

RESEARCH

Open Access



# Dealing with missing data in laboratory test results used as a baseline covariate: results of multi-hospital cohort studies utilizing a database system contributing to MID-NET<sup>®</sup> in Japan

Maki Komamine<sup>1,2\*</sup> , Yoshiaki Fujimura<sup>3</sup>, Masatomo Omiya<sup>4</sup> and Tosiya Sato<sup>1</sup>

## Abstract

**Background** To evaluate missing data methods applied to laboratory test results used for confounding adjustment, utilizing data from 10 MID-NET<sup>®</sup>-collaborative hospitals.

**Methods** Using two scenarios, five methods dealing with missing laboratory test results were applied, including three missing data methods (single regression imputation (SRI), multiple imputation (MI), and inverse probability weighted (IPW) method). We compared the point estimates of adjusted hazard ratios (aHRs) and 95% confidence intervals (CIs) between the five methods. Hospital variability in missing data was considered using the hospital-specific approach and overall approach. Confounding adjustment methods were propensity score (PS) weighting, PS matching, and regression adjustment.

**Results** In Scenario 1, the risk of diabetes due to second-generation antipsychotics was compared with that due to first-generation antipsychotics. The aHR adjusted by PS weighting using SRI, MI, and IPW by the hospital-specific approach was 0.61 [95%CI, 0.39–0.96], 0.63 [95%CI, 0.42–0.93], and 0.76 [95%CI, 0.46–1.25], respectively. In Scenario 2, the risk of liver injuries due to rosuvastatin was compared with that due to atorvastatin. Although PS matching largely contributed to differences in aHRs between methods, PS weighting provided no substantial difference in point estimates of aHRs between SRI and MI, similar to Scenario 1. The results of SRI and MI in both scenarios showed no considerable changes, even upon changing the approaches considering hospital variations.

**Conclusions** SRI and MI provide similar point estimates of aHR. Two approaches considering hospital variations did not markedly affect the results. Adjustment by PS matching should be used carefully.

**Keywords** Pharmacoepidemiology, Observational data, Database, Laboratory test item, Missing data

\*Correspondence:

Maki Komamine

[komamine-maki@pmda.go.jp](mailto:komamine-maki@pmda.go.jp)

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

In April 2018, the operation of the Medical Information Database Network (MID-NET<sup>®</sup>) began as a national project aimed at utilizing real-world data for drug safety assessments in Japan [1–4]. MID-NET<sup>®</sup> is a database systems network consisting of 10 collaborative organizations (23 collaborative hospitals) that can analyze data derived from claim data, diagnosis procedure combination data, and electronic medical record (EMR) data at the individual-level [2].

Laboratory test results derived from EMR data have detailed information on clinical symptoms [5] and are expected to be used for confounding adjustments in drug safety assessments. However, the appropriate use of these test results is difficult since a number of data obtained during routine medical care may be missing in datasets for analysis. Therefore, it is essential to appropriately select and apply existing methods to reduce bias due to missing data (hereinafter referred to as “missing data methods”).

Choosing a missing data method requires an understanding of the missing data sources and missing data mechanism [6–8]. Raebel et al., in their study using the US Food and Drug Administration Mini-Sentinel Distributed Database [6], reported that missing data sources of laboratory test results include testing outside of contracted laboratories, patient location where testing was conducted, patient clinical features, etc. In the study of the 10 MID-NET<sup>®</sup>-collaborative hospitals of the Tokushukai Medical Group [9], authors reported that patient background and setting of ordering system of laboratory tests were the main missing data sources. Although their impact was not evaluated because they were unobserved, the measurement policies by physicians and institutions were considered as potential sources as well. Missing data mechanism can be classified as missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR) based on the relationship of

missing data probability with missing data and observed values [10] (See Table 1).

In the utilization of laboratory test results contained in MID-NET<sup>®</sup> to confounding adjustments, it is not clear what impact different missing data methods have on the result. Furthermore, owing to the differences between hospitals in terms of the proportion of patients with missing data (hereinafter, “missing proportion”) and the relationship between missing data and patient background [9], hospital variations regarding missing data should be well-considered.

In this study, the application of missing data methods was carried out to two scenarios of drug safety assessment adjusted by one laboratory test item. We evaluated the impact that missing data methods/approaches to hospital variations had on the interpretation of effect estimates and results.

