

Characteristic	β -coefficient	95% confidence interval	P-value
Usage of airmat (vs. no)	3.52	1.10 to 5.95	0.01
Patient age, per 1-year increase	-0.04	-0.18 to 0.09	0.47
Patient gender, female (vs. male)	0.70	-4.50 to 5.91	0.76
Caregiver age, per 1-year increase	-0.02	-0.27 to 0.24	0.89
Caregiver gender, female (vs. male)	-1.23	-4.60 to 2.14	0.42
Family relationship, spouse (vs. others)	1.32	-5.05 to 7.69	0.64
Daily estimated time for caregiving, per 10-hour increase	0.10	-1.80 to 2.00	0.91
National index of long-term care need, the severest grade (vs. other grades)	-3.42	-5.96 to -0.88	0.02
Alzheimer-type dementia (vs. no)	3.10	-5.29 to 11.49	0.42
Violence by patients within past 3 months, (vs. no)	6.84	3.62 to 10.07	0.001

Table 3 Factors associated with total Burden Index of Caregivers score in limited patients who were free from pressure ulcers, using a multivariable linear regression model ($n = 115$)

scores than those who did not (β -coefficient = 3.52, 95%CI: 1.10–5.95, $P = 0.01$). In contrast, patients with the severest grade of long-term care need according to the national index were likely to have lower caregiver burden than patients with less severe grades (β -coefficient = -3.42, 95%CI: -5.96 to -0.88, $P = 0.02$) (Table 3).

Discussion

Our present study quantitatively revealed that informal caregivers whose patients were free from pressure ulcer had a higher caregiver burden than those who failed to prevent pressure ulcer. Of particular note, these results were derived using two validated scales of caregiver burden, the BIC and ZBI, confirming the robustness and reliability of the observed relationship. The association also tended to be clearly defined in analyses using categorized BIC scores, and the rate of patients who were free from pressure ulcers increased with increasing BIC score in a dose-dependant manner.

The participants in the present study were vulnerable and mostly bedridden, suggesting that they had a high risk of developing pressure ulcer [20]. This observed relationship therefore indicates an association between efforts to prevent pressure ulcer and higher caregiver burden. The difference in BIC score, categorized by the state of pressure ulcer, is shown in Table 2. Comparing the difference with other factors, a BIC score of 3.18 was almost six times larger than that from spending an additional 10 hours for home care per day. Preventing pressure ulcer is not only a time-consuming burden but also hard work which requires the knowledge and ability to carry out prevention protocols for informal caregivers [12,20]. Given the strong emphasis placed on the responsibility of caregivers in preventing pressure ulcer in Japan, it seems natural that caregivers who try to prevent pressure ulcer feel a higher degree of strain [21]. Furthermore, subgroup analysis showed a relationship between airmat use and relatively high burden in caregivers whose patients were free from pressure ulcers. All patients in the present study had severely limited activity and mobility, making them candidates for airmat use. Caregivers who prepared airmats for patients before development of pressure ulcers therefore seemed to be more concerned about patient well-being. Results from subgroup analysis indicated that dedicated caregivers were likely to have higher caregiver burden.

These findings potentially raise a conflicting issue in home care settings, given the general acceptance that airmat use reduces caregiver burden [22]. However, caregiver burden is a comprehensive concept which includes not only physical but also psychological burden regarding the 'meaning of caregiving' [8]. Dedicated caregivers are likely to strive not only to care for but also to cure their patients [23]. This attitude may exaggerate the burden with respect to the 'meaning of caregiving', because most vulnerable patients have little chance for recovery. To deal with psychological burden, present efforts to reduce physical burden, such as the introduction of airmat use and home care service, should be complemented by more intensive care towards informal caregivers, including counselling by psychiatric specialists. Home nurses should consider not only patient health but also psychological distress in caregivers. In home care settings, self-administered screening to assess depressive symptoms, such as the short version of the Center for Epidemiological Studies Depression Screening Index, may be feasible [24]. Furthermore, a society-wide paradigm shift from cure to care in managing home care patients appears necessary.

Our results also showed several other factors associated with caregiver burden. A statistically significant relationship was observed between violent behaviour in patients and increased caregiver burden, supporting previous investigations into the association between aggression in patients with Alzheimer-type dementia and caregiver burden [25]. In the subgroup analysis, the severest grade of long-term care need according to the national index was associated with lower caregiver burden. Although not statistically significant, this relationship was also seen in the main analysis. One possible explanation for this relationship may be that less impaired patients might require more continuous care and observation than completely bedridden patients. At the very least, this relationship might imply the intractability of caregiver burden in informal caregivers.

Several limitations of the study warrant mention. First, additional details about pressure ulcers were not ascertained in the medical records, so we were unable to evaluate the influence of other variables in the relationship between pressure ulcer preventing and caregiver burden. Second, we did not investigate the relationship between the preventive status of pressure ulcers and caregiver burden in other countries. Global generalization of our results therefore requires further investigation. Third, the cross-

sectional design of our study prevents us from determining causation or a temporal relationship. As a general limitation of observational studies, adjustment of unknown confounding factors highly associated with the investigated relationships was not possible.

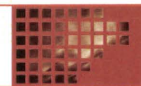
In conclusion, using a multicenter sample of vulnerable patients, our results show that informal caregivers whose patients were free from pressure ulcer are likely to have higher caregiver burden. We interpret our results to mean that the continuous effort required to prevent pressure ulcer might be associated with higher strain in informal caregivers. A society-wide approach towards managing psychological burden in informal caregivers is now necessary.

Acknowledgements

We thank the following home care nurses: Hitomi Kannon, Kazumi Harada, Kayoko Kinoshita, Michiko Nishino, Chieko Yoneda, Kyoko Oshima, Michiyo Otsuki, Atsuko Ueno, Yuko Sugibayashi and Kiyoko Tanjo. This study is supported by research grants from the Yuumi Memorial Foundation for Home Health Care, without restriction on publication.

References

1. Khaw, K. (1999) How many, how old, how soon? *BMJ*, 319 (7221), 1350–1352.
2. Schulz, R. & Beach, S. (1999) Caregiving as a risk factor for mortality: the Caregiver Health Effects Study. *JAMA*, 282 (23), 2215–2219.
3. Caqueo-Urizar, A. & Gutiérrez-Maldonado, J. (2006) Burden of care in families of patients with schizophrenia. *Quality of Life Research*, 15 (4), 719–724.
4. Morimoto, T., Schreiner, A. & Asano, H. (2003) Caregiver burden and health-related quality of life among Japanese stroke caregivers. *Age and Ageing*, 32 (2), 218–223.
5. Serrano-Aguilar, P., Lopez-Bastida, J. & Yanes-Lopez, V. (2006) Impact on health-related quality of life and perceived burden of informal caregivers of individuals with Alzheimer's disease. *Neuroepidemiology*, 27 (3), 136–142.
6. Zarit, S., Reever, K. & Bach-Peterson, J. (1980) Relatives of the impaired elderly: correlates of feelings of burden. *Gerontologist*, 20 (6), 649–655.
7. Abe, K. (2007) Reconsidering the Caregiving Stress Appraisal scale: validation and examination of its association with items used for assessing long-term care insurance in Japan. *Archives of Gerontology and Geriatrics*, 44 (3), 287–297.
8. Miyashita, M., Yamaguchi, A., Kayama, M., Narita, Y., Kawada, N., Akiyama, M., Hagiwara, A., Suzukamo, Y. & Fukuhara, S. (2006) Validation of the Burden Index of Caregivers (BIC), a multidimensional short care burden scale from Japan. *Health and Quality of Life Outcomes*, 4, 52.
9. Bennett, G., Dealey, C. & Posnett, J. (2004) The cost of pressure ulcers in the UK. *Age and Ageing*, 33 (3), 230–235.
10. Posnett, J. & Franks, P. (2008) The burden of chronic wounds in the UK. *Nursing Times*, 104 (3), 44–45.
11. Ferrell, B., Josephson, K., Norvid, P. & Alcorn, H. (2000) Pressure ulcers among patients admitted to home care. *Journal of the American Geriatrics Society*, 48 (9), 1042–1047.
12. Paquay, L., Wouters, R., Defloor, T., Buntinx, F., Debaillie, R. & Geys, L. (2008) Adherence to pressure ulcer prevention guidelines in home care: a survey of current practice. *Journal of Clinical Nursing*, 17 (5), 627–636.
13. Bergstrom, N., Braden, B., Laguzza, A. & Holman, V. (1987) The Braden Scale for Predicting Pressure Sore Risk. *Nursing Research*, 36 (4), 205–210.
14. Black, J., Baharestani, M., Cuddigan, J., Dorner, B., Edsberg, L., Langemo, D., Posthauer, M., Ratliff, C. & Taler, G. (2007) National Pressure Ulcer Advisory Panel's updated pressure ulcer staging system. *Dermatology Nursing*, 19 (4), 343–349; quiz 50.
15. Ikegami, N., Yamauchi, K. & Yamada, Y. (2003) The long term care insurance law in Japan: impact on institutional care facilities. *International Journal of Geriatric Psychiatry*, 18 (3), 217–221.
16. Arai, Y., Kudo, K., Hosokawa, T., Washio, M., Miura, H. & Hisamichi, S. (1997) Reliability and validity of the Japanese version of the Zarit Caregiver Burden interview. *Psychiatry and Clinical Neurosciences*, 51 (5), 281–287.
17. Kumamoto, K. & Arai, Y. (2004) Validation of 'personal strain' and 'role strain': subscales of the short version of the Japanese version of the Zarit Burden Interview (J-ZBI_8). *Psychiatry and Clinical Neurosciences*, 58 (6), 606–610.
18. Rubin, D. (1987) Multiple Imputation for Nonresponse in Surveys. New York: John Wiley & Sons.
19. van Buuren, S., Boshuizen, H. & Knook, D. (1999) Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in Medicine*, 18 (6), 681–694.
20. Reddy, M., Gill, S. & Rochon, P. (2006) Preventing pressure ulcers: a systematic review. *JAMA*, 296 (8), 974–984.
21. Murayama, S., Oe, M., Sanada, H., Okuwa, M., Sugama, J. & Matsui, Y. (2004) Introduction of caregiver ability scale related to pressure ulcer development and evaluation of reliability. *Japanese Journal of Pressure Ulcers*, 6 (4), 647–651.
22. Jonsson, A., Lindén, M., Lindgren, M., Malmqvist, L. & Bäcklund, Y. (2005) Evaluation of antidecubitus mattresses. *Medical & Biological Engineering & Computing*, 43 (5), 541–547.
23. Webb, C. (1996) Caring, curing, coping: towards an integrated model. *Journal of Advanced Nursing*, 23 (5), 960–968.
24. Andresen, E., Malmgren, J., Carter, W. & Patrick, D. (1994) Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). *American Journal of Preventive Medicine*, 10 (2), 77–84.
25. Teri, L. (1997) Behavior and caregiver burden: behavioral problems in patients with Alzheimer disease and its association with caregiver distress. *Alzheimer Disease and Associated Disorders*, 11 (Suppl. 4), S35–8.



