　３．悪い

②　この１週間で、発作止めが必要な程度の発作はありましたか？

１．はい（発作の回数：＿＿＿回／週）　　２．いいえ

③　この１週間で発作で眠れない日はありましたか？

１．はい（眠れなかった日数：　　日／週）　２．いいえ

【高血圧の方】

血圧（　　／　　mmHg）　（検査日：　　月　　日）

【高脂血症の方】

１．検査結果がある　　２．検査を受けていない／わからない

↓

総コレステロール（Ｔ-cho）（　　mg／dl）
HDL コレステロール　（　　mg／dl）
LDL コレステロール　（　　mg／dl）
中性脂肪（ＴＧ）　（　　mg／dl）（検査日：　　月　　日）

【膠原病の方】

1．検査結果がある　　2．検査を受けていない／わからない

血沈（ESR）（　　mm）　（検査日：　　月　　日）

【関節リウマチの方】

① ご自分で数えたとき、
全身で痛む関節はいくつありますか？　痛む関節の数 [　　] ヶ所

② 血沈・CRP の検査値について
1．検査結果がある　　2．検査を受けていない／わからない

血沈（ESR）（　　mm）
CRP （　　mg/dl）　（検査日：　　月　　日）

【アレルギー性鼻炎の方】

この 1 週間で、何回くらい症状（くしゃみ・鼻水・鼻づまり）が出ましたか？
1 日を午前・午後にわけ、1 週間分 14 回のうち、症状が出た回数をお答え下さい

この 1 週間で [　　] 回／14 回

【アトピー性皮膚炎の方】

① この 1 週間で、1 日あたり平均して何回くらいかゆみを感じましたか?

1 日平均 [　　] 回

② 症状は全身のどの部分ですか？　あてはまるものに○をつけてください。
頭・顔・首・胸・腹部・背中・おしり・手・腕・足・ひじ・ひざ

―すべての方にうかがいます―
最近のお体の調子や、健康のために
心がけていることについて
ご自由にお書きください

お手数ですが**糖尿病をお持ちの方**は以下の質問にお答えください

答え方：あなたの考えでは、以下に示すような糖尿病に関することがらが、
あなたにとってどのくらい問題になっていますか？
それぞれの質問項目について、最も当てはまる答の番号に○をつけてください。

例えば、
ある質問項目があなたにとって、心配でもなく、あてはまらず、問題になっていなければ、
“1”に○をつけて下さい。もしそのことでたいへん悩んでおられれば、“5”に○をして下さい。
それぞれの質問について、1 から 5 の 5 段階の中から番号で選んでください。

項目すべての度合いを表す数字に○をつけてありますか、もう一度ご確認下さい

1． 自分の糖尿病の治療法（食事療法、運動療法、飲み薬、インスリン注射、自己血糖測定など）について、はっきりとした具体的な目標がない	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	1・・・・・2・・・・・3・・・・・4・・・・・5	
2．自分の糖尿病の治療法がいやになる	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	1・・・・・2・・・・・3・・・・・4・・・・・5	
3．糖尿病を持ちながら生きていくことを考えるとこわくなる	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	1・・・・・2・・・・・3・・・・・4・・・・・5	
4．糖尿病の治療に関連して、周りの人たちから不愉快な思いをさせられる（例えば、他人があなたに何を食べるべきか指示するなど）	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	1・・・・・2・・・・・3・・・・・4・・・・・5	
5．食べ物や食事の楽しみを奪われたと感じる	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	1・・・・・2・・・・・3・・・・・4・・・・・5	
6．糖尿病を持ちながら生きていくことを考えるとゆううつになる	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	1・・・・・2・・・・・3・・・・・4・・・・・5	
7．自分の気分や感情が糖尿病と関係しているかどうかが分からない	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	1・・・・・2・・・・・3・・・・・4・・・・・5	
8．糖尿病に打ちのめされたように感じる	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	1・・・・・2・・・・・3・・・・・4・・・・・5	
9．低血糖が心配である	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	1・・・・・2・・・・・3・・・・・4・・・・・5	
10．糖尿病を持ちながら生きていくことを考えると腹が立つ	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	1・・・・・2・・・・・3・・・・・4・・・・・5	
１１．つねに食べ物や食事が気になる	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	1・・・・・2・・・・・3・・・・・4・・・・・5	

１２．将来のことや重い合併症になるかもしれないことが心配である	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
１３．糖尿病を管理していくことから脱線したとき、罪悪感や不安を感じる	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
１４．自分が糖尿病であることを受け入れていない	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
１５．糖尿病をみてもらっている医者に対して不満がある	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
１６．糖尿病のために、毎日多くの精神的エネルギーや肉体的エネルギーが奪われていると思う	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
１７．糖尿病のせいでひとりぼっちだと思う	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
１８．自分が糖尿病管理のために努力していることに対して、友人や家族は協力的でないと感じる	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
１９．自分が今持っている糖尿病の合併症に対処していくことが難しいと感じる	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
２０．糖尿病を管理するために努力しつづけて、疲れ燃え尽きてしまった	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	

