Trajectories of change after a health-education program in Japan: decay of impact in anxiety, depression, and patient-physician communication

Min Jeong Park [1, 2], **Joseph Green** [Corresp., 3], **Hun Sik Jung** [4], **Yoon Soo Park** [5]

1 Department of Nursing, College of Nursing, Konyang University, Daejeon, South Korea

2 Department of Health Communication, University of Tokyo, Tokyo, Japan

3 Graduate School of Medicine, University of Tokyo, Tokyo, Japan

4 College of Global Business, Konyang University, Nonsan, South Korea

5 College of Medicine, University of Illinois at Chicago, Chicago, United States

Corresponding Author: Joseph Green
Email address: jgreen@gol.com

Background. Health education can benefit people with chronic diseases. However, in previous research those benefits were small, and reinforcement to maintain them was not effective. A possible explanation is that the benefits *appeared* to be small and reinforcement *appeared* to be ineffective because those analyses mixed data from two latent groups: one group of people who needed reinforcement and one group of people who did not. The hypothesis is that mixing the data from those two different groups caused the true effects to be "diluted."

Methods. To test that hypothesis we used data from the Chronic Disease Self-Management Program in Japan, focusing on anxiety, depression, and patient-physician communication. To identify latent trajectories of change after the program, we used growth-mixture modeling. Then, to find out which baseline factors were associated with trajectory-group membership, we used logistic regression.

Results. Growth-mixture modeling revealed two trajectories – two groups that were defined by distinct patterns of change after the program. One of those patterns was improvement followed by backsliding: decay of impact. On anxiety and depression the decay of impact was large enough to be clinically important, and its prevalence was as high as 50%. Next, logistic regression analysis revealed that being in the decay-of-impact group could be predicted from multimorbidity, low self-efficacy, and high scores on anxiety or depression at baseline. In addition, one unexpected finding was an association between multimorbidity and *better* patient-physician communication.

Conclusions. These results support the hypothesis that previous findings (i.e. *apparently* small effect sizes and *apparently* ineffective reinforcement) actually reflect "dilution" of large effects, which was caused by mixing of data from distinct groups. Specifically, there was one group with decay of impact and one without. Thus, evaluations of health education should include analyses of trajectory-defined groups. These results show how the group of people who are most likely to need reinforcement can be identified even before the educational program begins. Extra attention and reinforcement can then be tailored. They can be focused specifically to benefit the people with the greatest need.

Trajectories of change after a health-education program in Japan: decay of impact in

anxiety, depression, and patient-physician communication

Min Jeong Park[1,2], Joseph Green[3], Hun Sik Jung[4], Yoon Soo Park[5]

[1]Department of Nursing, College of Nursing, Konyang University, Daejeon, Republic of Korea

[2]Department of Health Communication, The University of Tokyo, Tokyo, Japan

[3]Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

[4]College of Global Business, Konyang University, Nonsan, Republic of Korea

[5]College of Medicine, University of Illinois at Chicago, Chicago, IL, USA

Corresponding author:

Joseph Green[3], <jgreen@m.u-tokyo.ac.jp>

ABSTRACT

21   **Background**. Health education can benefit people with chronic diseases. However, in previous

22   research those benefits were small, and reinforcement to maintain them was not effective. A

23   possible explanation is that the benefits *appeared* to be small and reinforcement *appeared* to be

24   ineffective because those analyses mixed data from two latent groups: one group of people who

25   needed reinforcement and one group of people who did not. The hypothesis is that mixing the

26   data from those two different groups caused the true effects to be "diluted."

27   **Methods**. To test that hypothesis we used data from the Chronic Disease Self-Management

28   Program in Japan, focusing on anxiety, depression, and patient-physician communication. To

29   identify latent trajectories of change after the program, we used growth-mixture modeling. Then,

30   to find out which baseline factors were associated with trajectory-group membership, we used

31   logistic regression.

32   **Results**. Growth-mixture modeling revealed two trajectories – two groups that were defined by

33   distinct patterns of change after the program. One of those patterns was improvement followed

34   by backsliding: decay of impact. On anxiety and depression the decay of impact was large

35   enough to be clinically important, and its prevalence was as high as 50%. Next, logistic

36   regression analysis revealed that being in the decay-of-impact group could be predicted from

37   multimorbidity, low self-efficacy, and high scores on anxiety or depression at baseline. In

38   addition, one unexpected finding was an association between multimorbidity and *better* patient-

39   physician communication.

40   **Conclusions**. These results support the hypothesis that previous findings (i.e. *apparently* small

41   effect sizes and *apparently* ineffective reinforcement) actually reflect "dilution" of large effects,

42   which was caused by mixing of data from distinct groups. Specifically, there was one group with

43 decay of impact and one without. Thus, evaluations of health education should include analyses

44 of trajectory-defined groups. These results show how the group of people who are most likely to

45 need reinforcement can be identified even before the educational program begins. Extra attention

46 and reinforcement can then be tailored. They can be focused specifically to benefit the people

47 with the greatest need.

48

49   **INTRODUCTION**

50   Nothing lasts forever. Yet a worthy goal of health education is for its benefits to be

51 sustained. In some studies, benefits of health education have been found to last longer than 6

52 months (Barlow, Wright, Turner, & Bancroft, 2005; Barlow et al., 2008; Brady et al., 2013),

53 while in others the findings are more nuanced – some improvements do not endure (Caplin &

54 Creer, 2001; Clark, 2003; Franks, Chapman, Duberstein, & Jerant, 2009; Hennessy et al., 1999;

55 Krebs, Prochaska, & Rossi, 2010; Lorig, Ritter, Laurent, & Fries, 2004; Norris, Lau, Smith,

56 Schmid, & Engelgau, 2002). Particularly with regard to the self-management educational

57 interventions that are used to help people reduce the impact of chronic diseases, it was recently

58 noted that those interventions might need improvements to ensure that they have longer-lasting

59 impact (Miller, Lasiter, Ellis, & Buelow, 2015), which at least implicitly acknowledges that their

60 immediate benefits can decay after their (short-lived) demonstrated period of effectiveness. To

61 sustain these programs' benefits, reinforcement has been recommended (Clark, 2003; Green,

62 1977; Newman, Steed, & Mulligan, 2009), but reinforcement has generally not been found to be

63 useful (Glasgow, Toobert, Hampson, & Strycker, 2002; Lorig & Holman, 1989; Lorig, Ritter,

64 Laurent, & Plant, 2006; Lorig, Ritter, Villa, & Piette, 2008; Nguyen, Carrieri-Kohlman, Rankin,

65 Slaughter, & Stulbarg, 2005; Riemsma, Taal, & Rasker, 2003).

66   In this context, we consider the Chronic Disease Self-Management Program (CDSMP),

67 which can improve health status and can increase the frequency of desirable health-related

68 behaviors (Franek, 2013; Whitelaw, Lorig, Smith, & Ory, 2013). While the benefits of the

69 CDSMP are "statistically significant," some of them have also been described as minimal

70 (Franek, 2013) or moderate (Brady et al., 2013). The program's developers attributed these

71 "modest" effect sizes to heterogeneity in the clinical and demographic characteristics of the

72   participants (Lorig, Ritter, Laurent, & Plant, 2006), which points to a need to understand

73   differences among participants and factors that might contribute to larger and more-sustained

74   benefits.

75   In general, treating participants as homogeneous conceals true heterogeneity. For example,

76   an intervention may be very useful in some participants, but that fact will not be recognized if

77   one examines only the average for the group as a whole (Moynihan, Henry, & Moons, 2014). In

78   addition, important heterogeneity in treatment effects can occur not only across groups but also

79   over time. Sustained improvement in some participants could obscure relapse in others.