Effect of feedback in promoting adherence to an exercise programme: a randomized controlled trial

Masaaki Shakudo MPH,¹ Misa Takegami RN MPH PhD,² Ai Shibata PhD,³ Miki Kuzumaki MT,⁴ Takahiro Higashi MD PhD,⁵ Yasuaki Hayashino MD PhD MPH,⁶ Yoshimi Suzukamo PhD,⁷ Satoshi Morita PhD,⁸ Michio Katsuki MD⁹ and Shunichi Fukuhara MD DMSc FACP¹⁰

¹Postgraduate, ⁶Lecturer, Department of Epidemiology and Healthcare Research, Kyoto University School of Public Health, Kyoto, Japan

²Assistant Professor, Department of Epidemiology and Healthcare Research, Kyoto University School of Public Health, Kyoto, Japan and Research Associate, Institute for Health Outcomes and Process Evaluation Research, Kyoto, Japan

³Research Associate, Faculty of Sport Sciences, Waseda University, Saitama, Japan

⁴Chief of Public Health Service, ⁹Director, Hokuriku Institute of Wellness and Sports Science, Ishikawa, Japan

⁵Associate Professor, Department of Public Health/Health Policy, the University of Tokyo, Tokyo, Japan

⁷Assistant Professor, Department of Physical Medicine and Rehabilitation, Tohoku University Graduate School of Medicine, Miyagi, Japan and Senior Researcher, Institute for Health Outcomes and Process Evaluation Research, Kyoto, Japan

⁸Professor, Department of Biostatistics and Epidemiology, Yokohama City University Medical Center, Yokohama, Japan

¹⁰Professor, Department of Epidemiology and Healthcare Research, Kyoto University School of Public Health, Kyoto, Japan and Scientific Advisory Board, Institute for Health Outcomes and Process Evaluation Research, Kyoto, Japan

Keywords

aerobic, epidemiology, intervention, motivation, prevention, wellness

Correspondence

Masaaki Shakudo
Department of Epidemiology and Healthcare Research
Kyoto University School of Public Health
Yoshida-Konoecho
Sakyo-ku
Kyoto 606-8501
Japan
E-mail: masa-s@tkf.att.ne.jp

Accepted for publication: 17 August 2009

doi:10.1111/j.1365-2753.2009.01342.x

Abstract

Objective We investigated whether providing participants in an exercise programme with regular feedback on their exercise progress affected their adherence to the programme regimen.

Method We conducted a randomized controlled trial. Adult men and women with borderline hypertension and a body mass index ≥ 25.0 were randomized to two intervention groups (groups A and B) and one control group (group C) and were prescribed regular aerobic exercise. During the 12-week study period, group A was provided with both feedback information on their exercise progress and a health letter, while group B was provided with the health letter only. The main outcome measure was exercise performance, per cent achievement of target exercise level (%) defined as the number of weeks during which the exercise target was reached divided by the number of weeks in the programme. Results were compared using the Kruskal–Wallis test.

Results A total of 105 study subjects were randomized into three groups (A, $n = 37$; B, $n = 37$ and C, $n = 31$). Per cent achievement of target exercise level during the 12-week period was highest in group A (26.5%), followed by groups B (22.9%) and C (17.4%) ($P = 0.36$). Subjects who received regular feedback during the exercise programme tended to have higher exercise performance.

Conclusions In improving adherence to exercise intervention, the provision of regular feedback to participants in an exercise programme may be an effective intervention.

Introduction

Reduced physical activity is a major public health concern in many developed countries. The importance of physical activity and exercise in the primary prevention of disease [1,2] is increasingly recognized, and regular exercise and the maintenance of cardiovascular and respiratory function are now considered important components of an integrated approach to the prevention of chronic diseases [3].

Although programmes promoting aerobic exercise are prevalent throughout Japan, concern has been raised that participants may not adhere to the programmes that are of extended duration and may tend to stop exercising particularly soon after the programme is finished. A prior study, for example, showed that more than half of the programme participants stopped exercise 2–4 months after cessation [4]. Effective interventions in improving adherence to exercising remain unknown. Given that the determinants of long-term adherence include the maintenance of motivation, personal

experience of the benefits of exercise and willingness to continue the exercise programme [5], we hypothesized that the provision of regular feedback to participants on their progress may encourage them to continue doing exercise. Studies in other countries have shown that subjects provided with regular feedback on their progress during an exercise programme achieve higher physical activity levels and are less likely to drop out of the programme than those without such feedback; however, a method of specific intervention has yet to be established [6,7]. The establishment of exercise programmes for health promotion thus requires the creation of effective feedback methods and confirmation of their efficacy.

Here, we conducted a randomized controlled trial to assess whether the regular provision of information on the physical effects of exercise and the maintenance of log books recording the details of the exercise affect the adherence to exercise.

Methods

Subjects

Subjects were adult men and women with borderline hypertension and a body mass index (BMI) ≥ 25.0 at five companies in Komatsu City, a coastal city on the Japan Sea, Ishikawa Prefecture, who underwent a routine physical examination from April to May 2005. Borderline hypertension was defined as a systolic blood pressure (SBP) of 130–159 mmHg or diastolic blood pressure (DBP) of 85–99 mmHg, both of which are categorized as high normal to mild hypertension by the Japanese Society of Hypertension.

Study participants were recruited by representatives of each company and the overview of the exercise programme was posted on a company bulletin board. Employees willing to participate in the study were briefed on the contents of the first session and the necessary cautions to be observed. We also explained favourable effects of aerobic exercise on a variety of chronic conditions. A total of 111 participants aged 25–68 years volunteered to participate and all of them agreed to sign the letter of study consent. After the informed consent was obtained from all participants, we measured height, weight and blood pressure, and we conducted submaximal step test to predict maximal oxygen consumption (Vo₂max). Six participants were excluded owing to normal blood pressure or BMI, leaving 105 subjects as participants in the trial.

Study design and intervention

The study was conducted as a randomized controlled trial. The subjects were stratified by company, disease (obesity, hypertension, both) and age (age 20–40, 41–69 years), then randomized into three groups, A ($n = 37$), B ($n = 37$) and C ($n = 31$).

All subjects undertook a home-based aerobic exercise programme during the study period, which included any or all of walking, jogging and swimming at the discretion of the individual. Target exercise dose for each participant was set to meet the exercise recommendations proposed by the American College of Sports Medicine in 1990, namely aerobic exercise of more than 20 minutes per session and three sessions per week [8]. The minimum duration of any particular exercise within a session was set at 10 minutes and above. We set exercise intensity at moderate intensity or higher according to the definition by the American College of

Sports Medicine. An exercise of moderate intensity is equivalent to 3–6 metabolic equivalents [9].

Subjects in group A were provided with feedback based on their exercise and a 'health letter', which is a newsletter on health issues every 2 weeks, while Group B was given the health letter alone every 2 weeks. Group C was a control group that was instructed to submit exercise records without any intervention. Comparison of groups A and B allowed us to evaluate whether the provision of regular feedback to participants in an exercise programme affected their adherence, while comparison of groups B and C indicated the effect of regular contact with participants via periodic health letter. The newsletter contained general health-related topics without specific reference to the effects of exercise.

Content of feedback intervention

We quantitatively analysed exercise records submitted by group A subjects and returned the resulting data sheets to them as feedback via regular mail or fax. Feedback contents included the following items:

- 1 a graph of body weight and blood pressure (daily changes over the preceding 2 weeks; changes during the study period; elapsed exercise time; overall increase or decrease since the start of the programme; and variation from normal weight and blood pressure);
- 2 exercise history (daily exercise achievement represented by coloured boxes on calendar boxes);
- 3 pedometer record (cumulative walking distance illustrated on a map from Komatsu City through famous sightseeing spots in Japan; cumulative steps, average steps and walking distance per day; maximum/minimum steps; and maximum/minimum walking distance); and
- 4 caloric expenditure (cumulative caloric expenditure over the preceding 2 weeks, calculated from the pedometer recordings).

Data collection

We collected the following information at the first study session: sex, age, height, BMI, SBP and DBP, Vo₂max, medical history, history of present illness, current medication, smoking history, drinking habit and stage of change in exercise behaviour [10].

All participants were instructed to record pedometer, weight and blood pressure data by disease and daily exercise achievement on a specific sheet everyday. Daily target exercise level was met if the subjects performed more than 20 minutes' aerobic exercise per day. The subjects submitted the exercise records to the data centre by mail or fax once every 2 weeks, with the due date written on each record sheet in advance to facilitate compliance. When records were more than 3 days overdue, we sent a reminder by mail or fax.

