Discussion Paper No. 547

COMPARISON BETWEEN TREATMENT OUTCOMES IN ISCHAEMIC HEART DISEASE USING SURVEYS OF MEDICAL CARE IN JAPAN

Rei Watanabe
and
Yasushi Ohkusa

July 2001

The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

29 June, 2001

Comparison between Treatment Outcomes in Ischaemic Heart Disease Using Surveys of Medical Care in Japan *)

Rei Watanabe

Graduate School of Economics, Kyoto University

and

Yasushi Ohkusa

Institute of Social and Economic Research, Osaka University

Correspondence:Yasushi Ohkusa, Institute of Social and Economic Research, Osaka University, 6-1 Mihogaoka Ibaraki Osaka, Japan

Tel:81-6-6879-8566 Fax:81-6-6878-2766

e-mail:ohkusa@iser.osaka-u.ac.jp

Summary

Objective The treatments of ischaemic heart disease have drastically changed due to the developments in coronary-anterial bypass graft operations and coronary intervention procedures. The treatment outcomes of these procedures are compared using observational data.

Method The data used are microdata from Surveys of Medical Care (SMC) in Japan from 1992 to 1998, which are claim data, not using a randomized controlled trial. Hence, endogeneity in the choice of treatment should be important and thus both the instrumental variables method and the nonparametric propensity score matching method are adopted for estimation.

Result Although the instrumental variables method and nonparametric propensity score matching method produce different results in many cases, predominance of the latter over the instrumental method is confirmed. Although claimed with some reservations, patients treated with precutaneous transluminal coronary angioplasty (PTCA) have significantly lower possibility of severe heart failure than those treated with coronary artery by-pass grafting (CABG). The results of comparing invasive procedures with preserving treatment depend on the existence of constraints. In data for more acute situations, invasive procedures have a significantly higher possibility of severe heart failure than preserving treatment.

Keywords: Ischaemic heart disease, Coronary intervention, Percutaneous transluminal coronary angioplasty, Coronary aorta bypass grafting, Instrumental variables method, Nonparametric propensity score matching method

1 Introduction

Heart disease is the second most common cause of death in Japan, and more than half these deaths are caused by ischaemic heart disease. The risk factors include age, hyperlipidaemia, smoking, hypertension, and the importance of ischaemic heart disease is naturally rising in Japan.

Before the 1970s, the main treatments for ischaemic heath disease were conservative, such as rest and medication. In the 1970s, treatment changed drastically with the development of coronary artery bypass grafting (CABG) and coronary intervention using coronary catheterization. These new treatments have decreased mortality from ischaemic heart disease, but they are invasive and expensive. Indications for these treatments have to be considered carefully. Recently, the practice of evidence-based medicine (EBM) has been established, and many randomized controlled studies are being performed for collecting evidence of the outcomes of each treatment.

Some clinical studies of ischaemic heart disease have results contradicting other studies. For example, some show evidence that percutaneous transluminal coronary angioplasty (PTCA) has better outcomes than percutaneous transluminal coronary recanalization for patients with acute myocardial infarction[1)], but others provide opposite evidence[2)]. By integrating many clinical studies, the American Cardiology Association and American Heart Association have produced guidelines for the management of patients with acute myocardial infarction. In these guidelines, they divide various diagnostic procedures and treatments into three classes, of which class 1 shows positive evidence for its use and effect, class 2 shows conflicting evidence, and class 3 shows negative evidence. These guidelines propose recommendations and help clinicians' decision-making[3)]. Although these guidelines are based on many clinical studies examining cost-effectiveness as well as short-term outcomes and long-term prognoses, such a randomized trial has some serious problems, such as cost in both money and time, generalizing from a specific kind of patient to other patients, and

ethical problems arising when patients might be randomized out of a treatment suspected to be effective[4]. Instead of a randomized control study, one study uses observational data from patients' medical records and investigates the incremental effects of intervention therapy for acute myocardial infarction by using the distance between their houses and hospitals as valid instrumental variables[5].

Some researchers point out that the effects of treatments with high fixed and marginal costs, such as operations and coronary interventions, are over-estimated. Treatments for ischaemic heart disease include low technical innovations with low fixed and marginal costs, such as aspirin or β-blockers, as well as the use of advanced technical innovations. One study concludes that both low and high technical interventions contributed significantly to the improvements of ischaemic heart disease[6], and another found that low technical innovations probably accounted for the bulk of the heart attack mortality decline between 1975 and 1995[7].

Unfortunately, in Japan, no large randomized controlled studies have been carried out, and no treatment guidelines based on many clinical studies have been established. Clinicians usually make decisions on diagnostic procedures and treatments with the aim of improving in-hospital prognoses, and make decisions without consideration of costs. Needless to say, clinical decision-making must not only be grounded on evidence from published articles, but also on each clinician's experience and each patient's conditions. Considering the differences of outcomes of surgery or intervention therapy for ischaemic heart disease among medical providers and the need for efficient use of medical resources, large clinical studies are needed in Japan.

With this background, this paper examines outcomes of the treatment for ischaemic heart disease using observational microdata of the SMC gathered from claim data.

2 Data

The data are from Surveys of Medical Care (SMC), which are conducted every year in June by the Ministry of Health and Welfare. The surveys are stratified samples of approximately 200,000 claims data from the government-managed health insurance and national health insurance scheme, written by medical providers all over Japan. Although it is the most comprehensive survey of medical care in Japan, there are several problems with the SMC for research.

