# 論文の内容の要旨

論文課題　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）

氏名　朴敏廷 (Park Min Jeong)

**Background**

The burden of chronic illnesses is increasing worldwide. People with chronic diseases of course require medical treatment, but they also need a "patient-professional partnership, involving collaborative care and self- management education" [5]. They require education to promote skills and strategies for handling the daily problems caused by chronic illnesses [9].

One example of self-management education is Stanford University's generic Chronic Disease Self-Management Program (CDSMP), which is based on self-efficacy theory. The CDSMP is successful. It can decrease fatigue, disability, health-related distress, limitations on social activities, the number of hospitalizations, and the number of inpatient days. Also, the CDSMP increased self-reported health, communication with physicians, cognitive symptom management, and the frequency of physical exercise. One conclusion was: "it is possible to educate patients with different chronic diseases successfully in the same intervention at the same time" [49].

Long- term studies are rare, but researchers generally believe that "Short-term effects are rarely maintained over long intervals" [68] and that "effects tend not to be maintained." [69] This is called decay of impact [55] (Figure 1). To prevent or reverse that decay, reinforcement programs have been recommended [55, 77], but those programs are not effective [29, 30, 68, 78, 79, 80]. The reason they are not effective is unknown. However, if a subgroup has decay of impact, then reinforcements will appear to be ineffective because only some of the people who receive the reinforcements in fact need them. That is, the benefits of reinforcements may be diluted by whole-group analyses that include their lack of effect among people who do not need them. Reinforcements can be optimized on the basis of information about the decay of impact, but previous studies give no such information. The decay itself has never been studied, and there is no method for predicting which participants will have decay. Thus, this thesis addresses two main questions: 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?

**Methods Participants, program, and data collection**

The participants were adults with chronic diseases who voluntarily participated in the CDSMP. 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. Each discussion group had two lay leaders, who facilitated and managed the discussion.

All data were collected via self-administered questionnaires [82]. There were 8 outcome measures: self-rated overall health status, pain during the previous 2 weeks, the use of cognitive techniques to cope with symptoms, the use of proactive methods for improving communication with medical doctors, anxiety in the past week, depression in the past week, health-related distress in the past month, and self-efficacy to manage chronic conditions. The baseline questionnaire also asked about age, schooling, civil status, diagnoses, etc. Baseline data were collected before the first group-discussion session. Follow-up data were collected 3, 6, and 12

months later (Figure 2).

**Operational definition of decay of impact**

The operational definition of decay of impact had two criteria: (#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. Data that met those two criteria were categorized as decay of impact. To apply that definition, first, an estimated true score ($t'$) was computed [86] for each observed baseline score, best score, and last score, using test-retest reliability [82]:

$t'$ = (reliability x (observed score – mean score)) + mean score. The next step was to construct ranges of true scores ($t$) that were most consistent with observed scores ($x$), that is, $\sigma t.x$. [86]. Then, confidence intervals (CI) were made for each $t'$: CI = $t' \pm \sigma t.x$ .

Scores were considered to be different only if their CIs did not overlap. Non-overlap of CIs was used to judge whether the two criteria mentioned above were met. If there was no decay of impact, then the pattern was categorized as improvement, deterioration, or no change (Figures 3).!

**Timing, prevalence, and magnitude of decay of impact**

Timing: 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. Prevalence: For each outcome measure, the percentage of participants who had decay of impact was computed. Magnitude: For each instance of decay of impact, the magnitude of the decay was defined as the difference between the best value and the last value. To allow comparisons between outcomes measured on different scales, the magnitude of decay as a percent of the maximum possible decay (i.e., % of the full-scale value) was computed.

**Predictors of having decay of impact**

To identify predictors of having decay of impact, classification trees [87-90] were used, with 28 independent variables: 4 socio-demographic variables, 11 clinical variables, 11 baseline values, and 2 values measured at the 3-month follow-up (self-efficacy at 3 months and perceived positive change). The two classifications used were having decay of impact and having improvement only. To avoid overfitting, the trees were "pruned" [92] according to Brieman's 1-SE rule [87]. The classification trees were evaluated according to 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 curve [90, 92, 93].

**Results Participants and data collected**

Usable data were obtained from 364 participants. Many of them were middle-aged. Almost 80% were women. The length of time since the diagnosis of their chronic disease varied widely, from less than 1 year to more than 60 years. More than 40% of them had more than one diagnosis, and more than 15% of them had more than two diagnoses. The most common diagnoses were allergic disease, cardiovascular disease, connective tissue disease, diabetes, and rheumatic disease (Table 3 and Appendix 6). ⌴

**Patterns of change over time**

In contrast to the small changes over time at the whole-group level, the pattern-defined subgroups with improvement and with decay of impact had large changes and were easy to identify (for example, Figure 4b).