## Methods

### Study's scope

We considered a drug safety assessment to estimate an exposure effect using the Cox proportional-hazards model, which is commonly used in cohort designs. Assuming one laboratory test item as a confounder, we applied five missing methods.

### Database and target hospitals

This study used the database system for MID-NET<sup>®</sup>-collaborative organizations of Tokushukai Medical Group (hereinafter, “Tokushukai database”), which has the largest number of MID-NET<sup>®</sup>-collaborative hospitals (10 hospitals; Supplementary Table S1). Supplementary Table S2 demonstrates the data items used for analysis.

### Definition of missing data

As per a previous study, missing data were defined as “data that would be meaningful for the analysis but not available during a specific period before the first

**Table 1** Missing data mechanisms and their examples

Missing data mechanism	Description	Example
<b>Missing completely at random (MCAR)</b>	The probability that values are missing is unrelated to either the specific missing values that should have been obtained or the set of observed values.	Missing data due to equipment failure.
<b>Missing at random (MAR)</b>	The probability that values are missing depends on the set of observed values but is further unrelated to the specific missing values that should have been obtained.	Missing data of blood glucose can be said to be MAR given age, if younger patients are less (or more) likely to have their blood glucose measured than older patients.
<b>Missing not at random (MNAR)</b>	The probability that values are missing is related to the specific missing values that should have been obtained, in addition to the ones actually obtained.	If there are data that are unobtained but can influence the missingness of blood glucose, missing data of blood glucose cannot be said to be MAR but said to be MNAR.

prescription date” [9]. The specific period, based on the results of this prior study, was set to 90 days.

### Missing methods

MAR, the reasoning behind which is explained in [Missing methods](#) section, was assumed as the missing data mechanism in this study. The following four methods were considered as missing data methods providing unbiased results (hereinafter, “MAR-based methods”) when both the MAR assumption and the correct model specification used in the missing data method (hereinafter collectively referred to as “missing data models”) were correct: single regression imputation (SRI), multiple imputation (MI), inverse probability weighted (IPW) method, and likelihood-based method [11–14]. In the MID-NET® database system, SAS version 9.4 (SAS Institute Inc., Cary, NC, USA) can be used. Since SRI, MI, and IPW methods are implemented in SAS version 9.4, we adopted these MAR-based methods.

In this study, five missing methods were applied in order to evaluate their impact on the effect estimation of the outcome model. The methods included the above three MAR-based methods, a method excluding a laboratory test item from baseline covariates (Exclusion), and a complete case (CC) method (Table 2). Table 2 provides a brief description of the three MAR-based methods [7, 15–23] and the settings of this study. In the application of MAR-based methods, two approaches were adopted to consider hospital variations in missing data. The hospital-specific-approach implements a missing data model within each hospital cohort. The overall-approach, consequently, implements a missing data model to the overall cohort (combined hospitals cohorts) and uses hospital as a fixed effect covariate (Fig. 1).

### Confounding adjustment and outcome model

For the confounding adjustment methods, we adopted two propensity score (PS) methods (PS weighting and PS matching) as well as an outcome model method with confounding factors as covariates (regression adjustment). The confounding factors, other than the laboratory test item, were patient related factors (see Table 3) and the corresponding hospitals. Since the target population in the scenarios of this study is an exposed population, the standardized mortality ratio weighting (SMRW) [27] was used for PS weighting.

Then, the point estimate of adjusted hazard ratio (aHR) and 95% confidence interval (CI) were calculated using the Cox proportional-hazard model. In PS matching, stratified Cox proportional-hazard model by matched pairs was used. Robust standard error was used to calculate 95% CI. In the combination of PS methods and MI, aHR estimation was performed for  $m$  imputed datasets,

and  $m$  estimates were combined (“within approach”) [34] (see Supplementary Fig. S1).