80   Reinforcement given to all can appear to be ineffective, even if it is quite useful to some.

81   In previous research, subsets of participants were defined by socio-demographic

82   characteristics, personality factors, etc., and not by patterns of change after the intervention

83   (Franks et al, 2009; Harrison et al., 2012; Jerant, Chapman, Duberstein, & Franks, 2010; Reeves

84   et al., 2008; Swerissen et al., 2006; Smeulders et al., 2010). In contrast, LW Green (1977)

85   discussed five distinct patterns (i.e. trajectories) of change after health education. He referred to

86   the pattern in which good outcomes do not endure as *decay of impact* (also sometimes called

87   relapse or backsliding). Its essence is simple: deterioration after improvement. Outside of health

88   education, a similar concept is used commonly in research on the effectiveness of treatments for

89   addiction (Hendershot et al., 2011; Menon & Kandasamy, 2018). Within the field of health

90   education, Green (1977) may have been the first to point out that decay of impact can cause the

91   benefits of an intervention to *appear* to be larger or smaller than they really are, depending on

92   when the outcome is measured. That difference between the apparent change and the real change

93   can be minimized only by measuring the outcome repeatedly over a relatively long period of

94   time. Hennessy et al. (1999) measured self-efficacy for condom use 5 times during 1 year after

95   an HIV-prevention educational program, and they found clear evidence of decay of impact 3

96   months after enrollment in the program. Caplin and Creer (2001) identified various

97   characteristics of people with decay of impact 7 years after an asthma self-management program:

98   For example, the people with decay of impact had initially been less *self*-motivated to participate

99   in the program. Nonetheless, with only a few exceptions, over more than 40 years since Green

100   (1977) described decay of impact in health education, that concept seems to have received little

101   attention from researchers. Stated simply, decay of impact as a person's pattern of change after

102   chronic-disease self-management education has been studied only very rarely. But knowing the

103   magnitude of decay, and knowing when and in whom it occurs would be very useful in

104   evaluating a health-education program's effectiveness, and also in targeting interventions.

105   Information about decay of impact can give planners an objective basis for deciding whether

106   reinforcement is needed, when it is needed (Hennessy et al., 1999), and who is most likely to

107   need it (Park, Green, Ishikawa, & Kiuchi, 2012).

108   Taking seriously the possibility of decay of impact entails defining groups by their

109   patterns of change, i.e. the trajectories of their outcomes during follow-up. Using data collected

110   in two waves, Nolte, Elsworth, Sinclair, & Osborne (2007) may have been the first to study

111   groups defined by their change after health education. An extension from patterns defined using

112   two waves of data to those defined using four was published a few years later (Park et al., 2013).

113   However, without the use of a well-established method for analyzing longitudinal data, doubts

114   remain regarding different trajectories in outcomes. In this current study, using data collected

115   before and after the CDSMP we applied Growth-Mixture Modeling (GMM) to empirically

116   identify latent groups that were characterized by their patterns of longitudinal change, and we

117   subsequently used logistic regression to identify baseline factors contributing to group

118   membership (Cook, Karriker-Jaffe, Bond, & Lui, 2015; Hibbard, Mahoney, Stock, & Tusler,

119   2007; Muthén, 2004; Rabe-Hesketh, Skrondal, & Pickles, 2004; Ram & Grimm, 2009). This

120   approach allows more granularity in evaluating the program's effects and more accuracy in

121   assessing needs for reinforcement.

122   Many outcomes of the CDSMP have been studied (Lorig et al., 1996), but here we

123   focused only on symptoms of anxiety, symptoms of depression, and the use of proactive

124   techniques to improve patient-physician communication. Thus, the focus is not on the

125   longitudinal change in participants' chronic illnesses themselves (with the possible exception of

126   the small number of participants who had clinical depression). Trajectories of change in anxiety

127   and depression could reflect changes in the application of coping skills learned in the CDSMP,

128   while patient-physician communication is of course a health-related behavior. We chose those

129   three outcomes because all of them are relevant in many different chronic medical conditions,

130   and thus the results will be important to a large proportion of all CDSMP participants. Also, one

131   area on which the CDSMP focuses specifically is learning and practicing better patient-physician

132   communication.

133   **METHODS**

134   **Participants**

135   In this study, we focused on a particular population: adults in Japan who have at least one

136   chronic medical condition and who participated in the CDSMP. In this implementation of the

137   CDSMP in Japan, the program accommodated people who were community-dwelling, were able

138   to attend weekly group meetings, and were able to communicate easily in Japanese. As in all

139   implementations of the CDSMP, the participants had various different chronic medical

140   conditions. The population of CDSMP participants in Japan consists only of adults (defined as

141 people who were at least 18 years old), but it is otherwise diverse with respect to age, duration of

142 chronic illness, and other socio-demographic and clinical characteristics (Appendix 1).

143   CDSMP participants were recruited through public service centers, outpatient clinics of

144 hospitals, and the Internet homepage of the Japan Chronic Disease Self-Management Association

145 (Japan Chronic Disease Self-Management Association, 2018). Some participants also found out

146 about the program through friends or acquaintances. Participation was voluntary. All people who

147 met the conditions described above comprised the population of interest. The study itself

148 involved a census (not a sample) of that population.

149 **The program**

150   Based on self-efficacy theory (Bandura, 2019), the program aims to build the

151 participants' skills in six areas: 1) handling pain, fatigue, frustration, and isolation, 2) exercising

152 to maintain and increase strength, endurance, and flexibility, 3) using medications appropriately,

153 4) improving communication with friends, family, and healthcare professionals, 5) achieving and

154 maintaining proper nutrition, and, 6) evaluating new therapies (Chronic Disease Self-

155 Management Program, 2018). Skills in those areas were taught and practiced during group-

156 discussion sessions that were held once each week for six weeks. Each group had two lay

157 facilitators who had undergone approximately 35 hours of training. A textbook was used as the

158 reference for the program's content (Lorig, Holman, Sobel, Laurent, & González, 2001).

159 **Measures**

160   Demographic and clinical information were collected using self-administered

161 questionnaires. Also included in the questionnaires were scales to measure health status, health-

162 related behaviors, psychological variables, etc. Those included self-efficacy to manage chronic

163 health conditions (on a 0-to-60 scale, coefficient alpha = 0.92) and the Hospital Anxiety and

164   Depression Scale (HADS, Matsudaira et al., 2009). The HADS asks about symptoms of anxiety

165   and of depression in the past week. Possible total scores on both the anxiety scale and the

166   depression scale range from 0 to 21, with higher scores reflecting more symptoms and more

167   frequent symptoms. Coefficient alpha for the depression scale of the HADS was 0.73 and for the

168   anxiety scale it was 0.84. Also included was a 3-item scale to measure communication with

169   physicians, with possible total scores ranging from 0 to 15 (coefficient alpha = 0.78). Higher

170   scores reflected more frequent use of proactive methods for good patient-physician

171   communication (Lorig et al., 1996).

172   **Study design and timing of measurements**

173   Data were collected four times over one year. Baseline data were collected before the

174   first group-discussion session. Follow-up questionnaires were sent by postal mail 3, 6, and 12

175   months later. A post-paid envelope addressed to the research office was included for returning

176   the questionnaire. To ensure that data collection was as complete as possible, one of the

177   researchers (MJP) was available to speak directly with participants and answer their questions

178   about the study. In addition, a reminder postcard was sent whenever a follow-up questionnaire

179   was not received by return mail within two weeks.

180   **Analyses**

181   To allow detection of decay of impact, the analyses were done using data from

182   participants who provided at least three waves of data (456/643; 71%). Unconditional quadratic

183   growth curves (Rabe-Hesketh, Skrondal, & Pickles, 2004) and conditional growth curves were

184   fit using the four-wave data with all participants. Time-variant (self-efficacy) and time-invariant

185   covariates (gender, educational status, partnered status, number of diagnoses, and history of

186   illness) were included in the growth-curve analysis, with anxiety, depression, and

187   communication as outcomes, to examine baseline differences and interactions over time.

188   Quadratic terms were specified on "time" by including a squared time (i.e. time x time) term.

189   Interaction terms between time-invariant covariates and time were included to examine factors

190   contributing to longitudinal changes.

191   GMM with quadratic growth curves were fit for 2, 3, and 4 latent groups. Posterior mode

192   estimation, a partly Bayesian approach, was used to obtain consistent parameter estimates and to

193   avoid boundary estimation issues (Park, Xing, &, Lee, 2018; Huang & Bandeen-Roche, 2004).