We held the programme-end assessment after the 12-week study period, at which time we evaluated BMI, blood pressure and Vo₂max by conducting submaximal step test for all participants to determine the effects of exercise.

To assess adherence to exercise after the programme ended, participants were sent a questionnaire 1 month after completion of the survey, which investigated how many days a week and for how many hours a day the participants were still performing aerobic exercise.

Table 1 Characteristics of subjects at baseline

	Group A (n = 37)	Group B (n = 37)	Group C (n = 31)
Age	39.3 ± 9.1	41.8 ± 10.9	40.4 ± 10.3
Male	30 (81)	28 (76)	27 (87)
Body mass index (kg m ⁻²)	27.5 ± 3.4	27.0 ± 3.0	26.5 ± 2.8
Systolic blood pressure (mmHg)	140 ± 18	141 ± 17	140 ± 14
Diastolic blood pressure (mmHg)	88 ± 11	86 ± 10	86 ± 11
Vo ₂ max (mL kg ⁻¹ minute ⁻¹)	33.6 ± 3.2	32.6 ± 5.7	34.4 ± 4.3
Present illness	8 (22)	5 (14)	9 (29)
Smoking	15 (32)	11 (49)	11 (36)
Excessive drinking	8 (22)	5 (14)	4 (13)
Habitual exercise	18 (49)	15 (40)	11 (35)

Data expressed as mean ± standard deviation or number of subjects (%).

The main outcome measure throughout the study was the % achievement of target exercise level (%), defined as the number of weeks during which the exercise target was reached divided by the number of weeks in the programme. When records of daily exercise achievement were missing, the subject was defined as not having exercised. Secondary outcomes included BMI, SBP/DPB, Vo₂max, and adherence to exercise and total exercise time (in minutes) per week at 1 month after the end of the programme. Adherence was defined as the maintenance of aerobic exercise for 3 or more days per week and for 20 minutes per day.

We submitted the protocol of this study to the Institutional Review Board of the Faculty of Medicine, Graduate School of Medicine, Kyoto University in March, 2005. This study was approved by the Institutional Review Board of the Faculty of Medicine, Graduate School of Medicine, Kyoto University.

Statistical analysis

We compared the % achievement of target exercise level in the three groups by the Kruskal–Wallis test. We also analysed the data after stratification by age (≤40 or >40 years) at the beginning of the exercise programme.

We compared BMI, SBP/DPB and Vo₂max at the first and second sessions by the paired *t*-test. Statistically significant differences were analysed for variance using the groups (A, B, C) as explanatory variables. Data from subjects taking antihypertensive medications were excluded from analysis of SBP and DPB.

We also compared differences among the groups in the proportion of adherent subjects and in exercise time per week in minutes at 1 month after the exercise programme by the Kruskal–Wallis test.

All analyses were performed using SPSS version 11.5 for Windows.

Results

Subjects

Average age of the 105 subjects was 40.5 ± 10.1 years. The number of subjects in groups A, B and C after randomization was 37, 37 and 31, respectively. Baseline characteristics of the subjects by group showed no statistically significant differences (Table 1).

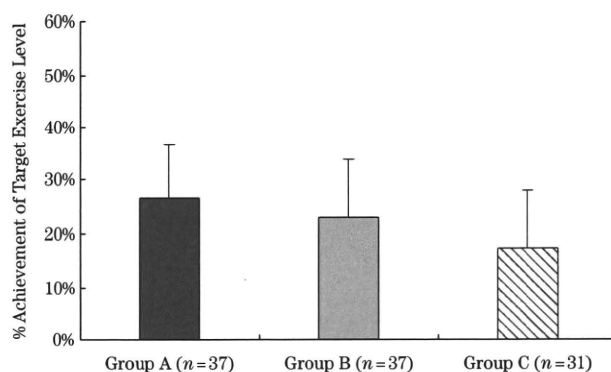


Figure 1 The % achievement of target exercise level. The provision of regular feedback tends to be effective in improving adherence to the programme.

A total of 99 subjects (34 in group A, 35 in group B, 30 in group C) attended the programme-end session.

Main outcome measure

The % achievement of target exercise level during the 12-week period was highest in group A (26.5%), followed by groups B (22.9%) and C (17.4%) ($P = 0.36$) (Fig. 1). Figure 2 compares the % achievement of target exercise level by age group.

Secondary outcome measures

Changes in physiological indicators

Changes in BMI before and after the exercise programme were 0.21 kg m⁻² in group A ($n = 34$), 0.21 kg m⁻² in group B ($n = 34$) and 0.05 kg m⁻² in group C ($n = 30$). BMI after the programme appeared to be lower in each group.

Changes in SBP before and after the programme trended downward by 0.9 mmHg in group A ($n = 30$) and 2.9 mmHg in group B ($n = 32$), whereas that in group C ($n = 29$) trended upward by 1.6 mmHg. Likewise, DBP in groups A and B decreased by 0.4 and 2.3 mmHg, respectively, whereas that in Group C increased by 0.4 mmHg.

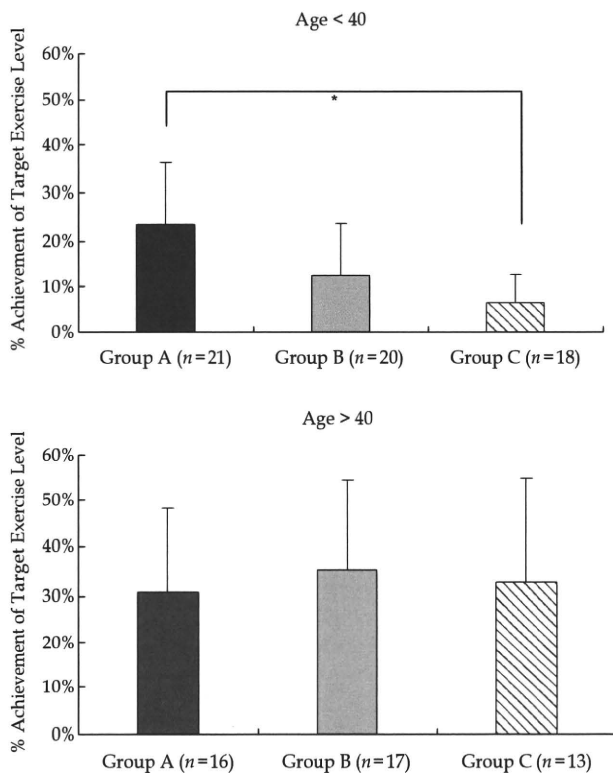


Figure 2 Comparison of the % achievement of target exercise level by age group. There was significantly different ($*P = 0.040$) between groups A (23.4%) and C (6.4%).

Among other variables, Vo_2max increased by $1.0 \text{ mL kg}^{-1} \text{ minute}^{-1}$ in group A ($n = 34$), $1.2 \text{ mL kg}^{-1} \text{ minute}^{-1}$ in group B ($n = 30$) and $0.8 \text{ mL kg}^{-1} \text{ minute}^{-1}$ in group C ($n = 29$).

Maintenance of exercise and total exercise time per week at 1 month

The number of respondents to the follow-up questionnaire evaluating exercise adherence at 1 month after cessation was 46, representing 43.8% of the original study subjects (48.6% in group A, 35.1% in group B and 48.4% in group C). The percentage of subjects maintaining exercise at 1 month was highest in group A (22%), followed by groups B (15%) and C (13%) ($P = 0.76$). Total exercise time per week was 108.1 hours in group A, 64.2 hours in group B and 90.7 hours in group C.

Discussion

In our study, subjects who received regular feedback during an exercise programme tended to have better adherence to the target exercise level than those receiving a periodic health letter only. This finding suggests that the provision of regular feedback may be effective in improving exercise adherence. Furthermore, the results also suggested that the provision of a newsletter to contact subjects alone may also promote exercise adherence.

In this study, stratified analysis by age indicated that subjects aged over 40 years performed exercise more often than those below 40 throughout the study. Subjects aged over 40 years may have been more concerned about their health, with greater motivation to stay fit, with or without the intervention. In subjects aged below 40, the finding that group A attained a higher target level than group C suggests that the provision of regular feedback in an exercise programme is more effective for younger participants. For younger people who are relatively less concerned about their health, active intervention to keep them motivated may enhance adherence to exercise.

Body mass index in groups A and B tended to decrease in a similar manner after the exercise programme, suggesting that the regular provision of a newsletter alone may be as effective for weight control as feedback. BMI is determined by other factors, such as diet and basal metabolism. Given that the effect of exercise on weight loss is approximately 30% of that of diet, and that an increase in physical activity alone does not easily contribute to weight loss [11], the present intervention may have been insufficient to bring about a visible effect on BMI.

Blood pressure tended to decrease after the exercise programme in both groups A and B, suggesting that blood pressure was unaffected by the provision of feedback. In contrast, Fagard reported that aerobic exercise under direct observation decreased SBP by 7.4 mmHg and DBP by 5.8 mmHg [12]. This difference in findings may have been due to the fact that our subjects performed exercise without direct observation. Given that exercise may be an effective intensity level to exert a hypotensive effect, the lack of direct observation in the present programme may have meant that intensity was insufficient, notwithstanding some increase in physical activity brought about by the feedback intervention itself. Consistent with a previous report [13], we further observed that blood pressure showed weekly fluctuations despite an overall downward trend, suggesting that short-term changes in blood pressure may not be a suitable monitor of the effects of exercise.

The feedback method used in this study was created to overcome various barriers to exercise adherence, such as time constraints on visits to gyms and poor accessibility [14]. The idea of this feedback contents can be meaningful in future study. Further investigation is required on the optimal frequency, period and content of the feedback intervention as a means of encouraging more voluntary exercise. It has been reported that a computer-based feedback system reduced the drop-out rate from an exercise programme [15], suggesting that contacting participants more quickly through text messages or email may be a more effective way of providing feedback. A previous study also suggested that the lack of enjoyment or amusement, as well as the lack of motivation, may reduce exercise adherence, especially when the programme is prolonged [16]. This finding highlights the importance of promoting adherence by making the programme more rewarding and less boring, such as by changing the content and providing feedback during the course of the programme, particularly for participants with no established exercise habit.