First, the survey is only conducted in June and does not cover all patients. Because there are not enough samples, the data are pooled for several years.

The second problem is the coding used for each disease and drug. With the disease classification, ischaemic heart disease consists of one group, but this corresponds to I20–I25 in ICD-10. This includes I20, Angina pectoris; I21, Acute myocardial infarction; I22, Subsequent myocardial infarction; I24, Other acute ischaemic heart diseases; and I25, Chronic ischaemic heart disease. From the SMC, acute myocardial infarction cannot be distinguished from chronic ischaemic heart diseases such as stable angina. With the coding of drugs, the SMC shows only a four-digit classification combining drug code and drug price determined by the government. However, many generic drugs have the same prices, and thus it is impossible to isolate them.

The third problem is related to the first one. The SMC uses a sample from claim data during one month, and there is no information about previous months and the following months, even if th sampled patients were in hospital dring the other months. From the Patient Survey of the Ministry of Health and Welfare, the duration of hospitalization of the ischaemic heart disease patient is 31.3 days on average. However, many patients have hospitalization across months.

The fourth problem is that the SMC is based on claim data, not on patients' clinical cards. In the clinical setting, physicians and other medical providers record everything

on the patient's clinical card. Using clinical cards, clerical workers make claim data for charging fees to insurers once a month. In claim data, clinical procedures are recorded comprehensively, not chronologically. In other words, there is no information about the date of the specified procedures, even though the information about drugs, tests and treatments performed are almost perfect. Furthermore, the SMC does not indicate death, discharge or continuation of therapy.

Owing to these problems, the SMC is not suited for analysis of the long-term prognosis and medical expenditure of hospitalization periods longer than one month. Hence, it is important to specify proxy variables that represent short-term outcomes and analyse the effects of operations and coronary interventions for ischaemic heart disease. The compared operations are CABG and coronary interventions such as PTCA. Additionally, preserving treatments with invasive procedures for the patients of ischaemic heart disease with emergent and intensive care are compared.

3 Comparisons between outcome of CABG with coronary interventions

First, CABG and coronary interventions such as PTCA are compared. The samples are the patients for which the diagnosis is ischaemic heart disease and coronary angiography is performed. Ischaemic heart disease is caused by stenosis or occlusion of the coronary artery supporting the heart muscle. Coronary angiography shows the pathological regions of the coronary artery, wall motion of the heart and haemodynamics of the heart in detail, and is always done before CABG and PTCA. By sampling the patients with coronary angiography from all the patients, samples in which patients' condition are so stationary that angiography is not needed, or so severe that invasive procedures such as angiography are impossible, are excluded. Note that, as for the second problem mentioned above, many cases are chronic, not acute. To exclude chronic cases, the data are limited to patients with

hospital stay less than one month, and with blood type tests in the corresponding month. Since blood type tests are always done at the first admission and not at readmission or re-operation, this indicates the first procedure in the hospital. By adding each constraint, the sample can be limited to patients in which the period from admission to corresponding treatment is short, and chronic cases can be excluded.

In this analysis, data include heterogeneous patients, and treatments are selected endogenously, not exogenously, as in a randomized control study. Corresponding to this endogeneity, we provide an account in 'the model' section.

Although PTCA corresponds to the usual balloon angioplasty (POBA) in the SMC, there are new devices such as the directional coronary atherectomy (DCA) and the transluminal extraction catheter (TEC). However, PTCA here is broadly defined as usual balloon angioplasty and other various new devices, because the data cannot provide information to distinguish them.

Finally, the adopted outcome measures are explained. The most important factor determining the short-term prognosis is pump function of the heart. Killip's classification[8] is the most classical, but Forrester's classification is broadly used in deciding treatment, in relation to patients' heart function[9]. In this classification, catecolamines are recommended for treatment of severe heart failure. Among catecolamines, dopamine and dobutamine are used most widely for heart failure, and norepinephrine and epinephrine produce such a strong constriction of peripheral vessels that they are not used as the first choice drug, but used for cases with the most severe malfunction of the heart. In particular, epinephrine is usually used for cardiac arrest. Accordingly, using of norepinephrine or epinephrine indicates that a patient has severe heart failure. As norepinephrine and epinephrine have the same code and price, there is no way to distinguish them.

From the fourth problem, the patients using these drugs include the ones before and after treatment. Thus, the patients in which heart failure has been very severe before treatment cannot be excluded, but uncompensated heart failure is one of the contra-indications for

coronary angiography[10]. We assume that a few patients were using these drugs for advanced coronary angiography.

To sum up, the use of norepinephrine or epinephrine is a proxy variable for severe heart failure, and we compare the effect of the PTCA with CABG on severe heart failure.

4 Comparison between outcome of preserving procedures with invasive procedures

The sample used cases with emergent or intensive treatments, and we compared invasive procedures with preserving treatments. To do this, the following procedure was adopted. First, the sampled cases were refined so that the hospitalization fee was for the emergency ward or intensive care unit, so as to limit acute cases and chronic patients managed to just after operations such as CABG. Moreover, to exclude chronic cases, hospitalization less than one month and blood type tests in this are required. The outcome measure for severe heart failure is the same.