194   Moreover, to obtain optimal estimation, we began with 100 sets of starting values and used a

195   combination of expectation-maximization algorithm followed by Newton-Raphson to obtain

196   parameter estimates and model fit indices. Finally, we examined the Jacobian matrix to be of full

197   rank to ensure local identification of results. This process was repeated for each growth-mixture

198   model allowing for the different latent groups. We decided on the final number of groups after

199   examining the relative fit (Bayesian Information Criterion [BIC]) and absolute fit (Proportion

200   Correctly Classified [P_c] based on posterior probability). Multiple logistic regression was used to

201   identify factors contributing to trajectory-group membership. Of the 456 participants whose data

202   were analyzed by GMM here, 369 were included in a previous non-GMM analysis (Park et al.,

203   2013). Data were analyzed with Latent Gold 5.1 (Belmont, MA), Stata 14 (College Station, TX),

204   and JASP (https://jasp-stats.org/).

205   **Ethics**

206   This study was approved by the University of Tokyo (number 1472–(2), Research Ethics

207   Committee, Graduate School of Medicine). Participation in the CDSMP and in this research

208   were voluntary. Informed consent was obtained in writing from all participants before the study

209   began.

210   **RESULTS**

211   **1. The participants**

212   Data from 456 participants were analyzed. Among them, 79% were women, 48% were

213   college educated, 52% were partnered (married or living with someone), and 47% had more than

214   one chronic condition. Details of multimorbidity are in Appendix 1.

215   **2. All participants considered together**

216   In the analysis with all participants considered together (Table 1), for the first six months

217   communication with physicians increased, while both anxiety and depression decreased.

218   However, by the end of the follow-up year all three outcomes had begun changing back toward

219   their baseline values. That is, the initial improvements appeared to be followed by at least some

220   backsliding.

221   Similarly, the growth-curve analysis also indicated that change over time was curvilinear

222   for all three outcomes: p values for the quadratic terms for anxiety, depression, and

223   communication were 0.001, 0.038, and 0.002, respectively (Table 2). In addition, higher self-

224   efficacy at baseline was associated with less anxiety, less depression, and better communication

225   with physicians.

226   Also associated with both anxiety and communication was the number of diagnoses, and

227   the regression coefficients for both were positive (Table 2). That is, the participants with more

228   comorbid conditions had greater anxiety at baseline. It is noteworthy that the participants with

229   more comorbid conditions also had *better* baseline scores on the scale measuring communication

230   with physicians. These associations did not change over time, as evidenced by the small

231   coefficients for the terms representing interaction with time.

3. Groups defined by their trajectories

232

233   For all three outcomes, the GMM results were similar: The BIC and the P_c both led to the

234   conclusion that the best-fitting models were those with two groups (Appendix 2).

235   For each outcome, those two groups began from substantially different baseline scores

236   (Figure 1). Also for each outcome, one group changed very little throughout the follow-up period

237   while the other changed noticeably within the first six months of follow-up, and then it reversed

238   course back toward the baseline value (Figure 1). That is, on each outcome some participants

239   were in a decay-of-impact group and the others were not. About half of the participants were in

240   the decay-of-impact group: anxiety 45.6%, depression 50.7%, and communication with

241   physicians 46.3%.

242   Participants who had decay of impact on one of the two mental-health outcomes (anxiety

243   or depression) were also likely to be classified as having decay of impact on the other one (Phi =

244   0.508, Appendix 3). However, participants who had decay of impact on one of the two mental-

245   health outcomes were no more or less likely to have decay of impact on communication (Phi =

246   0.095 and 0.043, Appendix 3).

247   On the two mental-health outcomes, the decay-of-impact group was the group with worse

248   baseline status: more symptoms, and more-frequent symptoms, of anxiety and depression. In

249   contrast, on communication with physicians the decay-of-impact group was the group with better

250   baseline status: more frequent use of the three specified methods for good patient-physician

251   communication. Also, by the end of the follow-up year the anxiety and depression scores had

252   decayed back to their respective baseline levels, whereas on communication the decay trajectory

253   was clear but the scores did not return to the baseline level, in other words the decay itself was

254   smaller on communication with physicians than on the mental-health outcomes (Figure 1).

255 **4. Contributors to group membership (Table 3)**

256   For all three outcomes, self-efficacy at baseline was associated with group membership.

257 Participants with higher self-efficacy were more likely to be in the group with lower anxiety at

258 baseline, in the group with lower depression at baseline, and in the group with better

259 communication at baseline.

260   There were no noteworthy associations between group membership and gender, civil

261 status (i.e. partnered or not), or the number of years of chronic-disease history. College education

262 was associated with group membership on only one of the three outcomes (anxiety).

263   For anxiety and for communication with physicians, the number of diagnoses was also

264 associated with group membership. Regarding anxiety, participants with more diagnoses were

265 more likely to be in the group with higher (i.e. worse) scores at baseline and subsequent decay of

266 impact. Regarding communication, participants with more diagnoses were more likely to be in

267 the group with higher (i.e. *better*) scores at baseline and subsequent decay of impact.

268   **DISCUSSION**

269 **All participants**

270   When all participants were considered together, all three outcomes improved over the

271 first six months. That improvement was followed by a small deterioration. Thus, even from the

272 least-detailed analyses, some decay of impact was evident (Table 1). That interpretation is

273 supported by the results of the growth-curve analyses: change over time was curvilinear

274 (quadratic) for all three outcomes (Table 2).

275   The growth-curve analyses also showed that higher self-efficacy at baseline was

276 associated with less anxiety, less depression, and better communication with physicians, which is

277 consistent with the theoretical basis of the CDSMP (Lorig & González, 1992).

278   Also evident at this level of analysis were associations with multimorbidity. Having more

279   diagnoses was associated with more anxiety, more depression, and *better* communication with

280   physicians. Of those three findings, the first two might be expected, but the third is particularly

281   interesting. It is also reflected in the analyses of membership in trajectory-defined groups, and so

282   we will discuss it below.

283   **Findings from GMM: groups defined by their trajectories**

284   For all three outcomes the results of GMM were consistent: Among all of the models

285   tested, the two-group models had the lowest BIC and the highest P_c. We are therefore confident

286   in saying that GMM revealed *two* latent groups among these participants in the CDSMP. In some

287   circumstances practical considerations could override the conclusions from those statistical

288   criteria, as described in Appendix 2.

289   As noted above in the Introduction, in some previous studies, subsets of CDSMP

290   participants were defined *a priori* and with reference to theory. In contrast, groups identified by

291   GMM are empirical, as is each participant's group membership. We note that the GMM

292   approach can lead to testable hypotheses (regarding multimorbidity, as described below), and it

293   can be used to answer important questions about whether similar phenomena also occur among

294   other groups and in different settings.

295   **Factors affecting self-management**

296   These findings are not inconsistent with the five categories of factors affecting self-

297   management that were identified by Schulman-Green et al. (2016). It is noteworthy that the

298   CDSMP does address factors in all five of Schulman-Green's categories: self-efficacy (which is

299   in the category of "personal/lifestyle characteristics"), problem-solving to cope with symptoms

300   (in the category of "health status"), social support via the CDSMP group sessions (in the

301 category of "resources"), action plans for ensuring proper nutrition (in the category of

302 "environmental characteristics"), and proactive patient-physician communication (in the

303 category of "health care system").

304   With regard to the decay of impact, we note that Table 3 shows associations of trajectory-

305 group membership with two factors that were identified by Schulman-Green, et al. (2016) as

306 influencers of self-management. Those two factors are self-efficacy (in their "personal/lifestyle

307 characteristics" category) and the number of diagnoses (i.e. comorbidity, in their "health status"

308 category). In addition to those two, the baseline level of patient-physician communication is in

309 Schulman-Green's category of "health care system." It is quite possible that factors in Schulman-

310 Green's two other categories ("resources" and "environmental characteristics") also affect the

311 decay of impact, and we believe this is an important topic for future research.

312 **Mental health, and reinforcement**

313   For both mental-health outcomes, the trajectory-defined groups differed in their baseline

314 status and in their pattern of change after the program. Regarding anxiety, approximately half of

315 the participants were in a group that began with relatively good scores, and they improved very

316 gradually over the following year. In contrast, the other half were in a group that began from a

317 high-anxiety baseline. That second group improved over the first 6 months, but by the time of the

318 12-month follow-up it had returned to its baseline level, and thus we refer to the latter group as

319 the decay-of-impact group. The same was true with regard to depression.