Several limitations deserve mention. First, the exercise records were filled out by the subjects themselves, so misclassification cannot be ruled out and we also cannot measure the subjects' exercise intensity accurately in the case of prescribing home-based

exercise; this issue can be the problem to be solved. Second, it is not clear if the study period was long enough to allow the effect of the intervention to be assessed. Third, although we found the provision of regular feedback tends to be effective intervention in improving adherence to the participants, the sample size was not sufficient to allow the detection of statistically significant differences among the groups. On the basis of the two-sided test ($\alpha = 0.05$, power of 80%), the enrolment of at least 172 subjects is required to detect the observed difference in exercise performance between groups A and C.

This study showed that the provision of regular feedback to participants in an exercise programme may be effective in improving adherence. Conventional exercise programmes can be valuable to bring in a feedback mechanism, such as that used in this study. Future studies with more subjects should confirm these findings with greater statistical power and explore an optimal intervention method of feedback.

Acknowledgements

This study was funded by the Hokuriku Institute of Wellness and Sports Science.

Conflict of interest

Masaaki Shakudo, Takahiro Higashi, Michio Katsuki and Shunichi Fukuhara have indicated no financial conflicts of interest.

Contributors

Masaaki Shakudo worked out research plan of this study and analysed the data, and also carried out the management of the study.

Misa Takegami advised the idea and design of this study, and also analysed and interpreted the data.

Takahiro Higashi and Yasuaki Hayashino advised the idea of this study, and also analysed and interpreted the data.

Yoshimi Suzukamo advised research plan and interpreted the data.

Satoshi Morita designed analysis plan of this study.

Ai Shibata, Miki Kuzumaki and Michio Katsuki carried out the management of the study.

Hokuriku Institute of Wellness and Sports Science provided research field and obtained funding management of the project.

Shunichi Fukuhara advised research plan, design and management of the project, and also obtained funding for this study.

Furthermore, all authors helped to interpret the findings and review the drafts of this paper.

References

1. United States Department of Health Services (1996) Physical Activity and Health: A Report of the Surgeon General. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion.
2. Blair, S. N. (1993) Physical activity, physical fitness, and health. *Research Quarterly for Exercise and Sport*, 64, 365–376.
3. Ball, K., Bauman, A., Leslie, E. & Owen, N. (2001) Perceived environmental aesthetics and convenience and company are associated with walking for exercise among Australian adults. *Preventive Medicine*, 33, 434–440.
4. Ward, A. & Morgan, W. (1984) Adherence pattern of healthy men and women enrolled in an adult exercise program. *Journal of Cardiac Rehabilitation*, 4, 143–152.
5. Sallis, J. F. & Hovell, M. F. (1990) Determinants of exercise behavior. *Exercise and Sport Sciences Reviews*, 18, 307–330.
6. Roemmich, J. N., Gurgol, C. M. & Epstein, L. H. (2004) Open-loop feedback increases physical activity of youth. *Medicine and Science in Sports and Exercise*, 36, 668–673.
7. Duncan, K. & Pozehl, B. (2003) Effects of an exercise adherence intervention on outcomes in patients with heart failure. *Rehabilitation Nursing*, 28, 117–122.
8. American College of Sports Medicine Position Stand (1990) The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness in healthy adults. *Medicine and Science in Sports and Exercise*, 22, 265–274.
9. Pate, R. R., Pratt, M., Blair, S. N., *et al.* (1995) Physical activity and public health. A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *The Journal of the American Medical Association*, 273, 402–407.
10. Oka, K. (2000) Recent trends of research on exercise adherence utilizing the transtheoretical model of behavior change. *Journal of Physical Education*, 45, 543–561.
11. Ross, R., Freeman, J. A. & Janssen, I. (2000) Exercise alone is an effective strategy for reducing obesity and related comorbidities. *Exercise and Sport Sciences Reviews*, 28, 165–170.
12. Fagard, R. H. (1999) Physical activity in the prevention and treatment of hypertension in the obese. *Medicine and Science in Sports and Exercise*, 31 (Suppl. 11), S624–S630.
13. Kiyonaga, A., Arakawa, K., Tanaka, H. & Shindo, M. (1985) Blood pressure and hormonal responses to aerobic exercise. *Hypertension*, 7, 125–131.
14. Trost, S. G., Owen, N., Bauman, A. E., Sallis, J. F. & Brown, W. (2002) Correlates of adults' participation in physical activity: review and update. *Medicine and Science in Sports and Exercise*, 34, 1996–2001.
15. Annesi, J. J. (1998) Effects of computer feedback on adherence to exercise. *Perceptual and Motor Skills*, 87, 723–730.
16. Robinson, J. I. & Rogers, M. A. (1994) Adherence to exercise programmes: recommendations. *Sports Medicine*, 17, 39–52.

その他：編集委員会依頼原稿（原著）

コンセンサスパネルを用いた 肺がん検診チェックリストの適切性検討と改訂案作成

東 尚弘*¹、町井涼子*²、青木綾子*²、濱島ちさと*²、中山富雄*³、佐川元保*⁴、斎藤 博*²

国のがん対策の一つの柱にがん検診受診率向上があげられているが、検診の効果を確保するには適切な精度管理が必要である。厚生労働省「がん検診に関する検討会」では検診機関に対して「事業評価のためのチェックリスト」を策定・普及を図っているが、継続的な精度向上のためにはチェックリストもまた適切性の検証と改善を繰り返すことが望まれる。今回、評価基準を定める際の国際的な標準的手法を用い、全国各地の検診従事医師8名のコンセンサスパネルでチェックリスト検証と改訂案策定を行った。結果、現行のチェックリストは、一項目（「1日あたり実施可能な人数を明らかにしている」）以外は適切という評価であったが、さらなる改善のため検討会議では様々な修正・追加が行われた。本過程は現段階では研究上の試みであるが、地域的に多様な委員による検討と定型的な手順に基づく合意がまとめられており、次回の公式的な改訂時に活用可能と考えられる。

キーワード： 肺がん検診、事業評価、精度管理、コンセンサスパネル、ヘルスサービス研究
J Thorac CT Screen 2010;17:106-114

平成19年施行されたがん対策基本法^[1]に基づき発行された国のがん対策推進基本計画では、5年以内にがん検診受診率を50%とすることが目標とされている^[2]。現状では平成19年国民生活基礎調査によると肺がん検診の受診率は40歳以上男性25.7%、女性21.1%^[3]と目標へは隔たりがあり、検診の必要性の認知を広げていくことが望まれる。

検診の受診率向上の一方で、検診の効果を確保するためには検診受診者が適切な検診を受けられるよう検診提供体制の質・精度管理も重要

である。その認識のもと、厚生労働省「がん検診に関する検討会」では、自治体を実施する住民検診を対象とし、「事業評価のための点検表（チェックリスト）」を乳がん、子宮がん、大腸がん、胃がん、肺がんと順次策定した^[4]。このチェックリストは、都道府県、市町村、検診実施機関版が設けられ、それぞれ満たすべき事項が評価基準として定められており、主に検診の技術・体制的指標の達成度の自己点検を図るための基準である。これらの周知徹底により検診の質の向上、精度管理が図られている。

チェックリストなど一般に評価指標を策定する際には、その妥当性を確保し、かつ現場での順守を促す意味でも現場の視点により評価することが重要である。しかも、その評価と改善の作業は継続的な改善を旨として繰り返し行うことが望ましい。今回、その手始めとして実際に検診実施機関に勤務して検診業務に従事してい

*¹ 東京大学医学系研究科健康医療政策学
〒113-0033 東京都文京区本郷7-3-1
e-mail: higashi@m.u.tokyo.ac.jp

*² 国立がん研究センターがん予防・検診研究センター検診研究部

*³ 大阪府立成人病センターがん予防情報センター疫学予防課

*⁴ 金沢医科大学呼吸器外科

る医師の意見によりチェックリストの適切性の評価を行い、さらに、次回改訂する場合の案を作成したので報告する。

対象と方法

＜概要＞

本研究は、米国RAND/UCLAで開発された適切性評価法(デルファイ変法)と呼ばれる方法に従ってコンセンサスパネルによりチェックリストの検証を行った^[5,6]。この方法は質評価の指標作成など、客観的な合意に基づく項目選択に使用され、判定の再現性^[7]や、本手法により作成された医療の質評価基準においてはその基準を用いた測定結果の予後予測における妥当性^[8]が示されている。本手法の一般的な全体の流れをFig. 1に示す。今回の具体的手順としては、まず様々な背景を持つ委員を8名選定、現行のチェックリストを各委員に送付し1～9のスケールで各項目の適切性の評価を依頼した。この第1回の評価は、原稿のチェックリストに対する委員の評価として機能する。さらに、その集計を元に3時間超の検討会議を行い、チェックリストとして改善すべき点については修正を加え、またそれらの適切性に関する議論をおこなった。これらの過程をふまえて再評価を行い、合意が得られた項目がチェックリスト改訂案として決定された。ただし、この改訂案は即公的

なチェックリストに反映されるものではなく次回改訂が提起された際の資料として扱われることは始めに確認した。

＜検診専門家パネルの選定＞

多様な意見を収集するために地理的に東日本と西日本を幅広く含めた関連各者を専門家委員会(以下パネル)に選定し、また、検診実施機関および大学関係者も選定した。実際のパネルのメンバーは以下の通り(敬称略、勤務先所在地順)となった。

国立病院機構北海道がんセンター	原田真雄
東北大学加齢医学研究所	遠藤千顕
ポートスクエア柏戸クリニック	瀧澤弘隆
化学療法研究所附属病院	小中千守
(初回評価のみ)	
石川県予防医学協会	木部佳紀
岐阜環境医学研究所	松井英介
岡山県健康づくり財団附属病院	西井研治
香川県立保健医療大学	佐藤 功