5 Estimation model

When using observational data, it is extremely important that selections of treatment are determined not exogenously, as in a randomized controlled study, but endogenously. To treat this problem of endogeneity, preceding investigations have used the distance between patients' houses and hospitals as an instrumental variable in evaluating the effect of cardiac catheterization from the 180,000 patients' records of acute myocardial infarction in the US[11].

On the other hand, the problem of endogeneity is often raised in econometrics. Although the instrumental variables method is used to solve this, the nonparametric propensity score matching method[12),13)] is widely employed and recognized for its advantages[14)]. It has now begun to be used in the analysis of treatment effects in the field of health economics[15)]. In

this paper, both the nonparametric propensity score matching method and instrumental variables methods are employed and compared for evaluation of treatment effects.

The nonparametric propensity score matching method is explained as follows. Consider patients with certain treatments or procedures i and j. Information on respective patients is denoted by X_i and X_j, treatment outcome of respective patients by Y_i and Y_j . The first stage in estimation is probit estimation with regard to treatment selection. The estimated participation probabilities of each types are denoted as $P(X_i)$ and $P(X_j)$. Then, the treatment effect is evaluated as:

$$\frac{\sum_i \{Y_i - \frac{\sum_j Y_j K[(P(X_i)-P(X_j))/h]}{\sum_j K[(P(X_i)-P(X_j))/h]}\}}{N} \tag{1}$$

where $K[\cdot]$ denotes the kernel estimator, and the Epanechnikov, biweight, triangular, Gaussian, Rectangular procedures are used[16]. h is set $0.9\sigma_{P(X_j)}N^{-0.2}$, which is regarded as the optimal value, and $\sigma_{P(X_j)}$ denotes the standard deviation of $P(X_j)$. N denotes sample size. This procedure is called kernel matching.

However some investigators recommend local linear matching rather than kernel matching with respect to efficiency[17]. This method is explained as:

$$\frac{\sum_i \{Y_i - \frac{\sum_j Y_j \{K_{ij} \sum_k K_{ik}(P(X_i)-P(X_k))^2 - K_{ij}(P(X_i)-P(X_j)) \sum_k K_{ik}(P(X_i)-P(X_j))\}}{\sum_j K_{ij} \sum_k K_{ik}(P(X_i)-P(X_k))^2 - (\sum_k K_{ik}(P(X_i)-P(X_k))^2}\}}{N} \tag{2}$$

where $K_{ij} = K[(P(X_i) - P(X_j))/h]$ and k denotes the summation of patients undertaking a certain treatment. Statistics estimated using equation (1) or (2) are called generically average treatment effects.

Further, the distributions of the average treatment effects are not specified. Accordingly, their confidence intervals are calculated using bootstrapping. Concretely, replication is done 1,000 times and bias correction is used[16].

5.1 Estimation model for comparison of PTCA and CABG

The explained variable is dichotomous, being 1 when either of two drugs is used and 0 when not used. Explanatory variables are patients' characteristics such as age and sex, providers' information such as number of beds in the hospital, type of hospital (non-profit, profit and public) and whether teaching hospital, and dichotomous treatment variables, being 1 when PTCA is performed and 0 when CABG is performed.

5.2 Estimation model for comparison between invasive procedures and preserving procedures

Almost all of the estimation equations are the same as in the previous subsection except for the treatment variables. The dichotomous variable is 1 when invasive procedures (PTCA or CABG or coronary angiography) are performed, and 0 otherwise.

6 Estimation result for comparison between PTCA and CABG

Table 1 shows the average of each sample, and Tables 2–5 indicate the estimation results from the whole sample. Tables 6–9 show results from data of less than one month hospitalization. As there were insufficient observations for the sample with blood type test, we omitted that estimation. Each sample produced results of the probit estimation, regressing outcome on treatment (PTCA or CABG) and other variables (Tables 2, 6), probit estimation for treatment selection (Tables 3, 7), probit estimation for outcome using estimated probability of treatment selection as instrumental variables (Tables 4, 8), and average treatment effect constructed by the nonparametric propensity score matching method using estimated probability of PTCA choice as P.

In simple probit estimation without consideration of endogeneity (Table 2), age and for-profit hospitals are significantly positive, and PTCA, age-squared and teaching hospital are significantly negative. The probability of severe heart failure increases until about 75

years of age, and then decreases. For-profit hospitals have higher probability of severe heart failure for their patients than public hospitals, and teaching hospitals have a lower probability than other hospitals. Most importantly, PTCA has a significantly lower probability of severe heart failure than CABG, by an amount of 35 percentage points. This means that PTCA reduces the probability of severe heart failure to 15% if it supposed to be a probability of severity of 50%

However, the choice of treatment should be an endogenous variable determined by various properties of patients and hospitals. In our data, we do not have detailed information of these properties, and thus this information is consigned to the error term, and all of the estimators have biases due to correlation between the error term and treatment selection. To examine such bias, Table 4 shows estimation results from the instrumental variables method. Comparing Table 4 and the result of simple estimation (Table 2), positive significance in for-profit hospitals and negative significance in PTCA and teaching hospital do not change. However, age, significant in Table 2, is not significant in Table 4; conversely, the number of beds, not significant in Table 2, is negatively significant in Table 4. PTCA has a more significantly lower probability of severe heart failure than CABG by 45 percentage points, and the direction of bias from the endogeneity seems to be zero. This result is confirmed also in Table 5 using the nonparametric propensity score matching method, and is robust from the definition of the kernel. Since the advantage of PTCA compared with CABG is 26% compared with 33% in this table , it is a little weaker than the instrumental variables results.