320   Dichotomization is undoubtedly dangerous (Harrell, 2019), and yet HADS scores are

321 used to separate people into categories of anxiety and depression severity. In Japan, the HADS

322 threshold score separating non-cases from possible and probable cases was 9 (Matsudaira et al.,

323 2009). The decay-of-impact trajectories on both anxiety and depression crossed that threshold

324   twice – first during the improvement occurring soon after the intervention ended, and then again

325   about 8 months later during the decay back toward the baseline value (Figures 1a and 1b).

326   Therefore, to the extent that the threshold of 9 is useful, both the improvement measured soon

327   after the CDSMP and also the deterioration measured near the end of the follow-up year were

328   clinically important.

329   "Average therapeutic trial results can mislead" (Moynihan et al., 2014), but GMM

330   provides more detail than average results. Here GMM showed that only some of the participants

331   had decay of impact. At least on anxiety and depression, both the existence of a decay-of-impact

332   group and the movement of that group between clinical categories support the idea that follow-

333   up interventions – reinforcement – should be offered to *some* of the participants. Had

334   reinforcement been given to all, it is unlikely that those in the group without decay of impact

335   would have benefitted from it, simply because they already had almost no psychological distress

336   – almost no room to improve. Rather than being expended on all of the participants, the

337   resources used to implement reinforcement should be saved for the people who need it, to help

338   them maintain their newly-improved status or perhaps improve further.

339   The present findings show how the small effect sizes and null results in published studies

340   of reinforcement could be underestimates. Specifically, when all participants are considered

341   together (Table 1), then any benefits to a group with decay-of-impact (Figure 1) will become at

342   least more difficult to detect, and possibly even completely effaced, as they are mixed with the

343   much smaller (or even zero) benefits to the non-decay group. This dilution is clearly shown in

344   the difference between Table 1 and Figure 1. Table 1 shows only a very small decay of impact,

345   because all participants were considered together, but Figure 1 shows substantial decay of impact

346   because GMM separated the participants into two trajectory-defined groups as it "unmixed" the

347   data. If one were to look at Table 1 alone, one could easily conclude that the benefits of the

348   program and also the decay of impact (i.e. the need for reinforcement) were both very small. But

349   that would be another example of how average results can mislead (Moynihan et al., 2014).

350   Figure 1 shows that almost half of the participants had both important benefits and important

351   decay of impact, and the difference between Table 1 (the diluted results) and Figure 1 (the post-

352   GMM results) shows how mixing the data from two different groups caused important effects to

353   be diluted.

354   Testing the effects of reinforcement is reasonable, but only in those whom reinforcement

355   could benefit: the decay-of-impact group. To identify that group at baseline, i.e., even before the

356   CDSMP begins, the present findings offer two potential criteria: a high baseline score (greater

357   distress) on the mental-health outcome of interest, and a low level of self-efficacy. With regard

358   to anxiety, a third criterion could be multimorbidity.

359   **Communication with physicians**

360   Similar to the results described above for mental health, regarding communication with

361   physicians GMM revealed two groups, each comprising about half of the participants, and those

362   two groups began from noticeably different baselines. One group (n = 245) started from a very

363   low baseline communication score (about 4 points) and it changed very little over the following

364   year (Figure 1). This could well indicate an unmet need. Specifically, by the standard implied in

365   the patient-physician communication scale, for those 245 participants substantial improvement

366   after the baseline measurement was possible, but it did not occur. This leads to at least three

367   research questions: (1) Are some participants in fact satisfied with a "low" level of

368   communication? (2) Was the program implemented as well as possible? and (3) Even if the

369 implementation was good, would those 245 participants have benefitted from a more-intensive

370 intervention with an even-greater emphasis on practicing communication skills?

371   Also noteworthy are the three criteria that might be used to pre-emptively identify the

372 participants who are most likely to need communication-skill practice: a low communication

373 score at baseline, low self-efficacy, and *uni*morbidity.

374 **Multimorbidity**

375   The participants with more diagnoses had better communication scores in the initial

376 growth-curve analysis (Table 2), and they were more likely to be in the trajectory-defined group

377 that had better communication scores throughout the year (Table 3). While the CDSMP has been

378 found to be particularly useful to people with multiple diagnoses (Harrison et al., 2012), here

379 multimorbidity was associated with a desirable health-related behavior even at baseline. To

380 address the apparent connection between having multiple diagnoses and communicating well

381 with physicians, we begin by noting that the communication scores reflect how often the

382 respondents do the following three activities: making a list of questions to ask one's physician

383 during clinic visits; asking one's physician about things that one wants to know or does not

384 understand regarding one's treatment; and discussing (with one's physician) personal problems

385 related to one's medical condition. The people with multiple diagnoses probably had more

386 experience being in health-related situations that were difficult to manage. To deal with those

387 difficulties, perhaps they began writing lists of questions, asking for clarification, and discussing

388 personal issues related to their diseases, simply because their health conditions were so complex.

389 We hypothesize that at least some of the people with multimorbidity had become accustomed to

390 doing those activities, and therefore in the domain of patient-physician communication they had

391 already become "expert patients" (Reeves et al., 2008) by the time the study began. To the extent

392   that better patient-physician communication results in better clinical care, this connection

393   between multimorbidity and good communication could account at least in part for the

394   documented association of multimorbidity with higher-quality care (Min et al., 2007).

395   Other factors could also be important. For example, self-selection might have played a

396   role. After all, participation in the program was voluntary (as it is worldwide). Personality can

397   moderate the effects of the CDSMP (Franks et al., 2009; Jerant et al., 2010), and perhaps it also

398   affects one's decision to participate. Among all of the eligible people with multimorbidity, those

399   who are less "conscientious" and less interested in self-managing their conditions would not

400   often make lists of questions, etc., and they might not have found the CDSMP to be attractive

401   and thus would not have participated. In contrast, the CDSMP might appeal to people like the

402   highly-proactive communicator with 8 chronic conditions who was described by Haslam (2015).

403   People with multiple diagnoses who take initiative in self-managing their condition(s) by writing

404   lists of questions, etc. could be over-represented among the program's participants.

405   As explanations of the association between multimorbidity and good patient-physician

406   communication, both the self-selection hypothesis and the already-an-expert-patient hypothesis

407   remain to be tested, and of course they are not mutually exclusive: both could be true.

408   **Generalizability**

409   Because this study involved not a sample but a census of the population of interest, the

410   common concerns about generalizability are not strictly applicable. Chronic-disease self-

411   management interventions have been implemented in many countries, but the country-specific

412   differences in those implementations are important (O'Connell, McCarthy, & Savage, 2018), and

413   we do not claim that the results shown in Figure 1, Table 3, etc. necessarily apply to CDSMP

414   participants outside Japan. Nonetheless, we do suggest that the approach and the methods

415   employed here could be quite useful in evaluations of the CDSMP and of other health-education

416   programs worldwide. As a minimum, evidence explaining why effect sizes appear to be small is

417   important wherever the CDSMP is used. In addition, a generalized application of GMM and

418   similar methods would ground the understanding of these programs' effects more firmly in the

419   empirical reality of trajectory-defined groups. Related analytic methods may also be useful,

420   including methods that allow analyses of individual participants' trajectories (Kozlowski, Pretz,

421   Dams-O'Connor, Kreider, & Whiteneck, 2013). A generalized application of the approach and

422   the methods used in the present study would also help to meet the current need for research on

423   long-term maintenance of self-management skills (Miller, Lasiter, Ellis, & Buelow, 2015).

424   **Cost-effectiveness**

425   Cost-effectiveness analyses of the CDSMP indicate that it results in net savings (Ahn et

426   al., 2013). Its cost-effectiveness has also been demonstrated through analysis of the incremental

427   cost-effectiveness ratio per quality-adjusted life year (Basu et al., 2015). Formal cost-

428   effectiveness analyses of the prediction of CDSMP trajectory-group membership are beyond the

429   scope of the present study. Nonetheless, it is reasonable to expect the costs to be low.