＜現行のチェックリスト＞

今回評価対象としたチェックリスト項目は、がん検診に関する検討会で作成された肺がん検診のためのチェックリスト(検診機関用)26項目^[4]であり、それぞれ、受診者への説明に関して4項目、問診および撮影の精度管理に関し

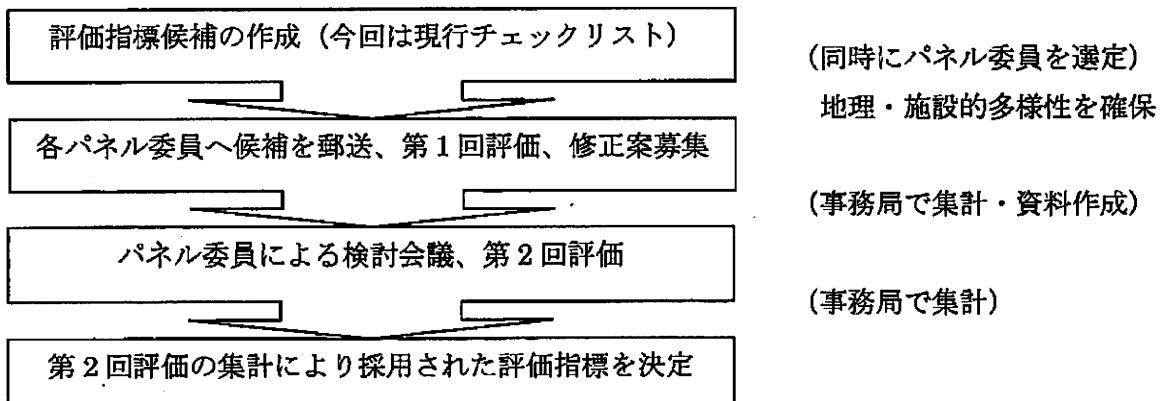


Fig. 1 The flow of the RAND/UCLA appropriateness method

て6項目、X線撮影の精度管理に関して5項目、
喀痰細胞診の精度管理に関して7項目、システ
ムとしての精度管理に関して4項目からなっ
ていた。それぞれの項目は、第1回の評価対象となっ

ており、Table 1 に示す。

<評価過程>

各パネル委員に会議の約2週間前に、Table 2

Table 1 Current checklist for lung cancer screening provider facilities, proposed revisions, and appropriateness ratings

現行チェックリスト	評価*	会議後、チェックリスト改訂案	評価*
1 受診者への説明			
(1) 要精密検査となった場合には、必ず精密検査を受ける必要があることを事前に明確に知らせているか	9, A	受診者全員に対して、肺がん検診が有効であるためには、要精密検査となった場合には必ず精密検査を受ける必要があることを検査前に明確に知らせているか	9, A
(2) 精密検査の方法や内容について説明しているか	8, A	受診者全員に対して、検査前に精密検査の方法や内容について説明しているか	9, A
(3) 精密検査の結果の市町村への報告などの個人情報の取り扱いについて、受診者に対し十分な説明を行っているか	8, A	修正なし	9, A
(4) 禁煙及び防煙指導等、肺がんに関する正しい知識の啓発普及を行っているか	7, I	修正なし	9, A
2 問診および撮影の精度管理			
(1) 検診項目は、問診、胸部X線検査、および喀痰細胞診を行っているか	9, A	検診項目として、問診、胸部X線検査、および喀痰細胞診を行っているか	9, A
(2) 問診は喫煙歴および血痰の有無を聴取しているか	9, A	問診の中で喫煙歴および血痰の有無を聴取しているか	9, A
(3) 問診記録は少なくとも5年間は保存しているか	9, A	修正なし	9, A
(4) 肺がん診断に適切な胸部X線撮影を行っているか注1)	9, A	修正なし	9, A
(5) 撮影機器の種類（直接・間接撮影、ミラー・I.I.方式等）、フィルムサイズを明らかにしているか注2)	9, A	撮影機器の種類（直接・間接撮影、デジタル方式等）に応じた撮影法を行っているか（注2参照）	9, A
(6)			
3 胸部X線撮影の精度管理			
(1) 2名以上の医師によって読影し、うち一人は十分な経験を有した呼吸器または放射線の専門医を含めているか	9, A	2名以上の医師によって読影し、うち一人は十分な経験を有した熟練した呼吸器科または放射線科の医師を含めているか	9, A
(2) 2名のうちどちらかが「要比較読影」としたものは、過去に撮影した胸部X線写真と比較読影しているか	9, A	修正なし	9, A
(3) 比較読影した症例数を報告しているか	8, I		
(4) X線写真は少なくとも3年間は保存しているか	9, A		
(5)		胸部X線写真は少なくとも5年間は保存しているか	9, A
(6) X線検査結果は少なくとも5年間は保存しているか	9, A	胸部X線検査結果は少なくとも5年間は保存しているか	9, A
4 喀痰細胞診の精度管理			
(1) 喀痰細胞診は、年齢50才以上喫煙指数400もしくは600以上、あるいは年齢40才以上6ヶ月以内に血痰を有したものの、その他職業性など高危険群と考えられるものに行っているか	9, A	修正なし	9, A
(2) 細胞診の業務を委託する場合は、その委託機関（施設名）を明記しているか	9, A	修正なし	9, A
(3) 採取した喀痰は、2枚のスライドに塗抹し、湿固定の上、パニコロウ染色を行っているか	9, A	修正なし	9, A
(4) 固定標本の顕微鏡検査は、日本臨床細胞学会の認定を受けた細胞診専門医と細胞検査士が連携して行っているか注3)	9, A	修正なし	9, A

Table 1 Current checklist for lung cancer screening provider facilities, proposed revisions, and appropriateness ratings

(5)	がん発見例は、過去の細胞所見の見直しを行っているか	9, A	修正なし	9, A
(6)	標本は少なくとも3年間は保存しているか	9, A		
(6)'			標本は少なくとも5年間は保存しているか	9, A
(7)	喀痰細胞診検査結果は少なくとも5年間は保存しているか	9, A	修正なし	9, A
5	システムとしての精度管理			
(1)	精密検査結果及び治療結果の報告を、精密検査実施機関から受けているか注4)	9, A	精密検査結果(組織型や病期等)及び治療結果の報告を、精密検査実施機関から追跡収集する体制があるか	9, A
(2)	診断のための検討会や委員会(第三者の肺がん専門家を交えた会)を設置しているか	8, 5, A	診断のための検討会や委員会(外部の肺がん専門家を交えた会)を設置・参加しているか	9, A
(3)	都道府県がプロセス指標(受診率、要精検率、精検受診率、がん発見率、陽性反応適中率)に基づく検討ができるようデータを提出しているか	9, A	都道府県がプロセス指標(受診率、要精検率、精検受診率、がん発見率、陽性反応適中率)に基づく検討ができるようデータを提出することができるか	9, A
(4)	実施主体へのがん検診の集計・報告は、地域保健・健康増進事業報告に必要な項目で集計しているか	9, A	修正なし	9, A
(5)			要精検者に対して、結果通知時に精密検査の重要性を個別に知らせているか	9, A
(6)			検診結果の通知を実施する場合には、正確な通知を行うためのチェック体制があるか(責任者の明確化を含む)	9, A

注1) 肺がん診断に適切な胸部X線撮影：日本肺癌学会編集、肺癌取り扱い規約 改訂第6版より
 背腹方向撮影1枚による場合、適切な胸部X線写真とは、肺尖、肺野外側縁、横隔膜、肋骨横隔膜角などを含むように正しく位置づけされ、適度な濃度とコントラストおよび良好な鮮鋭度を持ち、中心陰影に重なった気管、主気管支の透亮像ならびに心陰影及び横隔膜に重なった肺血管が観察できるもの

注2) 撮影法：日本肺癌学会編集、肺癌取り扱い規約 改訂第6版より

1：間接撮影の場合は、100mm ミラーカメラと、定格出力150kV以上の撮影装置を用いて120kV以上の管電圧により撮影する。やむを得ず定格出力125kVの撮影装置を用いる場合は、110kV以上の管電圧による撮影を行い縦隔部の感度を肺野部に対して高めるため、希土類(グラデーション型)蛍光板を用いる。定格出力125kV未満の撮影装置は用いない

2：直接撮影の場合は、被検者-管球間距離を1.5m以上とし、定格出力150kV以上の撮影装置を用い、120kV以上の管電圧及び希土類システム(希土類増感紙+オルソタイプフィルム)による撮影がよい。やむを得ず100~120kVの管電圧で撮影する場合も、被曝軽減のために希土類システム(希土類増感紙+オルソタイプフィルム)を用いる

3：CRの場合は、120kV以上の管電圧及び散乱線除去用格子比12:1以上を使用して撮影し、適切な階調処理、周波数処理、ダイナミックレンジ圧縮処理などを施した画像として出力する事が望ましい

注3) 日本臨床細胞学会 細胞診精度管理ガイドライン参照

注4) 組織や病期把握のための治療など

* (中央値、一致度：A：一致、D：不一致、I：どちらでもない)

注5) 太字部分は変更点、灰色は評決により不採用となった項目

Table 2 Excerpt of the rating sheet

チェックリスト候補	チェックリストとしての適切性	コメント
1 受診者への説明		
(1) 要精密検査となった場合には、必ず精密検査を受ける必要があることを事前に明確に知らせているか	1 2 3 4 5 6 7 8 9	
(2) 精密検査の方法や内容について説明しているか	1 2 3 4 5 6 7 8 9	
(3) 精密検査の結果の市町村への報告などの個人情報取り扱いについて、受診者に対し十分な説明を行っているか	1 2 3 4 5 6 7 8 9	