### Scenarios

Based on our previous research [9], we selected the scenarios of two cohort studies that evaluated the relation between drugs and their known risks (see Supplementary Fig. S2). Scenario 1 was the risk of diabetes associated with antipsychotic drug use. In Scenario 1, the blood glucose level before prescription of antipsychotic drugs was the target laboratory test item. Scenario 2 was the risk of hepatic injury associated with HMG-CoA reductase inhibitors (statins) use. In Scenario 2, the target laboratory test item was either the alkaline phosphatase (ALP), alanine transaminase (ALT), low-density lipoprotein cholesterol (LDL-chol), or triglyceride (TG) levels before statins prescription. In Scenario 2, four laboratory test items were included in the baseline covariates of the outcome model, and CC method, SRI, MI, and IPW method were set for each laboratory test item. Table 3 shows the detailed settings of the scenarios, including cohort sizes and the number of complete cases.

### Protocol approval and statistical analysis

Our study protocol was approved by the Kyoto University Graduate School, Faculty of Medicine, and Kyoto University Hospital Ethics Committee in November 2018 (R1793). Statistical analyses were performed using SAS version 9.4 (SAS Institute, Cary, NC, USA).

### Results

Overall cohort sizes and the numbers of complete cases were summarized in Table 3 and Supplementary Figs. S3 and S4. Patient backgrounds including incidence rate are demonstrated in Supplementary Tables S3–S6 (Tables S3 and S5 for overall cohorts and complete cases, and Tables S4 and S6 for hospital cohorts).

#### Scenario 1

##### Confounding adjustment by PS weighting

The aHR of Exclusion was 0.52 (95% CI, 0.34–0.81) (Fig. 2). CC method, SRI, MI, and IPW method contained blood glucose level as a covariate. The aHRs of SRI and MI by the hospital-specific-approach were 0.61 (95% CI, 0.39–0.96) and 0.63 (95% CI, 0.42–0.93), respectively; the point estimates were relatively close, but the width of the 95% CI was slightly narrower in MI. Conversely, the aHRs of the CC and IPW methods by hospital-specific-approach were 0.78 and 0.76, respectively, and thus higher than those of SRI and MI.

The aHRs of SRI, MI, and IPW method by the overall-approach showed no substantial differences compared

**Table 2** Overview of the five missing methods

Brief description	Limitation/ Concern	Setting in this study
<p>Excluded a laboratory test item from baseline covariates. (Exclusion)</p> <p>Included only patients with laboratory test results (CC method)</p> <ul style="list-style-type: none"> <li>The method may still yield valid results only if the missing values are not related to the outcome.</li> <li>In the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement [24], it is mentioned that the main differences with the result of the CC method should be discussed when applying MI or IPW method.</li> </ul>		<ul style="list-style-type: none"> <li>Supplementary evaluation targets.</li> <li>Supplementary evaluation targets.</li> </ul>
<p>Accounted for missing laboratory test results using SRI</p> <ul style="list-style-type: none"> <li>The method that assumes an imputation model for the distribution of missing values given the observed values and imputes the missing values with the predicted values obtained.</li> </ul>	<ul style="list-style-type: none"> <li>Only one imputation model is created, so it does not take into account the uncertainty of lack of data.</li> </ul>	<ul style="list-style-type: none"> <li>Imputation model: linear regression model<sup>a</sup></li> <li>The result to be complemented is logarithmic transformation.</li> </ul>
<p>Accounted for missing laboratory test results using MI</p> <ul style="list-style-type: none"> <li>In this method, after assuming the imputation model, imputation is repeated multiple times (<i>m</i>) using random numbers. The parameter estimation of the outcome model is performed for <i>m</i> imputed datasets, and the <i>m</i> estimates are integrated by Rubin's rule, etc.</li> <li>The algorithm is complicated, but it can be implemented by available statistical analysis software such as SAS.</li> <li>Uncertainty between complementary models that single regression imputation cannot handle is considered.</li> <li>The imputation model should include not only the predictors of missing values but also the covariates of the outcome model.</li> <li>It is assumed that the distribution of missing values follows a multivariate normal distribution, but there are reports that it is robust even if this assumption does not hold [15].</li> </ul>	<ul style="list-style-type: none"> <li>Since it can be implemented by an existing program, it may be implemented without sufficient consideration of the settings of variables to be included in the model.</li> </ul>	<ul style="list-style-type: none"> <li>Imputation model: linear regression model<sup>a</sup>.</li> <li>The result to be complemented is logarithmic transformation.</li> <li>Imputation method: predicted mean matching<sup>b</sup></li> <li>Imputation count: 10 times<sup>c</sup></li> <li>Integration method: Rubin's rule</li> </ul>
<ul style="list-style-type: none"> <li>Imputation methods include regression imputation and predicted mean matching.</li> </ul>		

