The Effectiveness of Cost Reduction with Charge Displays on Test Ordering under the Health Insurance System in Japan: A Study Using Paper-based Simulated Cases for Residents and Clinical Fellows

Yusuke Hirota, Shingo Suzuki, Yoshiyuki Ohira, Kiyoshi Shikino and Masatomi Ikusaka

Abstract:
Objective  To determine whether or not displaying the cost of tests can help reduce charges on test ordering in Japan.

Methods  This study was conducted under the setting of a simulated first visit of an outpatient for general internal medicine in a secondary medical institution in Japan. We randomly assigned 27 residents and clinical fellows to Team A or B. The first half, without charges displayed on the ordering system, was designated the “non-display group,” and the participants of Team A selected tests for each paper-based simulated case (Q1-Q14), while the participants of Team B selected tests for Q15-Q28. The second half, which had charges displayed, was designated the “display group,” and the participants of Team A selected tests for Q15-Q28, while the participants of Team B selected tests for Q1-Q14. The main outcome measure was the difference in the cost of tests per paper-based simulated case between the non-display and display groups.

Results  The median (interquartile range) cost of tests per paper-based simulated case was 12,255 yen (5,040-23,695 yen) in the non-display group versus 9,425 yen (2,320-21,700 yen) in the display group, showing a decrease of 2,830 yen with charges being displayed (p=0.002).

Conclusion  Displaying the charges when ordering tests in paper-based simulated cases resulted in cost reduction. The adoption of this intervention may reduce health insurance costs under the health insurance system in Japan, which has features such as universal health coverage and universal access to care.

Key words: charge displays, cost reduction, health insurance system, universal access to care, universal health coverage

Introduction

Control of healthcare costs is a pressing issue in Japan. The Organization for Economic Co-operation and Development (OECD) reported that expenditure on healthcare was equivalent to 10.2% of Japan’s GDP in the 2013 fiscal year (eighth largest proportion among OECD nations), having risen over time (1). Universal health coverage and universal access to care have created an environment in Japan in which patients can easily access medical facilities. However, this environment is responsible for frequent consultations, and although the number of doctors per 1,000 persons is relatively low in Japan, at 2.4 (ranking 29th), the average number of consultations per patient each year is 12.9 (ranking 2nd) (1). Therefore, to facilitate spending a shorter time on each examination, diagnoses are now made predominantly via tests. In conjunction with the “fees-for-service” payment model, this is a significant factor associated with the rise in healthcare costs (2). An example of this type of diagnosis is when a physician attempts to minimize the time consumed by face-to-face medical treatment and physical...
examination of a patient complaining of headaches by ordering a cranial computed tomography (CT) scan. This issue is underscored by the fact that there were 101.3 CT and 46.9 magnetic resonance imaging (MRI) scans performed per 1 million persons in Japan in the 2013 fiscal year (both the highest rates in the world) (3).

In various countries overseas, intervention studies (4) and education (5), audit and feedback (6), system-based (7), and incentive or penalty interventions (8) have been shown to be effective in reducing the excessive performance of tests. However, their usefulness in Japan has not been confirmed. Among these initiatives, the system-based intervention, which involves displaying the charges for each test as part of the ordering process (9), is straightforward to implement, and this approach has already been shown to be effective in reducing costs in the United States (9), France (10), Sweden (11), and South Africa (12). Considering interventions in the healthcare systems of the countries involved, the United States and South Africa lack universal health coverage (13, 14); therefore, because the amount paid for the medical treatment by the patient increases depending on the content of the medical treatment they receive and the nature of their insurance scheme (15), people’s awareness of the associated costs can be expected to be higher in those countries than in Japan. In France and Sweden, where there is universal health coverage, it is possible to receive medical treatment at a low cost; however, limits are placed on access to expensive treatments through a system of personal doctors (primary care physicians) (16, 17). Alternatively, due to the universal access to care in Japan, patients can consult doctors at leading medical institutions at their own discretion and receive high-level and expensive tests and treatments, covering only 10-30% of the costs by themselves. Japanese patients’ awareness of costs is thus expected to be lower than in other countries.