430   Specifically, for the three outcomes of interest here, the present results indicate that predicting

431   trajectory-group membership might require completion of only the 14-item HADS, the 6-item

432   self-efficacy scale, the 3-item patient-physician communication scale, and a question-item

433   regarding multimorbidity – all at baseline only. The participants already fill out registration

434   forms before the start of the first CDSMP group session. If the burden on the participants must

435   be strictly minimized, then the necessary information could be gathered only once (at baseline),

436   and it would take only a few extra minutes.

437   While the costs would be expected to be low, the effectiveness (benefits) could be high,

438   because there is potential for prevention or minimization of the decay of impact. Specifically, as

439   shown in Table 3 and in Figure 1, baseline information might be used to predict who among the

440   participants will have decay of impact, and thus who is most likely to benefit from extra attention

441   during the program and from reinforcement after the program. This could strengthen the

442   empirical foundation for decisions about cost-effective allocation of those educational resources.

443   **Limitations**

444   The four waves of data collection over one year were more than enough to allow

445   detection of decay of impact, but more frequent measurement and longer follow-up would of

446   course be useful.

447   The number of diagnoses was self-reported. While we would have preferred to use

448   medical records, for many chronic conditions self-reported diagnosis is accurate enough for

449   research (Karison et al., 1999; Wada et al., 2009). The diagnosis of clinical depression was found

450   by Sanches-Villegas et al. (2008) to be over-reported, with a true-positive percentage of 74.2%.

451   In that study, among 62 people who self-reported a diagnosis of depression, 46 were found to be

452   true positives according to the SCID-I (46/62 = 0.742). It is possible that in the present study a

453   similar percentage of those who self-reported that they had been given a diagnosis of depression

454   would in fact have met the SCID-I criteria for depression. In that case, the number of true

455   positives would be estimated to be approximately 19 rather than the 26 who did self-report that

456   they had been given a diagnosis of depression (Appendix Table A1C): 0.742 x 26 = 19.3. This is

457   unlikely to have had a large effect on the main results or conclusions. The logistic regression

458   analysis did identify the number of diagnoses as a predictor of trajectory-group membership

459   (Table 3), but any effect of over-reporting of depression on that result was probably very small.

460   Specifically, $26 - 19 = 7$, and $7/456 = 0.015$. Thus, self-reporting of a diagnosis of depression

461   might have caused the number of participants with depression to have been over-estimated by

462   approximately 7, which is 1.5% of the total. In addition, it is important to remember that the

463   number of diagnoses was not included in the GMM analysis, so any inaccuracy or imprecision in

464   that number did not affect the GMM results (Figure 1 and Appendix 2). Also, the presence or

465   absence of any self-report of any specific diagnosis (depression, etc.) was not included in the

466   GMM analysis, so any inaccuracy or imprecision in those data also did not affect the GMM

467   results.

468   **Previous knowledge gaps, new knowledge, and applications**

469   One approach to the presentation of research is to explicate knowledge gaps, how those

470   gaps have been "filled," and the practical consequences thereof. With that in mind, we give such

471   information below.

472   Knowledge gap 1: The benefits of the CDSMP appear to be small (and reinforcement

473   appears to be ineffective), but the reason is not well understood. Filling knowledge gap 1, we

474   found how mixing of data from two distinct trajectory-defined groups can make the overall

475   benefits of the CDSMP appear to be small even though they are relatively large for some

476   participants. Also, one of those two trajectory-defined groups was characterized by decay of

477   impact. Given the new information obtained in filling that first gap, another knowledge gap

478   became clear.

479   Knowledge gap 2: Criteria for identifying people who are likely to have decay of impact

480   are not known. Filling knowledge gap 2, we found that the people who are most likely to have

481   decay of impact can be identified using baseline self-efficacy and the number of diagnoses

482 (Table 3). Also, people who eventually had decay of impact were those with higher baseline

483 scores on the outcome of interest (Figure 1).

484 Applications of new knowledge: This new knowledge can be useful in at least two ways.

485 First, it can be used to tailor reinforcement to the participants who are most likely to need

486 reinforcement, which could increase cost-effectiveness. Second, it can be used to establish two

487 new goals for CDSMP implementation and two new criteria for CDSMP evaluation: a low

488 prevalence of decay of impact, and a small magnitude of decay of impact (Park et. al, 2012).

489 **Considerations regarding theory**

490 As noted above, the theoretical foundation of the CDSMP is centered on the concept of

491 self-efficacy (Bandura, 2019). In light of the present results, other theories and conceptual

492 frameworks may also be useful, especially those emphasizing long-term outcomes (Miller,

493 Lasiter, Ellis, & Buelow, 2015). For example, the field of relapse prevention after treatment for

494 substance abuse has well-studied theories and practices (Hendershot, Witkiewitz, George, &

495 Marlatt, 2011; Menon & Kandasamy, 2018; Witkiewitz & Marlatt, 2004), some of which might

496 be adapted to inform strategies for preventing decay of impact after chronic-disease self-

497 management education. In addition, consideration should be given to the difference between

498 *changing* health-related behavior and *maintaining a new* health-related behavior for long-term

499 benefits (Rothman, 2000; Sciamanna et al, 2011; Joseph et al., 2016; Giacobbi, 2016).

500 **Conclusions**

501 GMM exposed two trajectory-defined groups, and the CDSMP clearly benefitted one

502 group more than the other. However, the group that benefitted also had substantial decay of

503 impact, and thus needed reinforcement. The decay-of-impact group comprised almost half of the

504 participants. At baseline (i.e., before the program began), the participants most likely to need

505   reinforcement were those with multimorbidity, those with low self-efficacy, and those who were

506   clinically anxious or depressed.

507   Once the participants who are likely to have decay of impact are identified, extra

508   attention and reinforcement can then be tailored. They can be focused specifically to benefit the

509   people with the greatest need.

510

511

512   Acknowledgments: The authors would like to express their gratitude to the Japan Chronic

513   Disease Self-Management Association, as well as to all of the people who participated in the

514   study. MJ Park is grateful for advice received from Y Yamazaki, for help and collaboration

515   provided by the self-management research team at the University of Tokyo, for administrative

516   assistance and technical support received from N Okamoto, and for advice and academic support

517   received from T Kiuchi and H Ishikawa.

518

519 **REFERENCES**

520

521 Ahn, S., Basu, R., Smith, M. L., Jiang, L,, Lorig, K., Whitelaw, N., Ory, M. G. (2013) The

522   impact of chronic disease self-management programs: healthcare savings through a

523   community-based intervention. *BMC Public Health*. 13, 1141. doi: 10.1186/1471-2458-

524   13-1141.

525 Bandura, A. (2019) Applying Theory for Human Betterment. *Perspectives on Psychological*

526   *Science*. 14(1), 12-15. doi: 10.1177/1745691618815165.

527 Barlow, J., Wright, C., Turner, A., & Bancroft, G. (2005). A 12-month follow-up study of self-

528   management training for people with chronic disease: Are changes maintained over time?

529   *British Journal Of Health Psychology*, 10(4), 589-599. doi: 10.1348/135910705x26317

530 Barlow, J., Turner, A., Swaby, L., Gilchrist, M., Wright, C., & Doherty, M. (2008). An 8-yr

531   follow-up of Arthritis Self-Management Programme participants. *Rheumatology*, 48(2),

532   128-133. doi: 10.1093/rheumatology/ken429

533 Basu, R., Ory, M. G, Towne, S. D. Jr., Smith, M. L., Hochhalter, A. K., Ahn S. (2015) Cost-

534   effectiveness of the chronic disease self-management program: implications for

535   community-based organizations. *Frontiers in Public Health*. 2015 Apr 27;3:27. doi:

536   10.3389/fpubh.2015.00027

537 Brady, T., Murphy, L., O'Colmain, B., Beauchesne, D., Daniels, B., & Greenberg, M., House,

538   M., & Chervin, D. (2013). A Meta-Analysis of Health Status, Health Behaviors, and

539   Health Care Utilization Outcomes of the Chronic Disease Self-Management Program.

540   *Preventing Chronic Disease*, 10. doi: 10.5888/pcd10.120112

541   Caplin, D., & Creer, T. (2001). A Self-Management Program for Adult Asthma. III. Maintenance

542   and Relapse of Skills. *Journal Of Asthma*, 38(4), 343-356. doi: 10.1081/jas-100000263

543   Chronic Disease Self-Management Program (SMRC). (2018). Retrieved from

544   http://patienteducation.stanford.edu/programs/cdsmp.html

545   Clark, N. (2003). Management of Chronic Disease by Patients. *Annual Review Of Public Health*,

546   24(1), 289-313. doi: 10.1146/annurev.publhealth.24.100901.141021

547   Cook, W., Karriker-Jaffe, K., Bond, J., & Lui, C. (2015). Asian American Problem Drinking

548   Trajectories During the Transition to Adulthood: Ethnic Drinking Cultures and

549   Neighborhood Contexts. *American Journal Of Public Health*, 105(5), 1020-1027. doi:

550   10.2105/ajph.2014.302196

551   Foster, G., Taylor, S. J. C., Eldridge, S., Ramsay, J., Griffiths, C. J. (2007) Self-management

552   education programmes by lay leaders for people with chronic conditions. *Cochrane*

553   *Database of Systematic Reviews* 4: Art. No.: CD005108.