**Table 2** (continued)

Brief description	Limitation/ Concern	Setting in this study
<p><b>Accounted for missing laboratory test results using the IPW method</b></p> <ul style="list-style-type: none"> <li>• The method uses an estimation formula weighted by the inverse of the probability of observing the data.</li> <li>• A model is assumed for the distribution of the observed probabilities given the observed values, and the complete case is weighted by the inverse of the obtained observed probabilities to estimate the outcome model parameters.</li> <li>• A logistic regression model is generally used for the model for the probability of observing the data.</li> </ul>	<ul style="list-style-type: none"> <li>• Parameter estimation of outcome model becomes unstable in the presence of patients with extremely large weight.</li> </ul>	<ul style="list-style-type: none"> <li>• Model for the probability of observing the data: logistic regression model<sup>a,d</sup></li> </ul>

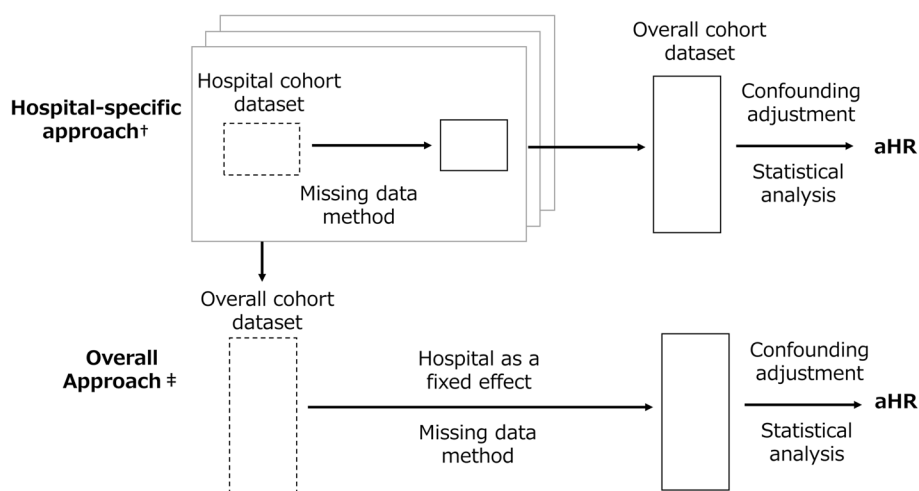
*Abbreviations:* CC Complete case, IPW Inverse probability weighted, MCAR Missing completely at random, MI Multiple imputation, SRI Single regression imputation

<sup>a</sup> Included patient background factors (sex, age, year of cohort entry, hospitalization, first visit, emergency care, class number of concomitant medications, complications, and concomitant medication), exposure variables, follow-up period, and event indicators. We referred to the study by White et al. [25] for the follow-up period and event indicator

<sup>b</sup> Predicted mean matching is a method of randomly selecting one observed value close to the predicted value obtained by the imputation model, and it was selected because it can take into account the underestimation of the variance that can occur by regression imputation

<sup>c</sup> Determined by referring to prior research by Raebel et al. [6] that used the Mini-Sentinel Distributed Database

<sup>d</sup> Including outcome information as covariates were referred to the study by Xu Q et al. [26]



**Fig. 1** Overview of the two approaches considering hospital variations. Abbreviations: aHR, adjusted hazard ratio. † In the overall approach, the effects of missing data sources at the hospital level, including unobserved missing data sources, can be considered as the hospital effect of the fixed effect, if not all. ‡ In the hospital-specific approach, the interaction of patient background factors with hospitals can be considered

with the estimates by the hospital-specific-approach (Fig. 3).

**Confounding adjustment by PS matching**

The aHRs of SRI and MI by the hospital-specific-approach were 1.10 (95% CI, 0.72–1.66) and 1.01 (95% CI, 0.34–2.97), respectively. Although there were no substantial differences in point estimates, the width of the 95% CI was larger in MI (Fig. 2). The aHR of the IPW method was 1.31 by the hospital-specific-approach but reduced to 0.94 by the overall-approach (Fig. 3).