Table 6 indicates the result of simple estimation with less than one month hospitalization. This shows that age and for-profit hospitals are significantly positive and age-squared is significantly negative. These are almost the same as for the whole data set. PTCA is negative but not significant. In other words, PTCA has a lower probability of severe heart failure, but the difference is not significant. The advantage is only eight percentage points, and much lower than for the whole data set. Estimation using the instrumental

variables method (Table 8) has almost the same result as in Table 6, and PTCA is negative but not significant. In the nonparametric propensity score matching method, the signs of the coefficients depend on the definition of the kernel. PTCA has a lower probability of severe heart failure than CABG for two kernels and no significant difference for the others. In the former cases, PTCA reduces the probability of severe heart failure by 9–31 percentage points, and this difference is smaller than for the whole data set. Note that some reservations are needed in interpretation because of the instability of the results. This may due to be poorness-of-fit in estimation in treatment selection caused by a small number of observations.

To sum up, the advantage of PTCA over CABG cannot be denied, but with reservations. Even in the case of lack of fit in the treatment selection without a significant relationship, the nonparametric propensity score matching method results are significant in some cases compared with the instrumental variables method approach. This may imply the methodological superiority of nonparametric propensity score matching method over the instrumental variables method.

7 Estimation results of a comparison between invasive procedures and preserving procedures

Next, we compare invasive procedures and preserving procedures for patients with intensive or emergency care. Estimations are performed on the whole data set and restricted data sets (with less than one month hospitalization or blood type test). Tables 11–14 indicate the estimation results for the whole sample, Tables 15–18 the results for less than one month hospitalization , and Tables 19–22 the results for blood type test. Each sample produces estimation results of probit estimation of outcome on treatment selection (invasive procedures or preserving procedures) and other variables (Tables 11, 15, 19), probit estimation for treatment selection (Tables 12, 16, 19), probit estimation of outcome on

estimated probability of treatment selection as instrumental variables (Tables 13, 17, 21), and average treatment effects constructed by the nonparametric propensity score matching method using the estimated probability of invasive procedures as the choice of P (Tables 14, 18, 22).

In the simple probit estimation without consideration of endogeneity (Table 11) and with the instrumental variable method (Table 13), there are no significant variables. Even though both are insignificant, invasive procedures increase the probability of severe heart failure in the simple probit estimation, and decrease it with the instrumental variables method. On the other hand, the nonparametric propensity score matching method does not show any significant effect except with one kernel. In this kernel, invasive procedures increase the probability of severe heart failure significantly by about seven percentage points compared with preserving procedures.

Table 15 indicates the results of simple estimation using the data for less than one month hospitalization. The number of beds and for-profit hospitals are significantly negative. Invasive procedures is negative but not significant; that is, invasive procedures have a lower probability of severe heart failure, but this difference is not significant. The estimation results using the instrumental variables method (Table 16) have no significant coefficients, but invasive procedures have a significantly lower probability of severe heart failure than preserving procedures. For the nonparametric propensity score matching method, Table 17 shows that the signs of the coefficients are mixed across the kernels. Invasive procedures have a lower probability of severe heart failure than preserving procedures in two cases, but no significant difference in three cases. In the former cases, the differences are 7–9 percentage points, and are inconsistent with the case for the whole data set.

Table 18 indicates the result of a simple probit estimation using the data with blood type test, and this shows that non-profit hospital is significantly negative. Invasive procedures have a lower probability of severe heart failure, but this is not significant. The instrumental variables method results in Table 20 show that female, teaching hospital and invasive

procedures are significantly negative, and thus invasive procedures decrease the probability of severe heart failure by 20 percentage points, in comparison with preserving procedures. In contrast, for the nonparametric propensity score matching method results shown in Table 21, invasive procedures have a higher probability of severe heart failure than preserving procedures for all kernels, and the difference is 9–16 percentage points.

The results of comparison between invasive procedures and preserving procedures are summarized as follows. With the instrumental variables method, there is no significant difference between the two procedures for the whole sample and the less than one month hospitalization sample, but predominance of invasive procedures over preserving procedures in the sample with blood type test is indicated. For the nonparametric propensity score matching method, results show a lack of robustness for the whole sample and less than one month hospitalization sample, and we found superiority of preserving procedures over invasive procedures in the blood type test sample regardless of kernel. The predominance of these two procedures is unstable across sample definition, and thus estimation interpretations should be made carefully. The results of the instrumental variables method and the nonparametric propensity score matching method are inconsistent, and thus the usefulness of the instrumental variable method seems to be limited.

8 Concluding remarks

Using microdata of the SMC based on observational patient data, this paper compares the outcomes of patients with ischaemic heart disease between PTCA and CABG, or of patients with intensive or emergent care between invasive procedures and preserving procedures. Taking endogeneity of treatment selection into consideration, the instrumental variables and nonparametric propensity score matching methods are employed for estimation. We found that there is no case in which both methods are consistent regarding significance. Of five sample sets (two from section 6 and three from section 7), one sample set (the sample set

with blood type tests in comparison of invasive procedures with preserving procedures) has opposite results between the two methods: the nonparametric propensity score matching method results show that invasive procedures have a significantly higher probability of severe heart failure than preserving procedures, but the instrumental variables method shows the opposite result.