２．あなたの身長・体重、糖尿病になって何年経つか、および糖尿病の型を教えてください

身長	体重	糖尿病になって	糖尿病の型（あてはまるものに○）	研究室使用欄
cm	kg	年　　ヶ月	１型 ２型 その他の型（具体的に　　）	

以上で質問は終わりです。ご協力ありがとうございました。
お手数ですが、もう一度お書き忘れがないかご確認をお願いします。
最後にこの研究調査に対するご意見・ご感想などありましたら、ご自由にお書きください。

Appendix 4. 3-month questionnaire

慢性疾患セルフマネジメントプログラム　に関する調査　③

ID

●●●お問合わせ先●●●

ご質問やご不明な点がございましたら、下記までお問合わせください。

東京大学大学院医学系研究科　健康社会学教室
セルフマネジメントプログラム評価研究チーム
担当：朴敏廷（パクミンジョン）・湯川慶子（ゆかわけいこ）
電話：０３－５８４１－３５１４
FAX：０３－５６８４－６０８３
Eメール：mjpark-tky@umin.ac.jp（受付時間：平日 10 時～17 時）

ワークショップの開始から３ヶ月が経過しました。このアンケートでは、あなたの現在の状態について教えてください。

調査日：　　年　　月　　日

【全般的な健康状態について】

あなたの健康状態は、全般的に見て、いかがですか。

（最もよくあてはまる番号 ひとつに○）

１．とてもよい
２．よい
３．普通
４．ややおもわしくない
５．おもわしくない

【症状について】

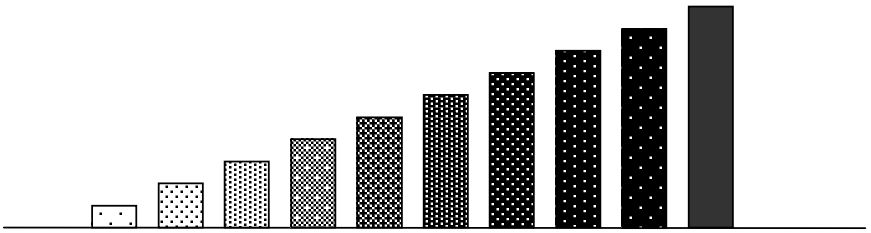
ここ１ヵ月間、次のように思ったり感じたりしたことがどのくらいの頻度でありましたか。

（それぞれ最もよくあてはまる番号 ひとつに○）

	全くなかった	たまにあった	時々あった	よくあった	ほとんどいつもあった	いつもあった
１）自分の健康上の問題で落ち込むことがあった	0	1	2	3	4	5
２）自分の将来の健康状態を考えると怖くなることがあった	0	1	2	3	4	5
３）自分の健康状態は人生における心配事のひとつだと思うことがあった	0	1	2	3	4	5
４）自分の健康状態は期待どおりにいっていないと感じることがあった	0	1	2	3	4	5