554   Franek, J. (2013). Self-management support interventions for persons with chronic disease: an

555   evidence-based analysis. *Ontario Health Technology Assessment Series*, 13 9, 1-60.

556   Franks, P., Chapman, B., Duberstein, P., & Jerant, A. (2009). Five factor model personality

557   factors moderated the effects of an intervention to enhance chronic disease management

558   self-efficacy. *British Journal Of Health Psychology*, 14(3), 473-487. doi:

559   10.1348/135910708x360700

560   Giacobbi, P., R. (2016) Theoretical, Critical, and Practical Reflections on the Long-Term

561   Maintenance of Health Behavior Change. *American Journal of Lifestyle Medicine*. 10(6),

562   377–380. doi: 10.1177/1559827616662435

563   Glasgow, R., Toobert, D., Hampson, S., & Strycker, L. (2002). Implementation, generalization

564   and long-term results of the "choosing well" diabetes self-management intervention.

565   *Patient Education And Counseling*, 48(2), 115-122. doi: 10.1016/s0738-3991(02)00025-3

566   Green, L. (1977). Evaluation and measurement: some dilemmas for health education. *American*

567   *Journal Of Public Health*, 67(2), 155-161. doi: 10.2105/ajph.67.2.155

568   Harrell, F. (2019). Dichotomania. http://www.fharrell.com/post/errmed/#catg Accessed on

569   February 25, 2019

570   Harrison, M., Reeves, D., Harkness, E., Valderas, J., Kennedy, A., Rogers, A., Hann, M., &

571   Bower, P. (2012). A secondary analysis of the moderating effects of depression and

572   multimorbidity on the effectiveness of a chronic disease self-management programme.

573   *Patient Education And Counseling*, 87(1), 67-73. doi: 10.1016/j.pec.2011.06.007

574   Haslam, D. (2015). "You're an Expert in Me": The Role of the Generalist Doctor in the

575   Management of Patients with Multimorbidity. *Journal Of Comorbidity*, 5(1), 132-134.

576   doi: 10.15256/joc.2015.5.65

577   Hendershot, C.S., Witkiewitz, K., George, W.H., & Marlatt, G.A. (2011) Relapse prevention for

578   addictive behaviors. *Substance Abuse Treatment, Prevention, and Policy*, 6, 17. doi:

579   10.1186/1747-597X-6-17

580   Hennessy, M., Bolan, G., Hoxworth, T., Iatesta, M., Rhodes, F., & Zenilman, J. (1999). Using

581   growth curves to determine the timing of booster sessions. *Structural Equation Modeling:*

582   *A Multidisciplinary Journal*, 6(4), 322-342. doi: 10.1080/10705519909540139

583   Hibbard, J., Mahoney, E., Stock, R., & Tusler, M. (2007). Do Increases in Patient Activation

584   Result in Improved Self-Management Behaviors? *Health Services Research*, 42(4), 1443-

585   1463. doi: 10.1111/j.1475-6773.2006.00669.x

586   Huang, G. H., & Bandeen-Roche, K. (2004). Building an identifiable latent class model with

587   covariate effects on underlying and measured variables. *Psychometrika*, 69, 5-32.

588   Japan Chronic Disease Self-Management Association. (2018). Retrieved from http://www.j-

589   cdsm.org/

590   Jerant, A., Chapman, B., Duberstein, P., & Franks, P. (2010). Effects of personality on self-rated

591   health in a 1-year randomized controlled trial of chronic illness self-management. *British*

592   *Journal Of Health Psychology*, 15(2), 321-335. doi: 10.1348/135910709x464353

593   Joseph, R. P., Daniel, C. L., Thind, H., Benitez, T. J., & Pekmezi, D. (2016) Applying

594   Psychological Theories to Promote Long-Term Maintenance of Health Behaviors.

595   *American Journal of Lifestyle Medicine*. 10(6): 356–368. doi:

596   10.1177/1559827614554594

597   Karison, E., Lee, I., Cook, N., Manson, J., Buring, J., & Hennekens, C. (1999). Comparison of

598   Self-Reported Diagnosis of Connective Tissue Disease with Medical Records in Female

599   Health Professionals: The Women's Health Cohort Study. *American Journal Of*

600   *Epidemiology*, 150(6), 652-660. doi: 10.1093/oxfordjournals.aje.a010064

601   Kozlowski, A.J., Pretz, C.R., Dams-O'Connor, K., Kreider, S., Whiteneck, G. (2013) An

602   introduction to applying individual growth curve models to evaluate change in

603   rehabilitation: a National Institute on Disability and Rehabilitation Research Traumatic

604   Brain Injury Model Systems report. *Archives of Physical Medicine and Rehabilitation*,

605   94(3):589-96. doi: 10.1016/j.apmr.2012.08.199.

606   Krebs, P., Prochaska, J., & Rossi, J. (2010). A meta-analysis of computer-tailored interventions

607   for health behavior change. *Preventive Medicine*, 51(3-4), 214-221. doi:

608   10.1016/j.ypmed.2010.06.004

609   Lorig, K., & González, V. (1992). The Integration of Theory with Practice: A 12-Year Case

610   Study. *Health Education Quarterly*, 19(3), 355-368. doi: 10.1177/109019819201900307

611   Lorig, K., & Holman, H. (1989). Long-term outcomes of an arthritis self-management study:

612   Effects of reinforcement efforts. *Social Science & Medicine*, 29(2), 221-224. doi:

613   10.1016/0277-9536(89)90170-6

614   Lorig, K., Holman, H., Sobel, D., Laurent, D., & González, V. (2001). *Living a Healthy Life*

615   *With Chronic Conditions: Self-management of Heart Disease, Arthritis, Diabetes,*

616   *Asthma, Bronchitis, Emphysema and Others* (Translated by Fusae, K.) (2nd ed.). Tokyo:

617   Nihon Kango Kyokai (in Japanese).

618   Lorig, K., Ritter, P., Laurent, D., & Fries, J. (2004). Long-Term Randomized Controlled Trials

619   of Tailored-Print and Small-Group Arthritis Self-Management Interventions. *Medical*

620   *Care*, 42(4), 346-354. doi: 10.1097/01.mlr.0000118709.74348.65

621   Lorig, K., Ritter, P., Laurent, D., & Plant, K. (2006). Internet-Based Chronic Disease Self-

622   Management. *Medical Care*, 44(11), 964-971. doi: 10.1097/01.mlr.0000233678.80203.c1

623   Lorig, K., Ritter, P., Villa, F., & Piette, J. (2007). Spanish Diabetes Self-Management With and

624   Without Automated Telephone Reinforcement: Two randomized trials. *Diabetes Care*,

625   31(3), 408-414. doi: 10.2337/dc07-1313

626   Lorig, K., Stewart, A., Ritter, P., González, V., Laurent, D., & Lynch, J. (1996). *Outcome*

627   *measures for health education and other health care interventions*. Thousand Oaks: Sage

628   Publications.

629   Matsudaira, T., Igarashi, H., Kikuchi, H., Kano, R., Mitoma, H., Ohuchi, K., & Kitamura, T.

630   (2009). Factor structure of the Hospital Anxiety and Depression Scale in Japanese

631   psychiatric outpatient and student populations. *Health And Quality Of Life Outcomes*,