However, the nonparametric propensity score matching method seems to be more flexible than the instrumental variables method in general. Moreover, the theoretical advantages of the nonparametric propensity score matching method over the instrumental variable method are well known. If we believe the results of the nonparametric propensity score matching method, PTCA has a significantly lower probability of severe heart failure than CABG, with some reservations, and invasive procedures have a significantly higher probability of severe heart failure than preserving procedures in the blood type test sample.

These results contain two important reservations. Firstly, there are several problems with the SMC as described in Section 2. These data take samples from clinical records in hospitals all over Japan, and report all medical procedures. This kind of data is unique compared with other countries. However, there is only a little information about diseases or patients' conditions. Thus, clinical outcome has to be inferred from corresponding clinical procedures. This procedure depends heavily on some assumptions, in cases where the order of the clinical procedures is unknown.

The second problem is endogeneity of treatment selection. Two estimation methods are employed, but the results are not robust between the methods. This seems to be caused by insufficient information about the clinical course, pathological lesion, detailed heart function and so forth, and we cannot control these factors, which are closely related to treatment selection. A more organized and complicated nonparametric propensity score matching method should overcome this informational limitation in the future.

Acknowledgement

This paper is part of a set of research results of the Research Group of Innovations of Institutions in Aging Society (director: Professor Toshiaki Tachibanaki, Kyoto University) funded by the Millennium Project of the Economic Planning Agency. The authors would like to thank members of the research group for active discussion and Ms. Tomoko Nagase for supporting the research. Naturally, mistakes in this paper belong not to the research group but to the authors.

References

[1] Grines CL, Browne KF, Marco J, Rothbaum D, Stone GW, O'Keefe J, Overlie P, Donohue B, Chelliah N, Timmis GC: A comparison of immediate angioplasty with thrombolytic therapy for acute myocardial infarction. The Primary Angioplasty in Myocardial Infarction Study Group. N Engl J Med 1993;328:673-9.

[2] The Global Use of Strategies to Open Occluded Arteries in Acute Coronary Syndrome Angioplasty Substudy Investigators: A clinical trial comparing primary coronary angioplasty with tissue plasminogen activator for acute myocardial infarction. N Engl J Med 1997;336:1621-1628.

[3] Ryan TJ, Anderson JL, Antman EM, Braniff BA, Brooks NH, Califf RM, Hillis LD, Hiratzka LF, Rapaport E, Riegel BJ, Russell RO, Smith EE, weaver WD: ACC/AHA guidelines for the management of patients with acute myocardial infarction. A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee on Management of Acute Myocardial Infarction). J Am Coll Cardiol 1996;28:1328-1428.

[4] McClellan M: Uncertainty, Health-Care Technologies, and Health-Care Choices. American Economic Review. 1995;85:38-44.

[5] Blanwald E: Shattuck Kecture-Cardiovasucular medicine at the turn of the millenium: Triumph, consensus, and opportunities. N Eng J Med 1997;337:1360-69.

[6] McClellan M, McNeil BJ, Newhouse JP: Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. JAMA 1994;272:859-66

[7] Heidenreich P, McClellan M: The consequences of technological change in medicine 1975–1995: A literature review and synthesis. Mimeo, Stanford University. 1998

[8] Killip T, Kimball JT: Treatment of myocardial infarction in a coronary care unit. A two year experience with 250 patients. Am J Cardiol 1967;20:457-64.

[9] Forrester JS, Diamond GA, Swan HJ: Correlative classification of clinical and hemodynamic function after acute myocardial infarction. Am J Cardiol 1977;39:137-45.

[10] Baim DM, Grossman W: Cardiac Catheterization. Angiography, and Intervention, 5th ed. William & Wilkins, 1996.

[11] McClellan M, Noguchi H: Validity and interpretation of treatment outcome estimates using observational data: treatment of heart attacks in the elderly. Preliminary Draft.

[12] Heckman JJ, Ichimura H, Todd P: Matching as an econometric evaluation estimator: Evidence from evaluating a job training program. Review of Economic Studies 1997;64:605-654.

[13] Heckman JJ, Ichimura H, Todd P: Matching as an econometric evaluation estimator. Review of Economic Studies 1998;65:261-294.

[14] Heckman JJ, Smith J: Evaluating The welfare State. in Strom S. (ed.) Frisch Centenary. CUP, 1997.

[15] Coyte PC, Young W, Croxford R: Costs and outcomes associated with alternative discharge strategies following joint replacement surgery: analysis of an observational study using a propensity score. J Health Econ 2000;19:907-29.

[16] Silverman BW: Density estimation for statistics and data analysis. Chapman & Hall, 1992.

[17] Todd P Behrman, Cheng Y: Evaluating preschool program when length of exposure to the program varies: A nonparametric approach. Presented at the 8th World Congress of the Econometric Society, 2000.

[18] Efron B, Tibshirani R: An introduction to the bootstrap¿ Chapman & hall/CRC, 1998.

Table 1: Averages

	All samples	Less than one month hospital-ization sample
Severe Heart Failure	.354067	.2876254
PTCA	.6220096	.6822742
Age	64.09569	64.08027
Female	1.232057	1.200669
Number of Beds (Log)	6.366292	6.276663
Non-profit Hospital	.1028708	.1036789
For-profit Hospitals	.2416268	.2073579
Teaching Hospital	.2105263	.1672241

Note: The number of samples is 223 in all samples, and 152 in samples less than One Month Hospitalization.