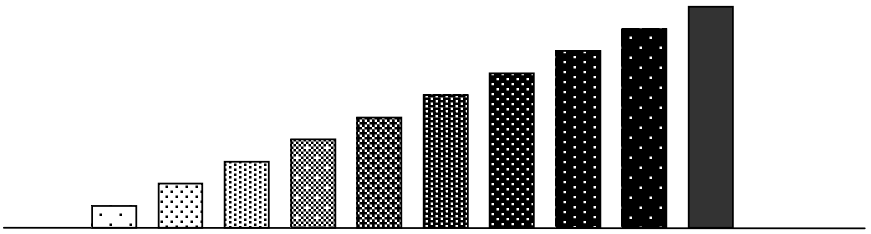
ここ２週間の状態についてうかがいます。

１）　ここ２週間のあなたの疲労の程度について、下の図であてはまる数字ひとつに○をつけてください。



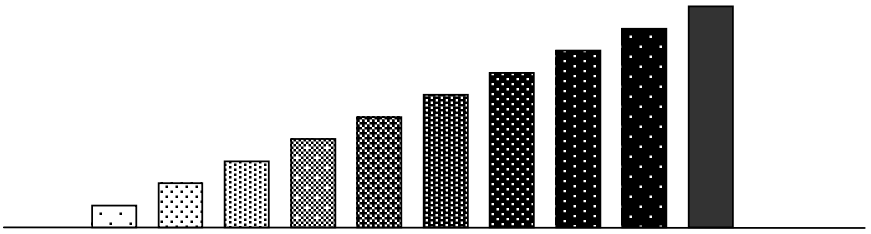
疲労なし　　きわめて耐え難い疲労

２）ここ２週間のあなたの息切れの程度について、下の図であてはまる数字ひとつに○をつけてください。



息切れなし　　きわめて耐え難い息切れ

３）ここ２週間のあなたの痛みの程度について、下の図であてはまる数字ひとつに○をつけてください。



痛みなし　　きわめて耐え難い痛み

【運動について】

あなたは次の運動を、ここ１週間で合計何時間くらい行いましたか。その１週間がいつも通りでなかったとしても、かまわずお答え下さい。

（それぞれ最もよくあてはまる番号 ひとつに○）

	なし	30 分未満/週	30～60 分/週	1～3 時間/週	3 時間以上/週
１）ストレッチまたは筋力トレーニング	0	1	2	3	4
２）ウォーキング（散歩など）・ジョギング	0	1	2	3	4
３）水泳または水中での運動	0	1	2	3	4
４）サイクリング（エアロバイクも含む）	0	1	2	3	4
５）マシンを使用した運動	0	1	2	3	4
６）その他の運動 （具体的に＿＿＿＿＿＿＿＿）	0	1	2	3	4

【症状への対処について】

気分が落ち込んだり、痛みや他の不快な症状があるとき、あなたはどのように対処していますか。

（それぞれ最もよくあてはまる番号 ひとつに○）

	全くしない	たまにする	時々する	よくする	ほとんどいつもする	いつもする
１）不快な症状から離れて、自分の体の一部ではないと感じるよう努める	0	1	2	3	4	5
２）不快な症状として考えずに、温かいとか無感覚だという何か他の感覚として考える	0	1	2	3	4	5
３）不快な症状から気をまぎらわすために、頭の中でゲームをしたり、歌を歌ったりする	0	1	2	3	4	5
４）部分ごとに体の筋肉をリラックスさせる	0	1	2	3	4	5
５）自分がどこか別のところにいるような想像したり、音声に導かれるイメージ法を行う	0	1	2	3	4	5
６）物事を前向きに考えるようにする	0	1	2	3	4	5

【日常の動作について】

現時点で、あなたは次のことをどのくらいできますか。　（それぞれあてはまる番号ひとつに○）

	何の困難もない	いくらか困難	かなり困難	できない
1）自分で身支度ができますか？（靴ひもを結ぶこと、ボタン掛けも含む）	0	1	2	3
2）就寝・起床の動作ができますか？	0	1	2	3
3）いっぱいに水が入っている茶碗やコップを口元まで運べますか？	0	1	2	3
4）屋外で平坦な道を歩くことができますか？	0	1	2	3
5）全身を洗い、タオルで拭くことができますか？	0	1	2	3
6）腰をまげ床にある衣類を拾うことができますか？	0	1	2	3
7）蛇口の開け閉めができますか？	0	1	2	3
8）車の乗り降りができますか？	0	1	2	3

【日常生活について】

ここ４週間、次の活動をする際、どのくらいの支障がありましたか。

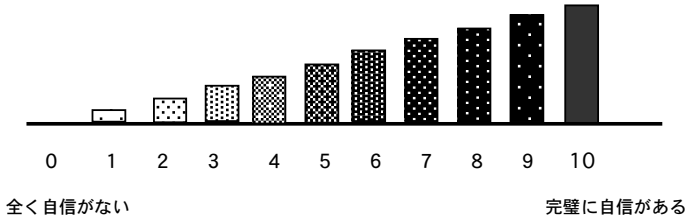
（それぞれあてはまる番号ひとつに○）

	全くなかった	少しあった	時々あった	よくあった	いつもあった
1）健康上の問題によって、家族・友人・近隣の人たちとのふだんの社会生活に支障がありましたか？	0	1	2	3	4
2）健康上の問題によって、趣味または娯楽に支障がありましたか？	0	1	2	3	4
3）健康上の問題によって、家事に支障がありましたか？	0	1	2	3	4
4）健康上の問題によって、用事や買物に出るのに支障がありましたか？	0	1	2	3	4