に例示したような評価スケールを評価過程と評価基準の説明書と共に電子メールで送付した。チェックリストの適切性としては、妥当性（現状の予算構成や状況においても検診実施機関はチェックリストを100%満たすべきか、また、その通りしていない場合に何らかの指導などが必要といえるか）、と明確性（要求されている事項の内容に曖昧さがないか）の2軸を定義して、その総合点として各パネル委員が判断して適切性を評価するように依頼した。評価結果は国立がんセンターがん予防・検診研究センター検診研究部に設けられた事務局で検討会議の1週間前までに回収し、会議当日までに集計結果を用意した。評価はこれまでのRAND/UCLA適切性評価法に従い各項目の「チェックリスト」としての適切性を「1=きわめて不適切、9=きわめて適切」のスケールで定量的な評価を依頼した（Table 2）。さらに、それぞれのチェックリストに関してコメント記入欄を設け、新しい意見、チェックリストの表現改訂案、また新しいチェックリストの追加の提案などの意見を収集し、検討資料としてまとめる作業を行った。検討会議の資料としては検討パネル委員の評点の分布およびコメントの一覧を示したが、どの委員が、何点の評点を付けたのか、またどのコメントを出したのかはわからないようにしてまとめた。

検討会議では、一つ一つのチェックリストについて評点の分布とコメントを吟味しながら検討した。検討が終わる毎に再度1~9のスケールでそのチェックリストとしての適切性を評価した。検討の過程でチェックリストの修正が提起された場合にはその場で反映させ、第2回の評価は反映後のチェックリストを対象とした。また、同時に追加が提案された項目についても検討し同様のスケールで評価した。

第2回の評点を集計し、中央値が7以上であり反対意見（1~3の評点）を付けた委員が3名以上でない項目を採用チェックリスト項目とし

た。また、同じ内容で期間のみが異なる場合など両方成り立たない複数の項目が採用された場合には、より高い中央値のあるもの、またはより合意の度合いが高いものを採用とした。最終的な結果は委員全員に回覧して承認を得た。

結果

<現状チェックリストに対する評価>

現行チェックリストへの評価の中央値と一致度をTable 1左欄に示し、チェックリスト改訂案とその評価について右欄に示す。第1回目の評価において、基準に従って不適とされた現行チェックリストは1項目あり、2-(6)「1日あたりの実施可能人数を明らかにしているか」であった。検討会議では1日あたりの実施可能人数は様々な条件によって変化するものであり、このような数字を公表・報告しても正確さに疑問があるとされた。この項目は、関連項目として修正することも困難と判断され、そのまま第2回の評価を行い不採用となった。

<現状に対する議論>

まず受診者に対する説明の章のなかで、受診者へのインフォームドコンセントを行う責任は検診対象者に受診を促す役割をもつ市町村にあるのか、実際の検診を提供する検診実施機関にあるのかが議論された。本来住民検診の実施主体は市町村であり、市町村が説明を行うべきものである。しかし検診対象者への連絡の時点から検診実施機関が委託を受けて実施している例も多く、実務的には説明を検診実施機関が行わざるを得ない。そもそも説明を行う責任の所在がどこにあるのか、統一的な結論を出すことの難しさが明らかになった。

また、検診対象者全員への説明と、要精検と判定されたものへの説明を概念的に区別することの必要性が議論された。前者はインフォームドコンセントという対象者保護や人権の尊重と

しての意味合いが強いものに対して、後者は精検受診勧奨という検診の精度管理という意味合いが強い。検診の確実な実施という意味での必要性は後者の方が強い意味をもつが、この点は現行のチェックリストに含まれていなかったため、この要精検者への重要性の説明が、「システムとしての精度管理」という章の中に追加された。

<チェックリストの修正・追加と第2回評価>

当初の26項目のチェックリストのうち12項目に対して何らかの修正が入り、4項目が追加されて、計30項目が第2回評価を受けた。そのうち4項目については不採用となり、最終的に26項目が採用された。検討中の議論は良く収束しており、採用された項目は全て半数以上の評価委員が「9=きわめて適切」と評価しており、6以下の評点を付けたものは全て2名以下であった (Table 1)。

考 察

本研究ではRAND/UCLA適切性評価法に従い、様々な地域の検診医師によるコンセンサスパネルの手法で肺がん検診事業評価のためのチェックリストの適切性評価と改訂案作成を行った。第1回の評価においては、現状のチェックリストは概ね適切と評価されたが、検討会議の過程で約半数のチェックリスト項目に関して表現の修正が提案され、また4項目が追加された。採用されたチェックリストは現行のチェックリストに現場の医師の意見を反映したものとなっており、次回改訂時にはその貴重な基礎資料となると考えられる。

議論の過程では、検診が非常に多様な制度の下で行われていることが再認識され、一律のチェックリストを設定することの難しさが明らかになった。まず、様々な地域のなかで検診には大きく分けて市町村の事業として行われる一般住民を対象とした住民検診と、戦域で労働安

全衛生法に基づく健康診断に付加して行われる戦域検診がある。概念上はこの両者を区別されるものであるが、実際市町村、戦域共に検診の実施については検診実施を専門とする検診実施機関へ委託することも多く、その委託に際して責任の分担が不明瞭であることも多い。戦域において行われる胸部X線検診が、がん検診なのか結核検診なのか、などが明確に区別されることも少ない。しかし、今回は、チェックリストの使用が行政的な文脈を想定されているため、住民検診を対象とすることで一致した。

しかし、市町村検診のなかでも、検診対象者に受診勧奨を通知する部分を市町村が行う地域や、検診実施機関が委託されて行う地域など様々である。そのため、チェックリストの最初の検診受診者への説明の中で、例えば「要精密検査となった場合に精検を受ける必要があることを説明する」が誰の責任かという点について、まず問題とされた。本来、これは検診受診勧奨を行う際に説明するのが妥当であり、市町村が受診勧奨を行っている場合には、すでに検診を受けようとの意思をもって検診現場に来院した受診者に対して、検診実施機関がインフォームドコンセントを行うことは順序が不適切ではないかという考え方もあった。しかし、これらの説明は非常に重要であることから、市町村の責任であると同時に検診実施機関の責任があっても良いという考え方で概ね一致した。

次にX線写真の読影をどのような資格を持った医師が行うのかという項目に対して問題が明らかになった。現行のチェックリストは「呼吸器又は放射線科の専門医」となっているが、学会の認定した専門医なのか、どの学会の専門医なのか、専門医を取得していないと問題なのか、などが議論になり、結局は「専門医」の言葉は取り除かれ「医師」となった。議論のなかではこのような読影のための「専門医」制度を作ったのは我々関連学会の責任という意見もあ

り、今後の制度整備の必要性が明らかになった。

今回の結果で引き続き課題となる可能性のあるものは、胸部X線写真および細胞診標本の保存期間である。医療法施行規則第21条の5ではX線写真の保存期間は、一般の医療機関において2年とされている^[9]。また、保険医療機関及び保険医療費担当規則9条では、「療養の給付の担当に関する帳簿及び書類その他の記録」として3年保存とされている^[10]。細胞診標本については、結果記録は診療録に準じた扱いを受けると解釈できるものの、標本そのものに関しては、特に法的に定められたものは無い。今回現行のチェックリストでは両方とも保存期間を3年としていたが、専門家パネルではそれが5年とすべきであると言う結果になっている。これらは検診機関においてはその精度管理上、たとえば過去のX線写真との比較読影の上で必要性が高く特に必要な事項であると考えられること、また電子的な保存が可能になったことで以前よりも延長して医療の役に立てることが可能になったためとも考えられるが、実行可能性などとの関連で、今後の議論が必要となると考えられる。

RAND/UCLA 適切性評価法と呼ばれる今回の定式手順には、定式化されない通常の会議に比していくつかの利点がある。まず、スケールを使った準備と決定の手順をとることで、主観的でわかりにくい思考内容が具体化される点である。もちろん各評価者のスケールのとらえ方が異なるために基礎的なばらつきや誤差が存在するのは否めないが、一定の傾向や意見の分布を示すには十分な可視化と考えられる。また、スケール化するためには、何を評点としているのかを必然的に明確化しなければならないため、チェックリスト項目とは何か、という点が最初に定義されその点でも議論の方向性が明確化される。次に、事前の評価により問題点について各委員が独自に考慮する機会が持たれ、さらにそれらが事前に明らかになって効率的に検討会

が進められる点である。さらに、委員の構成を意識的に多様化することで様々な視点からの意見を集約することが出来る。また評価の集計や個別コメントの提示の際には各委員がどの評点・コメントを出したのかを知らせないことで自由な議論や評価が可能になり不本意な他人からの影響のない自立した意見を表出することが可能である。結論の決定方法もあらかじめ定められており恣意性が入る余地はない。

逆に、このような試みの限界としては、検討パネル委員の構成や人選によって結果が変わる懸念がある。そのような懸念を解消するためには検討する人の数を増やして多くの意見を集めるのが良いかもしれない。しかし、そうすると逆に検討会で一人一人が意見を表明する責任や機会が薄れてしまうため議論の深みが生まれないうリスクが出てくる。そのため先行研究では多様性を確保したうえで10名前後の検討委員にとどめることが通例となっている。また広い意見を収集するために一旦10名前後の検討委員で結論を出した後に、別の場で再度承認を取られることもある。今回作成されたチェックリスト改訂案も最終結論ではなく、今後全国的に使われる前に厚生労働省の検討会などでの最終検討が行われるものと期待される。

本手法の最終効果としての死亡率減少については、検診における精度管理のためにチェックリストを使用する経験自体が新しいため、改訂の効果はもとよりチェックリストの順守による効果を検証した報告もまだみられていない。しかし、宮城県において肺がん検診の精度管理について市町村に対して質問紙による調査を繰り返したところ、市町村における肺がん検診の精度管理が向上したことが報告されており^[11,12]、このようなチェックリストの公表は有用であると考えられる。また、海外においてはコンセンサスパネルの手順により医療の質指標を作成し、その指標により測定された医療の質がその後の