Table 2: Estimation Results by Simple Probit (All Samples)

	Estimates	z-value	P-value	Marginal Effect
Age	.5966369	2.278	0.023	.17839615
Age^2	-.0042646	-2.089	0.037	-.00127514
Female	.1824614	0.402	0.688	.07255956
PTCA	-1.271928	-2.941	0.003	-.3625063
Number of Beds (Log)	-.0839713	-0.217	0.828	-.02510766
Non-profit Hospital	-.6163245	-1.145	0.252	-.2185034
For-profit Hospital	.9385328	2.336	0.019	.34456916
Teaching Hospital	-1.023907	-1.915	0.055	-.31989167
Constant	-19.25944	-2.218	0.027	

Note: The log likelihood is -118.22913, and quasi R^2 is 0.2338. The null hypothesis that all variables except with constant are 0 is rejected under 10% significant level.

Table 3: Estimation Results of Treatment Selection (All Samples)

	Estimates	z-value	P-value	Marginal Effect
Age	-.0469925	-0.231	0.818	-.0146637
Age^2	.0003431	0.211	0.833	.00010706
Female	-.0164387	-0.034	0.973	-.00554354
Number of Beds (Log)	-.6521113	-1.565	0.118	-.20348711
Non-profit Hospital	-.0280082	-0.052	0.959	-.00947649
For-profit Hospital	.8226628	2.079	0.038	.19909877
Teaching Hospital	-.5034769	-1.020	0.308	-.18805565
Constant	6.16708	0.849	0.396	

Note: The of likelihood is -122.61888, and quasi R^2 is 0.1112. The null hypothesis that all variables except with constant are 0 is rejected under 10% significant level.

Table 4: Estimation Results for Outcome using Instrumental Variable Method (All Samples)

	Estimate	z-value	P-value	Marginal Effect
Age	.372029	1.415	0.157	.12358848
Age^2	-.0026644	-1.307	0.191	-.0008851
Female	.0387424	0.083	0.934	.01539951
PTCA Selecting Probability(Estimates)	-9.112694	-2.037	0.042	-.45850835
Number of Beds (Log)	-1.335078	-1.778	0.075	-.44351456
Non-profit Hospital	-.324814	-0.665	0.506	-.12454906
For-profit Hospitals	2.679349	2.639	0.008	.53648192
Teaching Hospital	-2.238602	-2.381	0.017	-.44893838
Constant	1.680618	0.144	0.886	

Note: The log likelihood is -130.34474, and quasi R^2 is 0.1553. The null hypothesis that all variables except with constant are 0 is rejected under 10% significant level.

Table 5: Estimation Results for Outcome using Nonparametric Propensity Score Method (All Samples)

Kernel	Average	95%CI Upper	95%CI Lower
Epanechnikov	-.3262025	-.4336698	-.2349667
Biweight	-.6030771	-1.73575	-.197184
Triangular	-.3283879	-.4264668	-.2277659
Gaussian	-.2558149	-.3311637	-.1775348
Rectangular	-.2674383	-.3735231	-.1592936

Table 6: Estimation Results without accounting for Endogeneity (Less than One Month Hospitalization Samples)

	Estimates	z-value	P-value	Marginal Effect
Age	1.566249	2.189	0.029	.41259756
Age^2	-.0112827	-2.165	0.030	-.00297222
Female	-.3810702	-0.805	0.421	-.08747231
PTCA	-.3374903	-0.611	0.541	-.07921911
Number of Beds (Log)	-.6504081	-1.131	0.258	-.17133729
Non-profit Hospital	-.7992492	-1.290	0.197	-.14610727
For-profit Hospital	.9374392	2.005	0.045	.33579038
Teaching Hospital	-.4971548	-0.689	0.491	-.10735859
Constant	-49.41318	-1.971	0.049	

Note: The log likelihood is 699.40835, and quasi R^2 is 0.2884. The null hypothesis that all variables except with constant are 0 is rejected under 10% significant level.

Table 7： Estimation Results of Treatment Selection（Less than One Month Hospitalization　Samples ）

	Estimates	z-value	P-value	Marginal Effect
Age	.3392646	1.237	0.216	.08025054
Age^2	-.0028532	-1.323	0.186	-.0006749
Female	.1156836	0.199	0.842	.02496931
Number of Beds（Log）	-.1495782	-0.264	0.792	-.03538163
Non-profit Hospital	-.1138539	-0.189	0.850	-.02772146
For-profit Hospital	1.403338	2.767	0.006	.13964116
Teaching Hospital	-.6261293	-0.872	0.383	-.18888053
Constant	-8.272516	-0.822	0.411	

Note:The log likelihood is 64.809806, and quasi R^2 is 0.1416. The null hypothesis that all variables except with constant are 0 is not rejected under 10

Table 8： Estimation Results of Outcome using Instrumental Variable Method（Less than One Month Hospitalization Samples)

	Estimates	z-value	P-value	Marginal Effect
Age	2.096679	2.581	0.010	.53839867
Age^2	-.0155398	-2.553	0.011	-.0039904
Female	-.212923	-0.423	0.673	-.05045125
PTCA Choosing P-value（Estimates）	-5.3015	-1.313	0.189	-.18025783
Number of Beds（Log）	-.6767463	-1.137	0.256	-.17377923
Non-profit Hospital	-.878099	-1.452	0.147	-.14372993
For-profit Hospital	2.382303	1.872	0.061	.74867916
Teaching Hospital	-1.279259	-1.257	0.209	-.16612725
Constant	-61.99018	-2.338	0.019	

Note:The log likelihood is 68.383353, and quasi R^2 is 0.3043. The null hypothesis that all variables except with constant are 0 is not rejected under 10% significant level.