Furthermore, the fact that there is an economic benefit for physicians in carrying out tests (Japanese outpatient consultations are, by principle, paid on a fee-for-service basis) may lead some doctors to carry out expensive tests. However, Bates et al. suggested that the finding of no cost-reduction impact even when the charges for tests were displayed was a result of the nonchalance of the participating resident physicians regarding the costs of the tests and because they stood to make no economic gain or loss, in contrast to permanent physicians (18). It is therefore likely that resident physicians, who are unaffected by the financial incentives of the fee-for-service payment model, do not order tests for the income and instead order the necessary tests regardless of the cost.

Based on the above, we conducted the present study to assess the effect of displaying charges as part of the test-ordering process on cost reduction among Japanese resident physicians with a poor awareness of the tests’ costs (as universal access to care is guaranteed in Japan and there is universal health coverage) with a relatively small burden for patients at the point of use. If displaying test charges can promote behavior modification toward a reduction in unreasonable tests, this would help create an easily implemented strategy for inhibiting increases in test costs.

Materials and Methods

Participants

There were 28 candidates in this study, comprising 10 first-year and 10 second-year resident physicians who were training at the Department of General Medicine, Chiba University Hospital (hereafter, the Department), in fiscal year 2015 along with 8 clinical fellows at the Department. We obtained written informed consent from all of the participants. However, one first-year resident was excluded due to the fact that he urgently had to treat an emergency case. As a result, 27 candidates were ultimately enrolled as participants in this study.

The study was approved by the research ethics committees of the Graduate School of Medicine, Chiba University.

Setting

The study was carried out in April 2015 as an intervention study. The participants were told that this study was an investigation of simulated cases, involving delivering the most-suspected diagnosis and ordering clinical examinations just as they would in routine medical practice.

Assuming the setting of initial outpatient consultations at the department for general internal medicine at a secondary-care hospital in Japan, the participants were asked to consider 28 cases (Supplementary material 1) created with due regard for case mix. For each case, the age, gender, medical history, and presentation of symptoms during the medical examination of a putative patient were listed, and the participants were asked to describe the suspected disease and select from a list on a test-ordering sheet any number of blood, urine, physiological, or radiological tests or cultures (Supplementary material 2) that they would have ordered in routine medical care. There were two versions of the test-ordering sheet (Supplementary material 3): one that included the charge for each test and one that did not. Test charges were based on the medical fee remuneration points for fiscal year 2014 in Japan and were shown in yen. The time limit for answering the questions was set at 60 minutes, allowing 30 minutes for each half of the activity, which was found to be reasonable, considering the 2014 pilot study (unpublished), which examined residents on rotation at the Department.

Procedure

Figure shows the outline of this study. Participants at each level of training were randomly assigned to one of two teams: Team A (14 participants; 5 first-year residents, 5 second-year residents, and 4 clinical fellows) or Team B (13 participants; 4 first-year residents, 5 second-year residents, and 4 clinical fellows) (Table 1). In the first half of the trial,

Intern Med 58: 187-193, 2019 DOI: 10.2169/internalmedicine.0738-17

188
Outcomes of the study. Twenty-seven respondents were randomized to Team A (14 respondents), which started with question (Q) 1-Q14, or Team B (13 respondents), which started with Q15-Q28. The respondents initially solved model cases without charges displayed [n=378 examinations, 14 respondents ×14 questions (Q1-Q14) +13 respondents ×14 questions (Q15-Q28)] and then solved model cases with charges displayed [n=378 examinations, 14 respondents ×14 questions (Q15-Q28) +13 respondents ×14 questions (Q1-Q14)].