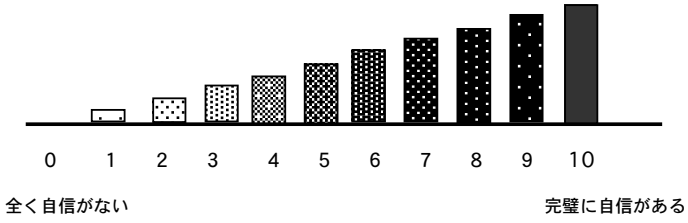
【実行できる自信について】

現時点で、あなたが次のことを実行できる自信はどのくらいありますか。
（それぞれ最もよくあてはまる数字ひとつに○）

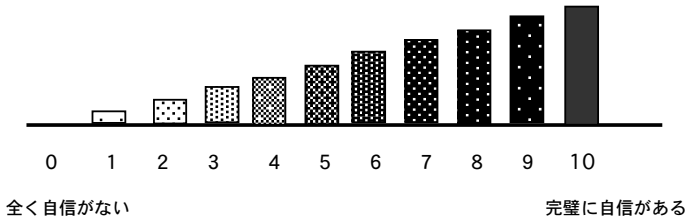
**1）病気による疲労があっても
やりたいことを実行できる自信は
どのくらいありますか？**



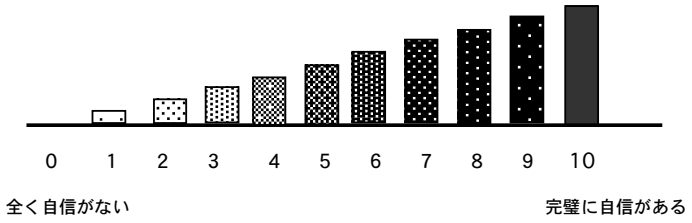
**2）病気による体の不快さや痛みがあっても
やりたいことを実行できる自信は
どのくらいありますか？**



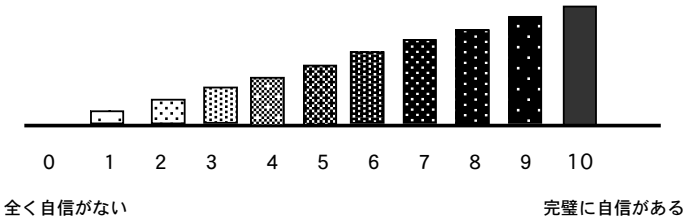
**3）病気による精神的苦痛があっても
やりたいことを実行できる自信は
どのくらいありますか？**



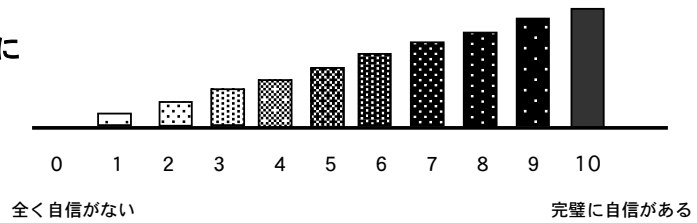
**4）その他の症状や健康問題があっても
やりたいことを実行できる自信は
どのくらいありますか？**



**5）医師にかかる回数が減るように
あなた自身の健康管理に必要な
さまざまなことを実行できる自信は
どのくらいありますか？**



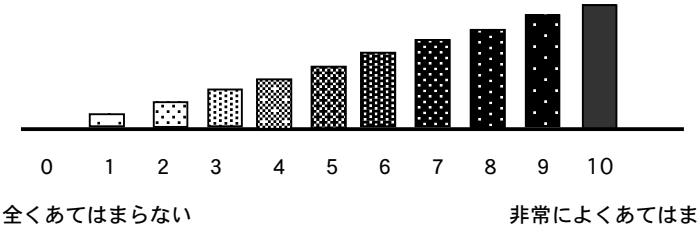
**6）病気による日常生活への影響が減るように
服薬以外のことも実行できる自信は
どのくらいありますか？**