Table 9: Outcome Estimates using Nonparametric Propensity Score Method (Less than One Month Hospitalization Samples)

	Average	95% CI Upper	95% CI Lower
Epanechnikov	-.0170832	-.1503817	.1186205
Biweight	-.0926619	-.1833897	-.0032472
Triangular	-.0733864	-.1926208	.0522043
Gaussian	-.3057324	-.4355384	-.1637794
Rectangular	.0194312	-.1149288	.1604738

Table 10: Averages

	All Samples	Less than One Month Hospital-ization Samples	Blood Type Test Samples
Severe Heart Failure	.2	.1795717	.2071006
Invasive Procedures	.4864198	.4629325	.4319527
Age	66.58148	65.24547	65.88462
Female	1.398765	1.357496	1.45858
Number of Beds (Log)	6.255634	6.283452	6.227614
Non-profit Hospital	.0567901	.0560132	.0710059
For-profit Hospital	.3	.3212521	.2899408
Teaching Hospital	.2802469	.2784185	.2781065

Note: The number of samples is 447 in all samples, 345 in Less than One Month Hospital-ization Samples and 193 in Blood Type Test Samples.

Table 11：Estimation Results of Simple Probit（All Samples）

	Estimates	z-value	P-value	Marginal Effect
Age	.0741479	1.068	0.285	.02259805
Age^2	-.0005246	-1.004	0.315	-.00015989
Female	-.2150448	-0.767	0.443	-.06286574
Invasive Procedures	.0131575	0.046	0.963	.00418386
Number of Beds（Log）	-.1300969	-0.817	0.414	-.03964964
Non-profit Hospital	.6017939	1.352	0.176	.2205717
For-profit Hospital	-.371838	-1.166	0.243	-.10180657
Teaching Hospital	.0640221	0.185	0.854	.02070098
Constant	-2.005997	-0.801	0.423	

Note:The log likelihood is 241.63305, and quasi R^2 is 0.0597. The null hypothesis that all variables except with constant are 0 is not rejected under 10% significant level.

Table 12：Estimation Results of Treatment Selection（All Samples）

	Estimates	z-value	P-value	Marginal Effect
Age	.0446887	0.761	0.447	.01593775
Age^2	-.0003181	-0.712	0.476	-.00011343
Female	-.3432881	-1.280	0.201	-.13144744
Number of Beds（Log）	.2528266	1.739	0.082	.09016785
Non-profit Hospital	-.9171804	-2.177	0.030	-.30588178
For-profit Hospital	-.7173538	-2.446	0.014	-.2534792
Teaching Hospital	-.2130057	-0.660	0.510	-.08306005
Constant	-2.148546	-0.973	0.331	

Note:The log likelihood is 279.65664, and quasi R^2 is 0.0947. The null hypothesis that all variables except with constant are 0 is rejected under 10% significant level.

Table 13： Estimation Results of Outcome using Instrumental Variable Method（All Samples）

	Estimates	z-value	P-value	Marginal
Age	.5793714	1.275	0.202	.17255549
Age^2	-.0041362	-1.274	0.203	-.0012318
Female	-4.087484	-1.202	0.229	-.2343995
Invasive Procedures Choosing P-value（Estimates）	-31.31108	-1.161	0.246	-.2344002
Number of Beds（Log）	2.542564	1.070	0.285	.75725776
Non-profit Hospital	-10.13291	-1.088	0.277	-.2344002
For-profit Hospital	-8.998187	-1.220	0.222	-.2344002
Teaching Hospital	-2.039323	-1.032	0.302	-.2315432
Constant	-9.269772	-1.226	0.220	

Note:The log likelihood is 236.679, and quasi R^2 is 0.0790. The null hypothesis that all variables except with constant are 0 is not rejected under 10% significant level.

Table 14： Estimation Results of Outcome using Nonparametric Propensity Score Method（All Samples）

	Average	95% CI Upper	95% CI Lower
Epanechnikov	.0690913	.010771	.1381654
Biweight	.0019177	-.0558639	.0623594
Triangular	.0623893	-.0016824	.120463
Gaussian	.0057966	-.0521428	.0693814
Rectangular	.0239519	-.0337745	.0846671

Table 15：Estimation Results of Simple Probit（Less than One Month Hospitalization　Samples）

	Estimates	z-value	P-value	Marginal Effect
Age	.0401907	0.487	0.627	.01056798
Age^2	-.000204	-0.314	0.754	-.00005364
Female	-.1112446	-0.343	0.732	-.02923743
Invasive Procedures	-.2025692	-0.615	0.539	-.05100845
Number of Beds (Log)	-.4093291	-1.647	0.100	-.10763148
Non-profit Hospital	.3665715	0.761	0.447	.11612355
For-profit Hospital	-.8165761	-2.023	0.043	-.14845607
Teaching Hospital	.444591	0.828	0.408	.14421667
Constant	.4466926	0.143	0.886	

Note:The log likelihood is 162.79032, and quasi R^2 is 0.1112. The null hypothesis that all variables except with constant are 0 is rejected under 10% significant level.