Table 1. Respondents’ Characteristics.

| Characteristics          | Survey Respondents, n (%) |
|--------------------------|----------------------------|
|                          | Team A                     | Team B                     |
| Total                    | 14 (52)                    | 13 (48)                    |
| Level of training        |                            |                            |
| First-year residents     | 5 (19)                     | 4 (15)                     |
| Second-year residents    | 5 (19)                     | 5 (19)                     |
| Clinical fellows         | 4 (15)                     | 4 (15)                     |
| Sex                      |                            |                            |
| Male                     | 9 (33)                     | 8 (30)                     |
| Age, year                |                            |                            |
| 25-30                    | 13 (48)                    | 11 (41)                    |
| 31-40                    | 1 (3.7)                    | 0 (0)                      |
| >40                      | 0 (0)                      | 2 (7.4)                    |
| Clinical experience, year|                            |                            |
| <1                       | 5 (19)                     | 4 (15)                     |
| <2                       | 5 (19)                     | 5 (19)                     |
| <6                       | 4 (15)                     | 3 (11)                     |
| >6                       | 0 (0)                      | 1 (3.7)                    |

Participants at each level of training were randomly assigned to one of two teams: Team A or Team B.

The following points were incorporated into question creation and into decisions about the method of answering: First, to prevent differences in the potential diagnosis arrived at based on the provided clinical information from affecting the tests participants ordered, the clinical information for each question comprised the typical medical history and presentation of symptoms during the medical examination for the disease concerned. Second, to avoid participants guessing the intention of the study and thereby affecting the study results and to avoid carry-over impact after participants became aware of the test charges, the decision was made not to use a cross-over method (where sheets with and without the test charges were used by two different groups at the same time). Instead, we administered the first half of the activity without charges displayed and the second half of the activity with charges displayed. Third, to adjust for the influence of any differences in the simulated cases on the questions or respondents, the participants were divided randomly into two teams, and in the first half (in which they were all answering as the “non-display group”), one team
was asked to answer Q1-Q14 and the other Q15-Q28, with each team being asked to answer the remaining questions in the second half (in which they were all answering as the “display group”).

**Outcome and measurements**

The primary clinical outcome was taken to be the difference in the total costs of tests per case between the 2 groups of data; namely, the non-display group data (the answers in the first half of the activity; n=378 “examinations”) and the display group data (the answers in the second half of the activity; n=378 “examinations”). In addition, the following differences between the two groups were established as secondary clinical outcomes: the difference in the total costs of tests per case from Q1-Q14 and Q15-Q28, the difference in the costs of non-radiological tests per case, the difference in the total number of tests per case, the difference in the proportion of “examinations” with the correct potential diagnosis, and the difference in the proportion of “examinations” with selected tests essential for the diagnosis. Please note that “tests essential for the diagnosis” for the purposes of this study were decided mainly on the basis of Japanese practice guidelines as determined by a focus group discussion among the faculty members of the Department. Furthermore, after the tests in this study were performed, we gave the participants a scale-based questionnaire to describe the study’s usefulness with free comments allowed [from 1 point (not useful at all) to 5 points (very useful)] subsequent to an explanation about the purpose of this study and commentary on the cases.

**Statistical analyses**

The Mann-Whitney U test was used to assess the differences between the two groups of data in the total costs of tests per case, the difference in radiological test costs per case, the difference in the costs of non-radiological tests per case, the difference in the total number of tests per case, the difference in the proportion of “examinations” with the correct potential diagnosis, and the difference in the proportion of “examinations” with selected tests essential for the diagnosis. Please note that “tests essential for the diagnosis” for the purposes of this study were decided mainly on the basis of Japanese practice guidelines as determined by a focus group discussion among the faculty members of the Department. Furthermore, after the tests in this study were performed, we gave the participants a scale-based questionnaire to describe the study’s usefulness with free comments allowed [from 1 point (not useful at all) to 5 points (very useful)] subsequent to an explanation about the purpose of this study and commentary on the cases.

Based on the results of the aforementioned pilot study, it was considered that, with the α error estimated as 0.05 and β error at 0.2 (with the power of detection at 0.8), the minimum sample size necessary to compare the difference between the non-display and display group data in test costs for the 28 cases was 13 participants; therefore, at least 14 participants were recruited for this study.