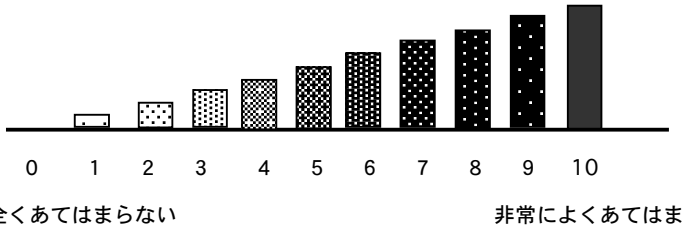
【生活の感じ方について】

あなたの人生に対する感じ方についてうかがいます。それぞれ 0 から 10 までのうち、
あなたの感じ方を最もよく表している数字ひとつに○をつけてください。

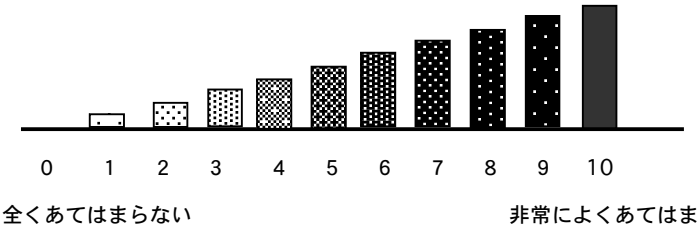
1）私は、日常生活で直面する困難や問題の
解決方法を見つけることができる



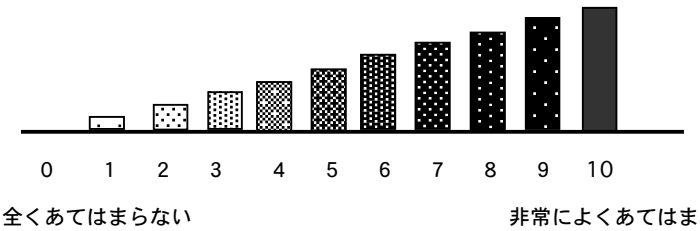
2）日常生活で直面する困難や問題の
いくつかは向き合い取り組むに値する、と
私は思えない



3）私は、日常生活で生じる困難や問題を
理解したり予測したりできる



4）私の日常生活は、
喜びと満足を与えてくれる



【心の状態について】

ここ１週間のあなたのご様子についてうかがいます。あなたはここ１週間どのように感じていますか。　（それぞれ最もよくあてはまる番号 ひとつに○）

１）緊張したり、気持ちが張りつめたりすることが	しょっちゅうあった	たびたびあった	ときどきあった	全くなかった
	0	1	2	3
２）むかし楽しんだことを今でも楽しいと思うことが	全く同じだけあった	かなりあった	少しだけあった	めったになかった
	0	1	2	3
３）なにか恐ろしいことが起ころうとしているという恐怖感を持つことが	しょっちゅうあって非常に気になった	たびたびあるがあまり気にならなかった	少しあるが気にならなかった	全くなかった
	0	1	2	3
４）物事の面白い面を笑ったり、理解したりすることが	いつもと同じだけできた	かなりできた	少しだけできた	全くできなかった
	0	1	2	3
５）心配事が心に浮かぶことが	しょっちゅうあった	たびたびあった	それほど多くはないが、ときどきあった	ごくたまにあった
	0	1	2	3
６）きげんの良いことが	全くなかった	たまにあった	ときどきあった	しょっちゅうあった
	0	1	2	3
７）楽に座って、くつろぐことが	かならずできた	たいていできた	たまにできた	全くできなかった
	0	1	2	3
８）仕事を怠けているように感じることが	ほとんどいつもあった	たびたびあった	ときどきあった	全くなかった
	0	1	2	3
９）不安で落ち着かないような恐怖感を持つことが	全くなかった	ときどきあった	たびたびあった	しょっちゅうあった
	0	1	2	3
１０）自分の顔、髪型、服装に関して	関心がなくなった	以前より気を配っていなかった	以前ほどは気を配っていなかったかもしれない	いつもと同じように気を配っていた
	0	1	2	3
１１）じっとしていられないほど落ち着かないことが	しょっちゅうあった	たびたびあった	少しだけあった	全くなかった
	0	1	2	3
１２）物事を楽しみにして待つことが	いつもと同じだけあった	以前ほどはなかった	以前よりも明らかに少なかった	めったになかった
	0	1	2	3
１３）突然、理由のない恐怖感（ﾊﾟﾆｯｸ）におそわれることが	しょっちゅうあった	たびたびあった	少しだけあった	全くなかった
	0	1	2	3
１４）面白い本や、ラジオまたはテレビ番組を楽しむことが	たびたびできた	ときどきできた	たまにできた	ほとんどめったにできなかった
	0	1	2	3

【医療との関わりについて】

1．あなたが医師を受診する際、次のことをどのくらい行いますか。

（それぞれ最もよくあてはまる番号ひとつに○）

	全くしない	たまにする	時々する	よくする	ほとんどいつもする	いつもする
1）医師に質問したいことのリストを用意する	0	1	2	3	4	5
2）治療について知りたいことや理解できていないことを質問する	0	1	2	3	4	5
3）病気にかかわるあなたの個人的な問題について話し合う	0	1	2	3	4	5