**Timing, prevalence, and magnitude of decay of impact**

Timing: The percentage of participants in whom the decay began at 3 months ranged from 26.1% to 61.4% (Table 5). Prevalence: Depending on the outcome, the prevalence of decay of impact ranged from 7% (pain) to 26% (self-rated health) (Figure 5). Magnitude: 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. The frequency distributions of magnitude of decay were right-skewed: some people had more than 50% decay,

and some had more than 60% decay, on some measures (Figure 6, Table 4).

**Predictors of having decay of impact (Table 7, Figure 8)**

In the classification trees the risks of misclassification were all less than 0.3. The best trees were those for predicting decay on coping, on anxiety, and on self-rated health. For those 3, the percentages of participants who were correctly classified with decay were greater than 70%, and the areas under the ROC curves were greater than 0.78.

In general, diagnoses were not associated with having decay of impact. There were only 2 exceptions. People with fibromyalgia syndrome were more likely to have decay on self-rated health, and people with Parkinson's disease were more likely to have decay on pain.

The most consistent predictor was the number of years since diagnosis, which was included in 6 of the 8 trees. In 5 of those 6, participants with longer disease histories were predicted to have decay of impact. However, in communication with medical doctors, participants with longer disease histories were predicted have improvement rather than decay.

## Discussion

### Summary of the main findings

The results of this study provide information about the prevalence, magnitude, and timing of the decay of impact. They also show 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 reinforcement is needed, in about 30% to 60% of the participants the decay began 3 months after the program started, so 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. Except for fibromyalgia syndrome and Parkinson's disease, diagnoses were not good predictors of having decay of impact. For self-rated health, coping, and anxiety, the classification trees gave good predictions of who would have decay of impact.

### Patterns of change, in the context of previous work

This is the first study to propose a single explanation of two important findings: that the effects of these programs are relatively small [22], and that reinforcement programs have failed [29, 30, 68, 79, 80]. The proposed explanation is that using whole-group summary statistics only, i.e. not analyzing pattern-defined subgroups, previous results reflect *dilution*. First, the present study's focus on pattern-defined subgroups rather than on whole- group analyses makes it clear that in previous studies the benefits to some participants were diluted by the lack of benefits to others. Second, 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. The solution, therefore, is to study pattern-defined subgroups, particularly decay of impact.

### 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, the timing of decay found here cannot be directly compared with previous findings. In related areas similar decay was found less than 6 months after the end of the intervention [74, 75], which is generally consistent with the present results.

### Predictors of decay

With no theory of the decay of impact in the context of this program, the search for predictors was exploratory, observational, and empirical. Particularly for self-rated health, coping, and anxiety, this exploratory analysis of many possible predictors resulted in good classification trees, even after the trees were pruned to avoid overfitting.

One noteworthy finding is that on 5 of the 8 outcomes participants with longer diseases histories were predicted 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 disease histories were predicted *not* 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 [97]. It is possible that people with more experience as patients might be more motivated to maintain their new self-management skills.

People with fibromyalgia syndrome and people with Parkinson's disease were predicted to have decay of impact on self-rated health and on pain, respectively, but no other diagnoses were included in the trees after pruning. 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].

**Conclusions and recommendations**

Decay of impact can occur on many different outcomes, and it can occur in up to one quarter of this program's participants. Among those who have decay, it can start as early as 3 months after the baseline measurement, i.e., 6 weeks after the end of the program, and its' magnitude can be as high as about 40% of the full-scale value.

Classification trees can give good predictions of who will have decay, and therefore who will need reinforcement, particularly for coping, anxiety, and self-rated health. Most diagnoses are not good predictors of 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.

These results can explain why previous studies found only small effects, and why reinforcements appeared to fail. Both of those findings could be caused by dilution; by mixture of data from people who have decay with data from people who do not.

Replication studies are needed, to determine the generalizability of these results. 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.

A conceptual model that includes self-efficacy, such as the theory of relapse prevention after treatment of addiction [77, 91] or the model proposed by LW Green and MW Kreuter [95, pages 160-161], might be adapted to fit the decay of impact after these educational programs. Rothman's theory of behavioral maintenance [96] might also be useful: "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." These theories should be operationalized and tested in the context of changes after self-management education for people with chronic diseases.

In relapse-prevention theory [91] social support is important, so it should be measured in future studies. In Rothman's theory of behavioral maintenance "perceived satisfaction with received outcomes" [94] is important, so that type of satisfaction should also be measured.

Patterns of change can be used to evaluate programs. Successful programs would be those after which very few people have decay of impact, or as those after which that decay is small.

Many people who have decay need reinforcement as early as 3 months after the start of the main program. If the resources needed for reinforcement are limited, then classification trees (at least for self-rated health, coping, and anxiety) can be used to predict who will have decay. Program administrators can then focus their limited resources on those participants.