Table 16：Estimation Results of Treatment Selection（Less than One Month Hospitalization　Samples）

	Estimates	z-value	P-value	Marginal Effect
Age	.0829551	0.922	0.357	.02864868
Age^2	-.0006732	-0.941	0.347	-.00023249
Female	-.6441801	-2.117	0.034	-.21512405
Number of Beds（Log）	.2396488	1.211	0.226	.08276313
Non-profit Hospital	-.6660112	-1.471	0.141	-.22088191
For-profit Hospital	-.6470829	-1.855	0.064	-.21589624
Teaching Hospital	-.2074055	-0.494	0.621	-.07747569
Constant	-2.836181	-0.892	0.372	

Note:The log likelihood is 209.61745, and quasi R^2 is 0.1045. The null hypothesis that all variables except with constant are 0 is rejected under 10% significant level.

Table 17： Estimation Results of Outcome using Instrumental Variable Method （Less than One Month Hospitalization Samples)

	Estimates	z-value	P-value	Marginal
Age	.1928973	0.809	0.418	.05067594
Age^2	-.0014347	-0.753	0.451	-.0003769
Female	-1.57529	-0.734	0.463	-.1858438
Invasive Procedures Choosing P-value （Estimates)	-6.635654	-0.704	0.482	-.1931546
Number of Beds （Log)	.1059676	0.140	0.889	.0278387
Non-profit Hospital	-1.190768	-0.511	0.610	-.1733162
For-profit Hospital	-2.403003	-1.035	0.301	-.1926156
Teaching Hospital	.0349378	0.045	0.964	.00972145
Constant	-1.528419	-0.357	0.721	

Note:The log likelihood is 162.84529, and quasi Rα is 0.1109. The null hypothesis that all variables except with constant are 0 is rejected under 10% significant level.

Table 18：Estimation Results of Outcome using Nonparametric Propensity Score Method（Less than One Month Hospitalization Samples）

	Average	95% CI Upper	95% CI Lower
Epanechnikov	-.0644408	-.1277267	.0080246
Biweight	-.0884281	-.1482805	-.0256893
Triangular	-.0596314	-.1262428	.0085508
Gaussian	-.072562	-.1380501	-.0053486
Rectangular	-.044513	-.1040212	.0212697

Table 19：Estimation Results of Simple Probit（Blood Type Test Samples）

	Estimates	z-value	P-value	Marginal Effect
Age	.024964	0.260	0.795	.00637125
Age^2	-.0001039	-0.145	0.885	-.00002651
Female	-.5475136	-1.116	0.264	-.12213132
Invasive Procedures	-.2815678	-0.647	0.517	-.07170522
Number of Beds (Log)	-.2903051	-1.138	0.255	-.074091
Non-profit Hospital	1.404206	2.377	0.017	.51501181
For-profit Hospital	-.5313012	-0.822	0.411	-.11951618
Teaching Hospital	.3107833	0.562	0.574	.10007992
Constant	.6818123	0.177	0.859	

Note:The log likelihood is 88.747851, and quasi R^2 is 0.1716. The null hypothesis that all variables except with constant are 0 is not rejected under 10% significant level.

Table 20：Estimation Results of Treatment Selection（Blood Type Test Samples）

	Estimates	z-value	P-value	Marginal Effect
Age	.2601578	1.838	0.066	.08028642
Age^2	-.0020923	-1.865	0.062	-.0006457
Female	-1.248489	-2.882	0.004	-.36477547
Number of Beds (Log)	.4424047	1.421	0.155	.136529
Non-profit Hospital	-.5958351	-0.977	0.329	-.21450428
For-profit Hospital	.2633366	0.432	0.666	.10475259
Teaching Hospital	-1.568982	-2.367	0.018	-.40426955
Constant	-8.14102	-1.652	0.098	

Note:The log likelihood is 105.12992, and quasi R^2 is 0.2140. The null hypothesis that all variables except with constant are 0 is rejected under 10% significant level.

Table 21: Estimation Results of Outcome using Instrumental Variable Method (Blood Type Test Samples)

	Estimates	z-value	P-value	Marginal
Age	.3454181	1.669	0.095	.0785599
Age^2	-.0026482	-1.636	0.102	-.000602
Female	-2.437649	-2.882	0.004	-.192299
Invasive Procedures Selecting Probability(Estimate)	-6.041505	-2.752	0.006	-.192774
Number of Beds(Log)	.0093336	0.041	0.967	.0021227
Non-profit Hospital	.1712881	0.238	0.812	.050305
For-profit Hospital	-.4380939	-0.668	0.504	-.096965
Teaching Hospital	-1.611602	-2.509	0.012	-.186192
Constant	-4.382809	-0.859	0.391	

Note: The log likelihood is 79.3828, and quasi R^2 is 0.2590. The null hypothesis that all variables except with constant are 0 is rejected under 10% significant level.

Table 22： Estimation Results of Outcome using Nonparametric Propensity Score Method（Blood Type Test Samples）

	Average	95% CI Upper	95% CI Lower
Epanechnikov	.1223203	.0416913	.214348
Biweight	.0874121	.0143986	.1618662
Triangular	.1621314	.0793257	.2544301
Gaussian	.0944667	.0232313	.1762032
Rectangular	.1441248	.0517966	.2363143