2．ここ６ヶ月間で、あなたは何回医師を受診しましたか。
（入院中の医師の回診や、救急外来への受診は除く）　□ 回

3．ここ６ヶ月間で、あなたは何回救急外来を利用しましたか。　□ 回

4．ここ６ヶ月間で、あなたは何回入院しましたか。　□ 回

ここ６ヶ月間で、あなたは合計、何泊入院しましたか。
（病院で過ごした夜の数を記入してください）　□ 泊

【ワークショップを受講して】

A. ワークショップが異なった疾患をもつ方々との集まりだったことに関して、お聞きします。

1）異なった疾患をもつ方々との集まりについてよかった点はありましたか？（番号ひとつに○）

なかった　　少しあった　　おおいにあった

1・・・・・・2・・・・・・3

＜2・3とお答えの方へ＞　具体的にはどんな点でしたか？

2）異なった疾患をもつ方々との集まりについて不満だった点はありましたか？（番号ひとつに○）

なかった　　少しあった　　おおいにあった

1・・・・・・2・・・・・・3

＜2・3とお答えの方へ＞　具体的にはどんな点でしたか？

B. ワークショップに参加することを通じて、あなたには次の点でどのような変化がありましたか？　（あてはまる番号ひとつに○）

1）気持ちが楽になったという感覚は	全く得られなかった	どちらかといえば得られなかった	どちらともいえない	どちらかといえば得られた	おおいに得られた
	0・・・・・・1	・・・・・・2	・・・・・・3	・・・・・・4	
2）少しずつでよい、無理しなくて良いという感覚は	全く得られなかった	どちらかといえば得られなかった	どちらともいえない	どちらかといえば得られた	おおいに得られた
	0・・・・・・1	・・・・・・2	・・・・・・3	・・・・・・4	
3）他人の助けになっているという感覚は	全く得られなかった	どちらかといえば得られなかった	どちらともいえない	どちらかといえば得られた	おおいに得られた
	0・・・・・・1	・・・・・・2	・・・・・・3	・・・・・・4	
4）物事をある程度冷静に受け止められるという感覚は	減った	どちらかといえば減った	どちらともいえない	どちらかといえば増えた	増えた
	0・・・・・・1	・・・・・・2	・・・・・・3	・・・・・・4	
5）できないことよりできることに目が向くように	全くならなkった	どちらかといえばならなかった	どちらともいえない	どちらかといえばなった	なった
	0・・・・・・1	・・・・・・2	・・・・・・3	・・・・・・4	
6）仲間と出会ったことによる心強さは	全く得られなかった	どちらかといえば得られなかった	どちらともいえない	どちらかといえば得られた	得られた
	0・・・・・・1	・・・・・・2	・・・・・・3	・・・・・・4	
7）何事にたいしても	悪い方向に考えるようになった	どちらかといえば悪い方向に考えるようになった	どちらともいえない	どちらかといえば良い方向に考えるようになった	良い方向に考えるようになった
	0・・・・・・1	・・・・・・2	・・・・・・3	・・・・・・4	

8）そのほかに、ワークショップに参加して、
肯定的に評価できる変化や得たものがありましたら、ぜひ教えて下さい。

４ページ　質問５で、
糖尿病、喘息、高血圧、高脂血症、
膠原病、関節リウマチ、アレルギー性鼻炎
アトピー性皮膚炎　があるとお答えの方に
うかがいます

【現在のあなたの体調について】

あなたの症状や、ここ３ヶ月内の検査結果について
差し支えない範囲で、次の質問にお答え下さい。

【糖尿病の方】

１．検査結果がある　　２．検査を受けていない／わからない

↓

ＨｂA1c　（　　％）　（検査日：　　月　　日）
空腹時血糖（　　mg／dl）（検査日：　　月　　日）

【喘息の方】

①　ピークフローを測定していますか？

１．定期的に測定している　　２．測定していない

↓

この１週間の平均値を平常値と比べると
１．よい　　２．ほぼ同じ　　３．悪い

②　この１週間で、発作止めが必要な程度の発作はありましたか？

１．はい（発作の回数：＿＿＿回／週）　　２．いいえ

③　この１週間で発作で眠れない日はありましたか？

１．はい（眠れなかった日数：　　日／週）　２．いいえ

【高血圧の方】

血圧（　　／　　mmHg）　（検査日：　　月　　日）

【高脂血症の方】

１．検査結果がある　　２．検査を受けていない／わからない

↓

総コレステロール（Ｔ-cho）（　　mg／dl）
HDL コレステロール　（　　mg／dl）
LDL コレステロール　（　　mg／dl）
中性脂肪（ＴＧ）　（　　mg／dl）（検査日：　　月　　日）

【膠原病の方】

1．検査結果がある　　2．検査を受けていない／わからない

血沈（ESR）（　　mm）　（検査日：　　月　　日）

【関節リウマチの方】

① ご自分で数えたとき、
全身で痛む関節はいくつありますか？　痛む関節の数 [　　] ヶ所

② 血沈・CRP の検査値について
1．検査結果がある　　2．検査を受けていない／わからない

血沈（ESR）（　　mm）
CRP （　　mg/dl）　（検査日：　　月　　日）

【アレルギー性鼻炎の方】

この 1 週間で、何回くらい症状（くしゃみ・鼻水・鼻づまり）が出ましたか？
1 日を午前・午後にわけ、1 週間分 14 回のうち、症状が出た回数をお答え下さい

この 1 週間で [　　] 回／14 回

【アトピー性皮膚炎の方】

① この 1 週間で、1 日あたり平均して何回くらいかゆみを感じましたか?