生存率と関連したという報告もあり¹⁸⁾、本邦による客観的な専門家の合意抽出によるチェックリストの使用が最終的に死亡率減少にもたらす効果も期待される。

一方で、今回のようなチェックリストを検証・改訂する活動は継続的な検証と改善の手始めという位置づけであり、今後も引き続き行っていく必要があると考えられる。検診事業そのものが一定の評価指標をもとに評価・改善を繰り返していくべきであると同様、評価指標も一定の期間毎に見直し・改訂を行うのが必須である。公的な検討会における結論も最終結論ではなく、検診を取り巻く状況や全国の検診精度が変化すれば再び改訂していく必要がある。また、そのために様々な研究活動の中で問題提起を行っていくことが、より良い評価活動と精度の向上に役立つと考えられる。

以上、我々は肺がん検診の事業評価のためのチェックリストを定式化された手順に従って評価、改訂案の作成を行った。概ね現行のチェックリストは適切と評価されたものの、改善の意見を問われた検討委員は多くの項目について改訂の提案を行った。今後このような意見をもとによりよい検診の質管理、向上がなされ、検診の有効性が高まることが期待される。

謝 辞

著者一同、研究に参加して頂いたパネル委員の皆様へ感謝致します。本研究は厚生労働省科学研究費補助金の助成を得ている。

文 献

- [1] がん対策基本法. <http://law.e-gov.go.jp/announce/H18HO098.html>
- [2] 厚生労働省：がん対策推進基本計画. <http://www.mhlw.go.jp/shingi/2007/06/dl/s0615-1a.pdf>
- [3] 厚生労働省：国民生活基礎調査. <http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa07/3-8.html>
- [4] がん検診事業の評価に関する検討会：今後のわが国におけるがん検診の事業評価の在り方について. <http://www.mhlw.go.jp/shingi/2008/03/s0301-4.html>
- [5] Fitch K, Bernstein SJ, Aguilar MD, et al: The RAND/UCLA Appropriateness Method User's Manual: RAND, 2001
- [6] Park RE, Fink A, Brook RH, et al: Physician ratings of appropriate indications for six medical and surgical procedures. *Am J Public Health*. 1986;76:766-772
- [7] Shekelle PG, Kahan JP, Bernstein SJ, et al: The reproducibility of a method to identify the overuse and underuse of medical procedures. *N Engl J Med*. 1998;338:1888-1895
- [8] Higashi T, Shekelle PG, Adams JL, et al: Quality of care is associated with survival in vulnerable older patients. *Ann Intern Med*. 2005;143:274-281
- [9] 医療法施行規則. <http://law.e-gov.go.jp/htmldata/S23/S23F03601000050.html>
- [10] 保険医療機関及び保険医療養担当規則. <http://law.e-gov.go.jp/htmldata/S32/S32F03601000015.html>
- [11] Sagawa M, Endo C, Sato M, et al: Four years experience of the survey on quality control of lung cancer screening system in Japan. *Lung Cancer*. 2009;63:291-294
- [12] 佐川元保、遠藤千顕、佐藤雅美、他：成人病検診管理指導協議会肺がん部会による肺癌集検に関わる精度管理調査. *肺癌* 2004;44:91-94

Evaluation and Revision of Official Checklists for Lung Cancer Screening Facilities in Japan

Takahiro Higashi^{*1}, Ryoko Machii^{*2}, Ayako Aoki^{*2}, Chisato Hamashima^{*2}
Tomio Nakayama^{*3}, Motoyasu Sagawa^{*4}, Hiroshi Saito^{*2}

^{*1} Department of Health Policy/Public Health, University of Tokyo

^{*2} Cancer Screening Research Division, Research Center for Cancer Screening and Prevention, National Cancer Center

^{*3} Department of Cancer Control and Statistics, Osaka Medical Center for Cancer and Cardiovascular Diseases

^{*4} Department of Thoracic Surgery, Kanazawa Medical University

Abstract

For cancer screening programs to effectively reduce cancer mortality, their quality must be controlled. The Japanese Ministry of Health, Labour, and Welfare has developed a "Checklist for the Evaluation of Screening Programs" and promulgated its use for this purpose. To achieve continuous improvement of quality of screening programs, continuous validating and updating the checklists themselves is important. In this study we evaluated use of this checklist for lung cancer screening programs using the RAND/UCLA appropriateness methods, a formal process of consensus among a panel of experts from different backgrounds. The process consisted of two rounds of ratings for appropriateness of each checklist item and a face-to-face meeting for discussion. The current checklist was rated in the first round of ratings and a draft revised checklist was created during the discussion and second ratings. All but one current checklist item was rated as appropriate, but the panelists proposed revisions for nearly half of the items during the process. This proposal will serve as a draft proposal for the official revision that will take place in a few years.

Key words: Lung cancer screening, Program evaluation, Quality control, Consensus panel, Health Services Research

J Thorac CT Screen 2010;17:106-114

連載

ヘルスサービスリサーチ(3)

「米国におけるヘルスサービスリサーチの教育課程」

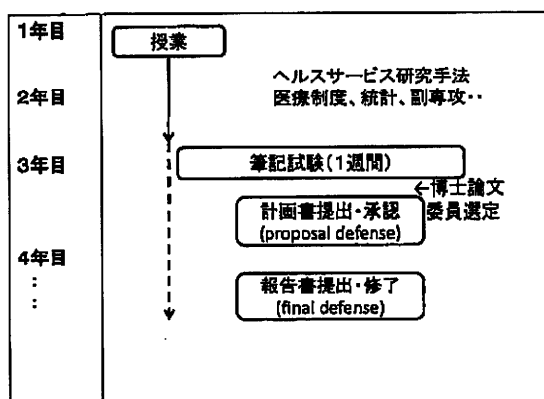
東京大学大学院医学系研究科健康医療政策学分野 東 尚弘

1. はじめに

ヘルスサービスリサーチの定義は時代とともに微調整が入り、米国医学研究所(Institute of Medicine)やヘルスサービスリサーチの学会である Academy Health (後述) による定義などが様々である¹⁻³⁾が、それらを総合すると臨床現場や健康政策の促進を目的に、医療の計画、配分、組織構造、提供過程や質、効果、効率、アウトカムを対象とした研究分野であり、資源配分、組織構成、財政構造に関する政策などに役立てることを目標としているといえる。この研究分野は、臨床医学、疫学、経済学、社会学、心理学、人口学、統計学などの技術を総合して遂行され、基本的に現実的な問題に対して研究結果から回答を見いだすことを目的とした応用研究センターの研究分野であるといえる(もちろん理論構築などの基礎研究分野も含んでいる)。ヘルスサービスリサーチは、米国に限定された研究分野ではないが、医療制度が非常に複雑で常に新しい試みがなされる(余地のある)米国で大きく発展している。日本ではあまり学問分野として「ヘルスサービスリサーチ」が意識されることが無かったが、医療に改革が必要であることが頻りに議論される時勢の中でその道しるべとして科学的な根拠をもたらすことが期待され、今後発展する分野であると言える。本稿においては筆者が2000年から2005年までカリフォルニア大学ロサンゼルス校(University of California at Los Angeles, 以下 UCLA)に留学した際に体験した当地における、ヘルスサービス博士課程(PhD in Health Services)のカリキュラムやその一端を紹介する。

なお、よく似た学位として DPH (Doctor of Public Health) があるが、UCLA のプログラムではどちらかというと専門職学位と考えられていた。ただし他大学の出身者の話を聞くとこの学位と内容の対応は大学院によって異なるようである。また、本稿では詳しく触れないが、修士課程については、専門職学位である MPH (Master of Public Health) の一系統としてヘルスサービス (Health Management

図1 UCLA ヘルスサービス博士課程



中心か Health Policy 中心にさらに分けられる), あるいは博士 (PhD) 課程へと続く研究職学位である MSHS (Master of Science in Health Services) のコースが設けられていた。取得する単位が MPH では幅広く公衆衛生をカバーするのに対して、MSHS では研究の組み立て方などに重点が置かれ、全く異なる課程となっていた。

2. UCLA における教育カリキュラム

カリキュラム全体の流れとしては、まずコースワークが2年有りそこで必須科目、および各分野系統の必須科目、選択科目をとらなければならない。その後、筆記総合試験 (Written comprehensive exam)、研究計画をたてると同時に博士論文審査委員を選び計画の承認 (Proposal Defense) を経て博士候補となり、計画通りの研究を遂行した後に最終審査を経て修了となる。以下に順に紹介する。

a) コースワークと筆記総合試験

コースワークは全体として、ヘルスサービスの専門レベルのコース、応用統計、研究計画法、研究(計画)発表会が必須であり、それに加えて各系統 (Cognate, 副専攻のようなもの) に分かれて必須科目が存在した。必須科目のヘルスサービスコース (Health Services 200) は基本的に米国の医療状況をみっちりたたき込まれた。ここでは、医療保険制

度に始まり、財政、特別な配慮が必要な集団（高齢者、児童、移民、ホームレス、低所得者など）の状況と健康行動の特徴、各国の医療制度、などが含まれる。講義も当然であるが宿題、課題が毎週のように出された。当然英語であり、筆者は初年度だったこともあり理解不十分なままたどたどしい英語でレポートを書くため、友人にみてもらったら、「ここは意味不明」、「ここは繰り返し」といって削られたあげく、返ってきた教官のコメントは「Too Short」という笑えないモノだったのをまだ覚えている。