632   7(1), 42. doi: 10.1186/1477-7525-7-42

633 Menon, J., & Kandasamy, A. (2018) Relapse prevention. *Indian Journal of Psychiatry*, 60(Suppl

634   4), S473–S478. doi: 10.4103/psychiatry.IndianJPsychiatry_36_18

635 Miller, W.R., Lasiter, S., Ellis, R., & Buelow, J.M. (2015) Chronic disease self-management: a

636   hybrid concept analysis. *Nursing Outlook*. 63(2), 154-161. doi:

637   10.1016/j.outlook.2014.07.005

638 Min, L., Wenger, N., Fung, C., Chang, J., Ganz, D., Higashi, T., Kamberg, C. J., Maclean, C. H.,

639   Roth, C. P., Solomon, D. H., Young, R. T., & Reuben, D. B. (2007). Multimorbidity is

640   Associated With Better Quality of Care Among Vulnerable Elders. *Medical Care*, 45(6),

641   480-488. doi: 10.1097/mlr.0b013e318030fff9

642 Moynihan, R., Henry, D., & Moons, K. (2014). Using Evidence to Combat Overdiagnosis and

643   Overtreatment: Evaluating Treatments, Tests, and Disease Definitions in the Time of Too

644   Much. *Plos Medicine*, 11(7), e1001655. doi: 10.1371/journal.pmed.1001655

645 Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques

646   for longitudinal data. In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology*

647   *for the social sciences* (pp. 345-368). London: SAGE.

648 Newman, S., Steed, L., & Mulligan, K. (2009). *Chronic physical illness* (p. 72). Maidenhead:

649   Open University Press.

650 Nguyen, H., Carrieri-Kohlman, V., Rankin, S., Slaughter, R., & Stulbarg, M. (2005). Is Internet-

651   based support for dyspnea self-management in patients with chronic obstructive

652   pulmonary disease possible? Results of a pilot study. *Heart & Lung: The Journal Of*

653   *Acute And Critical Care*, 34(1), 51-62. doi: 10.1016/j.hrtlng.2004.06.005

654   Nolte, S., Elsworth, G., Sinclair, A., & Osborne, R. (2007). The extent and breadth of benefits

655   from participating in chronic disease self-management courses: A national patient-

656   reported outcomes survey. *Patient Education And Counseling*, 65(3), 351-360. doi:

657   10.1016/j.pec.2006.08.016

658   Norris, S., Lau, J., Smith, S., Schmid, C., & Engelgau, M. (2002). Self-Management Education

659   for Adults With Type 2 Diabetes: A meta-analysis of the effect on glycemic control.

660   *Diabetes Care*, 25(7), 1159-1171. doi: 10.2337/diacare.25.7.1159

661   O'Connell, S., McCarthy, V. J. C., Savage, E. (2018) Frameworks for self-management support

662   for chronic disease: a cross-country comparative document analysis. *BMC Health*

663   *Services Research*, 25;18(1):583. doi: 10.1186/s12913-018-3387-0.

664   Park, M., Green, J., Ishikawa, H., Yamazaki, Y., Kitagawa, A., & Ono, M., Yasukata, F., &

665   Kiuchi, T. (2013). Decay of Impact after Self-Management Education for People with

666   Chronic Illnesses: Changes in Anxiety and Depression over One Year. *Plos ONE*, 8(6),

667   e65316. doi: 10.1371/journal.pone.0065316

668   Park, M., Green, J., Ishikawa, H., & Kiuchi, T. (2012). Hidden decay of impact after education

669   for self-management of chronic illnesses: hypotheses. *Chronic Illness*, 9(1), 73-80. doi:

670   10.1177/1742395312453351

671   Park, Y. S., Xing, K., & Lee, Y. S. (2018). Explanatory cognitive diagnostic models:

672   Incorporating latent and observed predictors. *Applied Psychological Measurement*, 42(5),

673   376-392.

674   Rabe-Hesketh, S., Skrondal, A., & Pickles, A. (2004). Generalized multilevel structural equation

675   modeling. *Psychometrika*, 69(2), 167-190. doi: 10.1007/bf02295939

676   Ram, N., & Grimm, K. (2009). Methods and Measures: Growth mixture modeling: A method for

677   identifying differences in longitudinal change among unobserved groups. *International*

678   *Journal Of Behavioral Development*, 33(6), 565-576. doi: 10.1177/0165025409343765

679   Reeves, D., Kennedy, A., Fullwood, C., Bower, P., Gardner, C., & Gately, C., Lee, V.,

680   Richardson, G., & Rogers, A. (2008). Predicting who will benefit from an Expert Patients

681   Programme self-management course. *British Journal Of General Practice*, 58(548), 198-

682   203. doi: 10.3399/bjgp08x277320

683   Riemsma, R., Taal, E., & Rasker, J. (2003). Group education for patients with rheumatoid

684   arthritis and their partners. *Arthritis & Rheumatism*, 49(4), 556-566. doi:

685   10.1002/art.11207

686   Rothman AJ. (2000) Toward a theory-based analysis of behavioral maintenance. *Health*

687   *Psychology*. 19(1 Suppl): 64–69.

688   Sanchez-Villegas, A., Schlatter, J., Ortuno, F., Lahortiga, F., Pla, J., Benito, S., & Martinez-

689   Gonzalez, M. A. (2008). Validity of a self-reported diagnosis of depression among

690   participants in a cohort study using the Structured Clinical Interview for DSM-IV (SCID-

691   I). *BMC Psychiatry*, 8(1), 43.

692   Schulman-Green, D., Jaser, S. S., Park, C., & Whittemore, R. (2016). A metasynthesis of factors

693   affecting self-management of chronic illness. *Journal of Advanced Nursing*, 72(7), 1469-

694   1489.

695   Sciamanna, C. N., Kiernan, M., Rolls, B. J., Boan, J., Stuckey, H., Kephart, D., Miller, C. K.,

696   Jensen, G., Hartmann, T. J., Loken, E., Hwang, K. O., Williams, R. J., Clark, M. A.,

697   Schubart, J. R., Nezu, A. M., Lehman, E., & Dellasega, C. (2011) Practices associated

698   with weight loss versus weight-loss maintenance. *American Journal of Preventive*

699   *Medicine*. 41(2), 159-66. doi: 10.1016/j.amepre.2011.04.009.

700   Smeulders, E., van Haastregt, J., Ambergen, T., Stoffers, H., Janssen-Boyne, J., Uszko-Lencer,

701   N., Gorgels, A., Lodewijks-van der Bolt, C., van Eijk, J., & Kempen, G. (2010). Heart

702   failure patients with a lower educational level and better cognitive status benefit most

703   from a self-management group programme. *Patient Education And Counseling*, 81(2),

704   214-221. doi: 10.1016/j.pec.2010.01.003

705   Swerissen, H., Belfrage, J., Weeks, A., Jordan, L., Walker, C., Furler, J., McAvoy, B., Carter,

706   M., & Peterson, C. (2006). A randomised control trial of a self-management program for

707   people with a chronic illness from Vietnamese, Chinese, Italian and Greek backgrounds.