**Results**

The background characteristics of the 27 participants in this study are shown in Table 1. The median value (interquartile range) of the total costs of tests per case was 12,255 yen (5,040-23,695 yen) for the non-display group and 9,425 yen (2,320-21,700 yen) for the display group, showing a difference of 2,830 yen, which was significantly lower for the display group (p=0.002) (Table 2). The median value of the total costs of tests per case from Q1-Q14 was 15,700 yen (5,905-24,460 yen) for the non-display group and 13,145 yen (2,000-21,700 yen) for the display group, showing a difference of 2,555 yen, which was significantly lower for the display group (p=0.03). The median value of the total costs of tests per case from Q15-Q28 was 9,160 yen (4,500-22,700 yen) for the non-display group and 6,555 yen (2,320-20,365 yen) for the display group, showing a difference of 2,605 yen, which was significantly lower for the display group (p=0.03) (Table 2). The median value of the costs of radiological tests per case was 1,420 yen (0-15,861 yen) for the non-display group and 1,420 yen (0-15,700 yen) for the display group, showing a difference of 0 yen, which was not significant (p=0.72) (Table 2). The median value of the costs of non-radiological tests per case was 7,050 yen (1,300-12,715 yen) for the non-display group and 4,685 yen (0-9,904 yen) for the display group, showing a difference of 2,365 yen, which was significantly lower for the display group (p<0.001) (Table 2). The median total number of tests per case was 7 (2-19) for the non-display group and 3 (1-13) for the display group, with the value for the display group being significantly lower (p<0.001) (Table 2).

The proportion of “examinations” with the correct potential diagnosis was 98.4% (372/378) for the non-display group and 97.6% (369/378) for the display group, showing no significant difference between the 2 groups (p=0.43) (Table 3). The proportion of “examinations” with selected tests essential for the diagnosis was 95.5% (361/378) for the non-display group and 96.0% (363/378) for the display group, again showing no significant difference between the 2 groups (p=0.72) (Table 3).

The mean scores on the questionnaire about the usefulness of this tests were 4.1 points among first-year residents, 4.4 points among second-year residents, and 4.1 points among clinical fellows (Supplementary material 4), with comments such as “It was a good opportunity to learn about test costs” and “In the future, I will consider those costs when ordering tests.”

**Discussion**

The results of this study conducted in first- and second-year residents and clinical fellows showed that, in simulated initial outpatient consultations at a department for general internal medicine at a secondary-care hospital, test costs were reduced by displaying the charges for each test as part of the ordering process. There were no marked differences between the two groups in terms of the diseases that the clinical information prompted the participants to diagnose, with the same results obtained between Q1-Q14 and Q15-Q28, which denies any confounding of the participants’ abil-
In an intervention study of second- and third-year residents and clinical faculty at family practice centers in the US, when participants were asked to choose the necessary tests for the case studies used, as in our study, the cumulative cost of tests ordered per patient by the “price-information group” was lower than that by the “control group.” Furthermore, no marked differences were seen between the two groups in terms of the ratio of minimum necessary test selection (19). However, the US family practice center study (19) did not consider differences between the two groups in terms of the correct potential diagnosis or the difference in the abilities of the participants; therefore, the existence of bias in those regards cannot be ruled out.

Table 2. Costs and the Number of Tests Per Simulated Case.

| Group                  | Median value of total cost of tests, yen (Interquartile Range) | Difference*, yen | p value† |
|------------------------|-----------------------------------------------------------------|-----------------|----------|
| Non-display group (n=378) | 12,255 (5,040-23,695)                                           | -               | -        |
| Display group (n=378)  | 9,425 (2,320-21,700)                                            | 2,830           | 0.002    |
| Q1-Q14                  |                                                                 |                 |          |
| Non-display group (n=196) | 15,700 (5,905-24,460)                                           | -               | -        |
| Display group (n=192)  | 13,145 (2,000-21,700)                                           | 2,555           | 0.03     |
| Q15-Q28                 |                                                                 |                 |          |
| Non-display group (n=196) | 9,160 (4,500-22,700)                                           | -               | -        |
| Display group (n=192)  | 6,555 (2,320-20,365)                                            | 2,605           | 0.03     |