1 日平均 [　　] 回

② 症状は全身のどの部分ですか？　あてはまるものに○をつけてください。
頭・顔・首・胸・腹部・背中・おしり・手・腕・足・ひじ・ひざ

―すべての方にうかがいます―
最近のお体の調子や、健康のために
心がけていることについて
ご自由にお書きください

お手数ですが**糖尿病をお持ちの方**は以下の質問にお答えください

答え方：あなたの考えでは、以下に示すような糖尿病に関することがらが、
あなたにとってどのくらい問題になっていますか？
それぞれの質問項目について、最も当てはまる答の番号に○をつけてください。

例えば、
ある質問項目があなたにとって、心配でもなく、あてはまらず、問題になっていなければ、
“１”に○をつけて下さい。もしそのことでたいへん悩んでおられれば、“５”に○をして下さい。
それぞれの質問について、１から５の５段階の中から番号で選んでください。

項目すべての度合いを表す数字に○をつけてありますか、もう一度ご確認下さい

１．　自分の糖尿病の治療法（食事療法、運動療法、飲み薬、インスリン注射、自己血糖測定など）について、はっきりとした具体的な目標がない	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
２．自分の糖尿病の治療法がいやになる	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
３．糖尿病を持ちながら生きていくことを考えるとこわくなる	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
４．糖尿病の治療に関連して、周りの人たちから不愉快な思いをさせられる（例えば、他人があなたに何を食べるべきか指示するなど）	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
５．食べ物や食事の楽しみを奪われたと感じる	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
６．糖尿病を持ちながら生きていくことを考えるとゆううつになる	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
７．自分の気分や感情が糖尿病と関係しているかどうかが分からない	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
８．糖尿病に打ちのめされたように感じる	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
９．低血糖が心配である	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
１０．糖尿病を持ちながら生きていくことを考えると腹が立つ	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
１１．つねに食べ物や食事が気になる	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	

１２．将来のことや重い合併症になるかもしれないことが心配である	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
１３．糖尿病を管理していくことから脱線したとき、罪悪感や不安を感じる	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
１４．自分が糖尿病であることを受け入れていない	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
１５．糖尿病をみてもらっている医者に対して不満がある	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
１６．糖尿病のために、毎日多くの精神的エネルギーや肉体的エネルギーが奪われていると思う	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
１７．糖尿病のせいでひとりぼっちだと思う	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
１８．自分が糖尿病管理のために努力していることに対して、友人や家族は協力的でないと感じる	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
１９．自分が今持っている糖尿病の合併症に対処していくことが難しいと感じる	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	
２０．糖尿病を管理するために努力しつづけて、疲れ燃え尽きてしまった	私にとってそれは全く問題ではない	私はそのことで大変悩んでいる
	１・・・・・２・・・・・３・・・・・４・・・・・５	

２．あなたの身長・体重、糖尿病になって何年経つか、および糖尿病の型を教えてください

身長	体重	糖尿病になって	糖尿病の型（あてはまるものに○）	研究室使用欄
cm	kg	年　　ヶ月	１型 ２型 その他の型（具体的に　　）	

以上で質問は終わりです。ご協力ありがとうございました。
お手数ですが、もう一度お書き忘れがないかご確認をお願いします。
最後にこの研究調査に対するご意見・ご感想などありましたら、ご自由にお書きください。

Appendix 5: Caution regarding interpretation of improvement and deterioration

Differences between the deterioration subgroup and the improvement subgroup should be interpreted only with extreme caution, because of the possibility of regression to the mean. This can be seen in the findings regarding anxiety. As shown in Figure 4g, the median anxiety scores in the improvement subgroup at baseline were clearly higher than those in the deterioration subgroup at the same time. That is, many participants with high anxiety scores at baseline had lower scores at subsequent times, and many participants with low anxiety scores at baseline had higher scores at subsequent times, which indicates some regression to the mean.

Appendix 6. Diagnoses designated as "other" in Table 3.