708   *Patient Education And Counseling*, 64(1-3), 360-368. doi: 10.1016/j.pec.2006.04.003

709   Wada, K., Yatsuya, H., Ouyang, P., Otsuka, R., Mitsuhashi, H., Takefuji, S., Matsushita, K.,

710   Sugiura, K., Hotta, Y., Toyoshima, H., & Tamakoshi, K. (2009). Self-reported medical

711   history was generally accurate among Japanese workplace population. *Journal Of*

712   *Clinical Epidemiology*, 62(3), 306-313. doi: 10.1016/j.jclinepi.2008.04.006

713   Whitelaw, N., Lorig, K., Smith, M., & Ory, M. (2013). Webinar: Findings from the CDSMP

714   National Study. Retrieved from https://www.ncoa.org/resources/webinar-findings-from-

715   the-cdsmp-national-study/

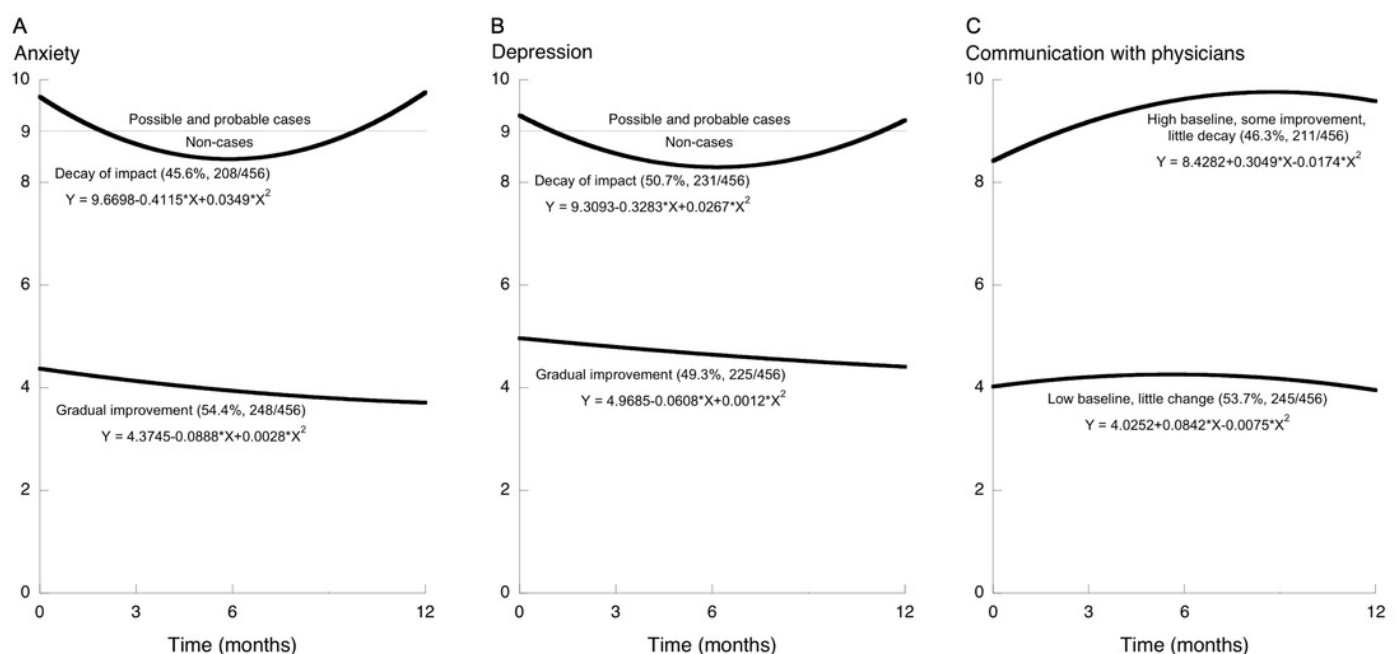
716   Witkiewitz K and Marlatt GA. Relapse prevention for alcohol and drug problems: that was Zen,

717   this is Tao. *American Psychologist* 2004; 59: 224–235.

718

Figure 1

Trajectories of change after health education, showing two trajectory-defined groups for each of the three outcomes

Growth-Mixture Modeling revealed two trajectory-defined groups for each outcome. On anxiety and depression higher scores are worse. On communication with physicians higher scores are better. For each outcome, one of those two groups had improvement followed by deterioration: decay of impact. For Anxiety and Depression, a score of 9 is the cutoff used in Japan to separate non-cases from possible and probable cases.

Descriptive statistics for all CDSMP participants considered together, baseline and follow up over one year

The outcomes discussed here are anxiety, depression, and communication with physicians. Self-efficacy was used as a mediator in subsequent analyses because of its importance in the theoretical basis of the CDSMP.

Factor	Baseline			3 Months			6 Months			12 Months		
	n	Mean	SD	*n*	Mean	SD	*n*	Mean	SD	*n*	Mean	SD
Self-efficacy: 0–60, higher scores are better	456	32.33	12.54	415	34.48	12.02	425	35.16	12.10	404	35.14	12.99
Anxiety: 0–21, lower scores are better	456	6.89	4.25	423	6.27	4.00	428	6.03	4.22	406	6.43	4.62
Depression: 0–21, lower scores are better	456	7.21	3.82	423	6.69	3.64	428	6.53	3.93	406	6.77	4.20
Communication: 0–15, higher scores are better	456	6.22	3.77	422	6.59	3.99	426	6.91	4.11	401	6.77	4.18

Results of growth-curve analysis, all CDSMP participants considered together ($n = 456$)

Notes:

1. Quadratic growth curve fit to data using full-information maximum likelihood (FIML) estimation.

2. Random effects indicate the variability in the fixed effect. For example, the 95% confidence interval for the slope (Time) of Anxiety is (from –.53 to .05) = (Fixed-effect estimate for Time) ± (Random-effect SD for Time x 1.96) = –.24 ± (.15 x 1.96).

3. Interaction terms with time indicate change in outcome for the time-invariant factor (i.e., male, college, partnered, number of diagnoses, and disease history) over time.

Factor	Anxiety			Depression			Communication		
	Coefficient	Std Err	*p*-value	Coefficient	Std Err	*p*-value	Coefficient	Std Err	*p*-value
Fixed effect									
Time	−.24	(.07)	.001	−.14	(.07)	.034	.20	(.06)	.001
Time x Time	.01	(.00)	.001	.01	(.00)	.038	−.01	(.00)	.002
Self-Efficacy	−.10	(.01)	< .001	−.11	(.01)	< .001	.03	(.01)	< .001
Male	−.06	(.41)	.887	.84	(.36)	.018	−.33	(.44)	.450
College Educated	−.97	(.34)	.004	−.04	(.30)	.882	.41	(.34)	.222
Partnered	−.02	(.34)	.947	−.23	(.30)	.443	.02	(.34)	.956
Number of Diagnoses	.37	(.16)	.019	.18	(.13)	.190	.61	(.15)	< .001
History (years)	−.02	(.01)	.164	−.04	(.01)	.004	−.01	(.01)	.643
Male x Time	.01	(.04)	.852	−.05	(.04)	.163	−.03	(.04)	.366
College x Time	.03	(.03)	.328	−.01	(.03)	.655	−.02	(.03)	.424
Partnered x Time	−.04	(.03)	.240	.01	(.03)	.648	.01	(.03)	.688
Number of Diagnoses x Time	.03	(.01)	.062	.02	(.01)	.174	−.01	(.01)	.300
History (years) x Time	.00	(.00)	.267	.00	(.00)	.397	.00	(.00)	.776
Intercept	1.19	(.57)	< .001	1.93	(.48)	< .001	3.98	(.48)	< .001
Random effect									
SD(Time)	.15	(.03)		.14	(.02)		.14	(.02)	
SD(Intercept)	2.83	(.12)		2.26	(.11)		3.14	(.11)	
SD(Residual)	2.64	(.08)		2.45	(.08)		2.17	(.08)	

Factors predicting membership in groups defined by their trajectory after the CDSMP

Notes:

1. Values in parenthesis are standard errors.

2. The results shown are from logistic regression. The 0-1 coding of group membership (which is the dependent variable) reflects the relative magnitudes of the baseline scores. For all outcomes, the group with the lower baseline score is coded "0" and the group with the higher baseline score is coded "1." Thus, the group with less anxiety at baseline is coded "0" while the group with more anxiety at baseline is coded "1." The same is true for depression. In contrast, the group with better communication (higher scores) at baseline is coded "1" and the group with worse communication (lower scores) at baseline is coded "0."

Factor	Anxiety			Depression			Communication		
	Adjusted Odds Ratio		*p*-value	Adjusted Odds Ratio		*p*-value	Adjusted Odds Ratio		*p*-value
Male	.92	(.23)	.730	1.50	(.38)	.112	.62	(.15)	.050
College	.59	(.12)	.009	.83	(.17)	.376	1.17	(.23)	.432
Partnered	.99	(.20)	.947	.78	(.16)	.227	1.17	(.23)	.418
Self-Efficacy	.97	(.01)	< .001	.95	(.01)	< .001	1.02	(.01)	.013
Number of Diagnoses	1.23	(.12)	.032	1.12	(.11)	.257	1.30	(.13)	.006
History (years)	.99	(.01)	.266	.99	(.01)	.244	1.00	(.01)	.683
Intercept	2.66	(.99)	.009	6.37	(2.49)	< .001	.29	(.11)	.001