Table 3. The Proportion of “examinations” with the Correct Potential Diagnosis and Selected Tests Essential for Diagnosis.

| Group                  | “Examinations” with the correct potential diagnosis, n (%) | p value* |
|------------------------|----------------------------------------------------------|----------|
| Non-display group (n=378) | 372 (98.4)                                              | -        |
| Display group (n=378)  | 369 (97.6)                                              | 0.43     |

| Group                  | “Examinations” with selected tests essential for diagnosis, n (%) | p value* |
|------------------------|-----------------------------------------------------------------|----------|
| Non-display group (n=378) | 361 (95.5)                                              | -        |
| Display group (n=378)  | 363 (96.0)                                              | 0.72     |

Q: question
* The value of the non-display group median less the display group median.
† Using the Mann-Whitney test

Intern Med 58: 187-193, 2019 DOI: 10.2169/internalmedicine.0738-17
ential for the diagnosis or in terms of the correct potential diagnosis, meaning that the same evaluations were made based on the clinical information provided. This makes it clear that displaying charges as part of the test-ordering process is effective in reducing costs. We can therefore conclude that informing physicians about charges reduces the excessive conducting of tests without resulting in the omission of tests essential to the diagnosis.

In a previous cross-sectional study, when asked to estimate the charges for 15 commonly ordered diagnostic tests, participating internal medicine residents and faculty members at an academic tertiary-care hospital showed poor knowledge of healthcare charges (20). Situations such as these have shown the effectiveness of educational inventions in reducing costs (5). However, as no mention has been made of who will perform this education or when (21), the burden on medical institutions from this kind of intervention is expected to be substantial, and its realization will not be easy. In the present study, as in other past reports (22), it was confirmed that the behavior of resident physicians regarding cost reduction can be modified by merely listing healthcare charges. Furthermore, the mean score on the questionnaire regarding the usefulness of this study was high in each grade, with positive comments for learning costs, indicating that displaying charges on test ordering would be helpful for physicians. This intervention is simple and should also have an educational effect, reminding physicians about test costs each time they order a test. Thus, this type of intervention is likely to be highly practical in the reduction of test costs.

The charge display did not help to reduce the costs of radiological tests in comparison to the non-charge display. Another study which investigated the costs of radiological tests for inpatients at a tertiary-care hospital in the United States also observed no cost reduction (23). The authors of that study suggested there would be less additive value to in-patients in a tertiary-care hospital, and the necessary tests were conducted for inpatients in a tertiary-care hospital, which are seen as necessary to examine the situation based on the clinical information provided. This makes it clear that displaying charges as part of the test-ordering process is effective in reducing costs in actual clinical practice. Looking ahead, it would be beneficial to verify whether or not displaying charges as part of the test-ordering process is effective in reducing costs in actual clinical practice in Japan.

The 7th Annual Conference of Japan Primary Care Association in Asakusa, Tokyo, Japan, June 11, 2016 (The Hinohara awardee presentation).

The authors state that they have no Conflict of Interest (COI).

Conclusion

This study showed that displaying the charges for tests as part of the test-ordering process in paper-based simulated cases caused even resident physicians in Japan (who have a low awareness of costs because they work within a healthcare system with universal health coverage and universal access to care) to modify their behavior and increased their awareness of medical costs when making a diagnosis. This finding suggests that this is a potential strategy for controlling medical costs. Looking ahead, it would be beneficial to verify whether or not displaying charges as part of the test-ordering process is effective in reducing costs in actual clinical practice in Japan.

References

1. OECD Health Statistics 2015. [Internet]. [cited 2017 Feb. 26]. Available from: http://www.oecd.org
2. Jones RS. Health-Care Reform in Japan: Controlling Costs, Improving Quality and Ensuring Equity. OECD Publishing, Paris, 2009.
3. Health at a Glance 2015: OECD Indicators. OECD Publishing, Paris, 2015.
4. Kobewka DM, Ronksley PE, McKay JA, Forster AJ, van Walraven C. Influence of educational, audit and feedback, system based, and incentive and penalty interventions to reduce laboratory test utilization: a systematic review. Clin Chem Lab Med 53: 157-183,
5. Davidoff F, Goodspeed R, Clive J. Changing test ordering behavior: A randomized controlled trial comparing probabilistic reasoning with cost-containment education. Med Care 27: 45-58, 1989.