Diagnosis	Number of patients
adrenal insufficiency	1
amblyopia	1
amyotrophic lateral sclerosis	1
anemia	2
ankylosing spondylitis	1
antiphospholipid antibody syndrome	1
aphasia	1
aplastic anemia	2
balance disorder	1
benign paroxysmal positional vertigo	1
biliary sludge	1
bipolar disorder type II	1
central retinal vein occlustion	1
cerebellar infarction	1
cervical disc hernia	2
cervico-omo-brachial syndrome	1
chromosomal abnormality	1
chronic glomerulonephritis	1
chronic hepatitis	1
chronic nephritis	2
chronic pancreatitis	1
chronic pharyngitis	1
chronic thyroiditis	1
complex regional pain syndrome	1
congenital male infertility	1
conjunctivitis	1
dry eyes	1
dysuria	1
endometriosis	1
esophageal achalasia	1
glaucoma	2
Graves' disease	4
growth hormone deficiency dwarfism	1
Hashimoto's disease	4
hemophilia A	1
hepatitis	2
hepatitis C	3
hyperthyroidism	2
hypopituitarism.	1
hypothyroidism	4
idiopathic avascular necrosis of the femoral hea	1
idiopathic small bowel dysfunction	1
idiopathic thrombocytopenic purpura	1
IgA nephropathy	1

Appendix 6. (continued) Diagnoses designated as "other" in Table 3.

Diagnosis	Number of patients
Klinefelter's syndrome	1
left ear hearing loss	1
low-back pain	4
lower limb disuse syndrome	1
lumbar vertebrae herniated disk	2
macroamylasemia	1
migraine	2
Minamata disease	1
multiple sclerosis	3
myasthenia gravis	1
neurofibromatosis	1
neurogenic bladder	1
obesity	2
ossification of posterior longitudinal ligament	4
ossification of the yellow ligament	1
osteoarthritis	2
osteoporosis	1
panic disorder	1
pemphigus	1
perceptive deafness	1
periodontal disease	1
polychondritis	1
post-nephrectomy	1
post-traumatic stress disorder	2
primary biliary cirrhosis	1
prostatic hypertrophy	1
pyelonephritis	1
rectal dysfunction	1
renal failure	3
retinal occult macular dystrophy	1
retinitis pigmentosa	5
retinopathy	1
Russell-Silver syndrome	1
sarcoidosis	1
schizophrenia	1
sciatica	1
sensitivity to cold	2
sinusitis	1
social anxiety disorder	1
spinal canal stenosis	2
spinocerebellar degeneration	3
subacute myelo-optico-neuropathy	1
sudden deafness	1
temporomandibular disorder	1
terminal ileitis	1
thrombocytosis	1
tinnitus	1
uterine fibroid	2
visual field disturbance	2
visual impairment	2

Appendix 7. There are various methods for growing classification trees. Table 7 and Figure 8 show the results for the method called CRT. The "unbiased" method called QUEST was also used. For 4 of the 8 outcomes QUEST gave worse predictions than CRT, and for the other 4 outcomes QUEST gave no predictions at all, as shown in the Table below.

Self-rated health

Growing method	CRT	QUEST (p = 0.32)
Number of nodes	> 1	> 1
Misclassification risk	0.252 ± 0.03	0.374 ± 0.03
% correctly predicted to have decay	71.0%	45.2%
Area under the ROC curve	0.789	0.645

Pain

Growing method	CRT	QUEST (p = 0.99)
Number of nodes	> 1	1 (no tree)
Misclassification risk	0.250 ± 0.05	0.368 ± 0.06
% correctly predicted to have decay	40.0%	–
Area under the ROC curve	0.680	–

Communication

Growing method	CRT	QUEST (p = 0.15)
Number of nodes	> 1	> 1
Misclassification risk	0.239 ± 0.04	0.333 ± 0.04
% correctly predicted to have decay	46.3%	24.1%
Area under the ROC curve	0.775	0.593

Coping

Growing method	CRT	QUEST (p = 0.16)
Number of nodes	> 1	> 1
Misclassification risk	0.186 ± 0.03	0.404 ± 0.04
% correctly predicted to have decay	81.7%	32.4%
Area under the ROC curve	0.863	0.574

Self-efficacy

Growing method	CRT	QUEST (p = 0.35)
Number of nodes	> 1	> 1
Misclassification risk	0.231 ± 0.03	0.244 ± 0.03
% correctly predicted to have decay	30.8%	26.9%
Area under the ROC curve	0.732	0.731

Health distress

Growing method	CRT	QUEST (p = 0.99)
Number of nodes (1 or > 1)	> 1	1 (no tree)
Misclassification risk	0.237 ± 0.03	0.32 ± 0.04
% correctly predicted to have decay	35.2%	–
Area under the ROC curve	0.696	–

Anxiety

Growing method	CRT	QUEST (p = 0.99)
Number of nodes	> 1	1 (no tree)
Misclassification risk	0.198 ± 0.04	0.396 ± 0.05
% correctly predicted to have decay	77.3%	–
Area under the ROC curve	0.832	–

Depression

Growing method	CRT	QUEST (p = 0.99)
Number of nodes	> 1	1 (no tree)
Misclassification risk	0.281 ± 0.04	0.36 ± 0.04
% correctly predicted to have decay	24.4%	–
Area under the ROC curve	0.683	–