研究計画法も必須科目であり、計画の各段階を辿って講義があった。内容は、科学的探求とは、リサーチクエスションとは、研究目的とは、セオリーとは、モデルとは、解析単位とは、変数とは、仮説とは、文献検索のやり方、デザインとは、サンプル法、妥当性、信頼性、倫理的配慮とは、などが含まれ10週間で概観する。毎週そのトピックに沿って自分の課題に応用して計画を書き進めなければならなかったのだが、じっくり一貫した計画など立てている暇がなく、まったく毎回ばらばらなことを書いてしまったのを覚えている。

応用統計はSASを使って線形回帰のやり方とそれらの前提、その検証の仕方などを習っていった。これは日本の高校の数学よりも数段やさしいので比較的楽だった。ただ、線形回帰などの方法がわかっていることだけではなく、その解釈を記述する練習をさせられたので、英語が苦勞だった。とくに「ベータ係数が有意など」という解釈は許されず、きっちりと「他の変数が同じの場合には、変数〇が1単位増加したら□がβ増加する」ということを言い回しから単位付きで書かされた。統計は有意差のための学問ではない、ということたたき込むのが目的だったのかもしれない。

それらと同時に副専攻のような形で系統 (Cognate) を選択する。系統の一覧を表1に示すが、筆

表1 UCLA 公衆衛生大学院におけるヘルスサービスの系統 (Cognate) 一覧

• 疫学	• Epidemiology
• 経済学	• Economics,
• 歴史	• History
• マネジメント	• Management
• 政策科学	• Political science
• 政策研究	• Policy studies
• 心理学	• Psychology
• 医療社会学	• Sociology
• 健康アウトカム	• Health care outcomes, research
• 薬剤経済学	• pharmaceutical economics

者は疫学系統を選択した。ヘルスサービスリサーチは多分野学際的であり基本的にヘルスサービスリサーチ者はジェネラリストであると考えられるが、その中でも副専攻的な研鑽を積む必要があるとの考えのようである。疫学は公衆衛生大学院の中にあるため、筆者は特に他の学部に出かけていく必要はなかったが、経済学系統を専攻した学生は経済学部へ授業を受けに行く必要があるし、心理学系統なら心理学部のコースをとる必要がある。驚いたことに医学部、経営大学院、法科大学院などの授業料の異なる学部 (School) は別として、コースは全ての大学院の学生に対して開かれており、全学で共通の単位登録システムを使っていたため、特別な許可や申請などなくともその学科の学生度同様に受講することができた。私も選択科目として、心理学部や教育学部、統計学部の統計関連のコースを受講した。

授業が終わった段階で、筆記総合試験 (Written comprehensive exam) が行われる。これは月曜から金曜までの5日間の内に5つ程度の課題について、持ち帰りでレポートを作成するというものである。解答を作成するにあたって人に尋ねてはならないが参考書やインターネットなど何をみても良いという規則になっていた。これは博士に値する能力とは知識を記憶していることではなく、どこを探せばどのような情報が入手できるかを的確に把握しており、それをまとめる能力の方が重要であるという思想に基づくものである。出題の題材としては「概念モデルとは何か」などの研究手法に関わるものから、「米国における処方薬の保険適用について」などの医療制度に関わるものまで幅広く出題された。ちなみに英語のNative checkをかける時間は無いので、特に文法的な間違いなどは意味が通る範囲で「理解する」ということであつた。

b) 研究計画と審査

筆記総合試験が終わると、次に博士論文の計画書 (Proposal) を書くことになる。この時点で博士論文審査委員会を構成する。構成員数は4人以上となっており、1人は自分の所属する部 (筆者の場合には Department of Health Services) 以外からメンバーを入れる必要がある。これで日程を決めて計画書審査会 (Proposal Defense) があり、ここで審査委員会により研究計画が博士論文として適切なものなのかを決定される。必要であれば (実際はほとんど必ず) 修正を要求されることになる。博士論文としての適切性は、一貫したテーマに関して3つの論文が発表できる程度大きさの研究であり現在の当該研究分野の発展に貢献するものかどうか、とされていた。(これは後に修正され、実際に3つの論文とし

てまとめても良い、ということに修正された。)

ここで研究計画が認められると、晴れて博士候補 (PhD Candidate) ということになる。この Proposal Defense が最も大変な過程であり、それが終わると後は計画書通りに遂行して結果を出せばよいということになる。その結果が出次第、最終論文を提出してその発表会を行う (Final Defense)。その結果計画書に記載されている解析がきちんと行われていることが確認されそれを正しく解釈していることが確認されると晴れて博士課程修了が認められるということになる。

この判定基準としては、結果として統計的な有意差などは全く関係ない。むしろ、計画を立てるにあたって、その課題が「結果として差があっても、なくてもどちらでも意義がある」研究が最も良いとされている。そのため、結果として差が無くても十分に考察ができる、ということになり、統計的有意差の有無が修了に影響することは無い。また、計画書は非常に重要であり、公式見解ではないものの現地で言われていたのは、計画書が審査委員と学生の契約書であり、計画書通りのことをきちんと遂行し結果を適切にまとめたのであれば、審査委員は修了を認めなければならないということであった。これは、研究を進めているうちに審査委員が異動になったりして入れ替わる、あるいは興味が変わってしまうことも考えられる。そうした時に審査委員の恣意的な意見によって学生に対して要求事項が増えて不当な負担となることを防ぐ意味合いがあるらしい。そのため、計画書がきちんとすれば、あとから「こんな解析もやってみたら？」などという意見が出て際限なく解析をしなければならなくなるということもない。

このような過程でなされる研究は、ヘルスサービスリサーチの専門誌に発表されるだけではなく、関連する臨床医学分野の雑誌に発表されることも多い。New England Journal of Medicine や JAMA などでも積極的にヘルスサービスリサーチが発表され、医療が基礎医学や臨床医学のみで構成されるわけではなく、社会的な要素が無視できないという認識の表れといえる。

3. 医師を対象としたプログラム (Clinical Scholars Program)

ヘルスサービスリサーチの教育は、何も公衆衛生大学院の中だけに限ったことではない。主なものにはロバートウッド・ジョンソン財団が主催する、「クリニカル・スカラーズ・プログラム」(Clinical Scholars Program) というものがあり、ここでは2

年以上の研修を終えた医師を対象として後期研修の一環として学ぶことができる。専門科は特に指定されておらず、幅広く医療に関する研究を志す医師のトレーニングの機会となっている。このプログラムは UCLA, ペンシルバニア大学, ミシガン大学, イェール大学が協力施設となっていて、それぞれ研究指導教官の指導の下で2年間の間に各自のテーマで研究をまとめることが課せられている。米国人に対する後期研修のため給料が支払われての研修となるが、筆者は UCLA のプログラムに無給ビジターとして参加した。また、他にも国立健康研究所 (National Institute of Health) が出資する, NRSA (National Research Service Award) フェローシップや、退役軍人部 (US Department of Veterans Affairs) の主催するフェローシップなどがあった。これらは、UCLA では合同で指導カリキュラムを作成していた。内容は非常に厳しい教育プログラムであり、1年目の夏から始まり、最初の3か月で、統計、政策、文章の書き方、論文抄読会、医療経済、法律の集中ゼミがありその後、実際の研究を行うことになる。また、頻繁に研究発表会があり個人の指導教員以外の教員からも研究の意義、方法、解釈などに関して徹底的な議論が行われる。ここでも何よりも、研究の意義が徹底的に議論され、「どのような結果が出ると思うのか」「そうであれば何が言えるのか」「逆の結果なら何が言えるのか」を、解析どころか、データを収集する前から吟味される。筆者は最初の会で、標準診療行為を記述してその実施率をもって医療の質とする「診療の質のプロセス指標 (QI)」の内容 (たとえば心筋梗塞患者が来院したらアスピリンを服用させる) に関して、一般医師が賛成するのかを調査したい、といったら「結果が賛成しない」だったら何が言えるのか、と問われ「賛成しないのだったらその原因を探りたい」と答えたら、「賛成しないのだったら教育が必要と考えられるが、すでにそれらの QI を使って診療の実態として標準が行われていないことがわかっているので、教育が必要なことに変わりはない。研究をする価値はあるのか」といわれてその場で却下されたのを覚えている。しかし、逆に研究そのものを吟味すると同時に、研究に対してきちんと意義を説明できるか、という能力を育てている側面もあり、アメリカらしい開拓精神かもしれない。

4. ヘルスサービスリサーチの学会 (AcademyHealth)

米国ではこのようなヘルスサービス分野に関する学会が存在する。もともと Association for Health

Services Researchとして1981年に始まり、2001年にAlpha Centerという、ヘルスサービスに関する解析、教育などの活動を行っている別の団体と合併してAcademy Healthという名称に変更になった。会員数は4,000人ほどの小規模な学会であるが、毎年6月頃に行われる学術大会では研究者だけではなく健康関連の政府関係者や米国外から関係者が参加し活発な議論が行われている。また、研究の方法論に関してもヘルスサービスリサーチで使われる最新の方法の教育セッションなども豊富である。詳しくは、<http://www.academyhealth.org>を参照されたい。

5. 結語

わが国においてヘルスサービスリサーチが全くなわけではない、しかし、現状では病院管理、看護研究、医療経済、経営などの分野に分かれておりあまり一貫した教育プログラムも無く独自に研究・教育を行っており、統合的な発展形とする努力が少な

かったように思われる。ヘルスサービスリサーチとしてこれらの研究活動や教育の連携を深め、政策提言や実務改善につなげるような研究を盛んにしていくことが、よりよい医療の構築に向けて公衆衛生研究者が果たしていくべき仕事の一つではないだろうか。

文 献

- 1) Academy of Health. Definition of Health Services Research. <http://www.academyhealth.org/About/content.cfm?ItemNumber=831&navItemNumber=514> (2010年6月24日アクセス可能)
- 2) Lohr KN, Steinwachs DM. Health services research: an evolving definition of the field. *Health Serv Res* 2002; 37: 7-9.
- 3) The Institute of Medicine. Health Services Research: Work Force and Educational Issues. Washington, D.C.: National Academy Press, 1995.