6. Stafford RS. Feedback intervention to reduce routine electrocardiogram use in primary care. Am Heart J 145: 979-985, 2003.

7. Tierney WM, Miller ME, McDonald CJ. Effect of cost-containment education on test ordering behavior. A randomized controlled trial comparing probabilistic reasoning with cost-containment education. Med Care 27: 45-58, 1989.

8. Martin AR, Wolf MA, Thibodeau LA, Dzau V, Braunwald E. A trial of two strategies to modify the test-ordering behavior of medical residents. N Engl J Med 303: 1330-1336, 1980.

9. Goetz C, Rotman SR, Hartoularos G, Bishop TF. The effect of charge display on cost of care and physician practice behaviors: a systematic review. J Gen Intern Med 20: 835-842, 2015.

10. Seguin P, Bleichner JP, Grolier J, Guillou YM, Mallédant Y. Effects of price information on test ordering in an intensive care unit. Intensive Care Med 28: 332-335, 2002.

11. Schilling UM. Cutting costs: the impact of price lists on the cost development at the emergency department. Eur J Emerg Med 17: 337-339, 2010.

12. Ellemendin S, Rheeder P, Soma P. Providing clinicians with information on laboratory test costs leads to reduction in hospital expenditure. S Afr Med J 101: 746-748, 2011.

13. Barnet JC, Vornovitsky M. Health Insurance Coverage in the United States: 2015. Government Printing Office, Washington, D. C., 2016.

14. Kumar A, Lagasnerie GD, Maieron F, Forti A. Pricing and competition in Specialist Medical Services: An Overview for South Africa. OECD Publishing, Paris, 2014.

15. Baker L, Bundorf MK, Royalty A. Private insurers’ payments for routine physician office visits vary substantially across the United States. Health Affairs 32: 1583-1590, 2013.

16. Chevreul K, Berg Brigham K, Durand-Zaleski I, Hernandez-Quevedo C. France: health system review. Health Syst Transit 17: 1-218, 2015.

17. Anell A, Glennärd AH, Merkur S. Sweden health system review. Health Syst Transit 14: 1-159, 2012.

18. Bates DW, Kuperman GJ, Jha A, et al. Does the computerized display of charges affect inpatient ancillary test utilization? Arch Intern Med 157: 2501-2508, 1997.

19. Cummings KM, Frisof KB, Long MJ, Hryniewicz G. The effects of price information on physicians’ test-ordering behavior. Ordering of diagnostic test. Med Care 20: 293-301, 1982.

20. Tek Sehgal R, Gorman P. Internal medicine physicians’ knowledge of health care charges. J Med Educ 3: 182-187, 2011.

21. Eisenberg JM. An educational program to modify laboratory use by house staff. J Med Educ 52: 578-581, 1977.

22. Feldman LS, Shihab HM, Thieman D, et al. Impact of providing fee data on laboratory test ordering: a controlled clinical trial. JAMA Intern Med 173: 903-908, 2013.

23. Durand DJ, Feldman LS, Lewin JS, Brotman DJ. Provider cost transparency alone has no impact on inpatient imaging utilization. J Am Coll Radiol 10: 108-113, 2013.

24. Schaefer GR, Matus HL, Goetz C, Arora VM. Magnetic resonance imaging monsters and surgical vampires. Arch Surg 146: 1333-1334, 2011.

25. The Japanese Association for Infectious Diseases. [Internet]. [cited 2017 Feb. 9]. Available from: http://www.kansensho.or.jp (in Japanese).

The Internal Medicine is an Open Access journal distributed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view the details of this license, please visit (https://creativecommons.org/licenses/by-nc-nd/4.0/).