**Regression adjustment**

The aHRs of SRI and MI by the hospital-specific-approach were relatively close in terms of the point estimates and the 95% CIs (Fig. 2).

**Scenario 2**

**Confounding adjustment by PS weighting**

The aHR of Exclusion was 1.32 (95% CI, 0.83–2.08). CC method, SRI, MI, and IPW method included each laboratory test item as a covariate. The aHRs of SRI and MI by hospital-specific-approach varied between 1.20 and 1.32 through all laboratory test items (Fig. 4). While there were no substantial differences in the point estimates of SRI and MI, the width of the 95% CI was slightly narrower in MI; for example, the aHR of SRI and MI for LDL-chol were 1.24 (95% CI, 0.78–1.96) and 1.20 (95% CI, 0.84–1.71), respectively. The aHR of IPW method for ALT was 1.21, whereas those for ALP and LDL-chol were close to 1. The aHRs of the CC method varied depending on the type of laboratory test item. Accordingly, the aHRs for ALT and LDL-chol were 1.29 and 1.08, respectively.

The aHRs of SRI and MI by the overall-approach ranged between 1.19 and 1.32 through all laboratory test items (Fig. 5). For some laboratory test items, especially ALP, the aHR of MI tended to be higher than those determined by the hospital-specific-approach.

**Confounding adjustment by PS matching**

For all laboratory test items, the differences in aHR between methods were higher and the width of 95% CI was greater than the values obtained with other confounding adjustment methods.

**Regression adjustment**

The aHRs of SRI and MI by the hospital-specific-approach ranged from 1.22 to 1.28 throughout all laboratory test items. For the results in each laboratory test item, both the point estimates and the 95% CIs were relatively close in SRI and MI (Fig. 4). However, the aHR of the IPW method was lower than those of SRI and MI.

**Discussion**

We evaluated the impact of five missing methods, approaches to hospital variations, and confounding adjustment methods on the effect estimation in two different scenarios. In [Missing methods](#) and [Approaches considering hospital variations](#) sections, we excluded discussions on PS matching in Scenario 2 were excluded and were, instead, included in [Confounding adjustment methods](#) section.

**Table 3** Details of scenarios 1 and 2

#	Scenario question/ Background	Study cohort	Target test item with missing	Outcome measure	Follow-up period	Patient-related factors <sup>d</sup>	Cohort size/Number (%) of complete cases
1	<p>Whether the risk of diabetes among SGA users is higher than that among FGA users?</p> <p>Glucose metabolic disorders are a known risk of SGA [28, 29]. Therefore, the diabetes risk of SGA is compared to that of FGA.</p>	<p>Inclusion criteria:</p> <ul style="list-style-type: none"> <li>• Patients initiated with any SGA<sup>a</sup> or FGA<sup>b</sup> during the study period (January 01, 2015, to December 31, 2017).</li> <li>• New users of monotherapy for SGA or FGA: Patients with no prescription for the drugs for &gt; 180 days prior to the first prescription of SGA or FGA.</li> </ul> <p>Exclusion criteria:</p> <ul style="list-style-type: none"> <li>• Patients with a diagnosis of dementia (ICD10: F01, F03, G30, G310, etc.) in the 180 days before the first prescription of any SGA or FGA.</li> <li>• Patients with a diagnosis of diabetes (ICD10: E10, E11, E12, E13, E14, O24) in the 180 days before the first prescription of any SGA or FGA.</li> </ul>	<p>Blood glucose</p> <p>The blood glucose or HbA1c before prescribing the drug was considered to be the potential confounder. The missing proportions at each hospital cohort were approximately 30–60% for HbA1c and approximately 5–40% for blood glucose. We chose blood glucose because we were interested in the effects of moderate missing data.</p>	<p>Blood glucose</p> <p>≥ 200 mg/dL or HbA1c(NGSP) ≥ 6.5%</p> <p>Note:</p> <ul style="list-style-type: none"> <li>• We convert HbA1c (JDS) to HbA1c (NGSP)</li> </ul>	<p>From the date of the first prescription of SGA or FGA to the first date of the following: occurrence of diabetes; end of exposure; end of study period (December 31, 2017); change to a different antipsychotic drug; or addition of a different antipsychotic drug; end of observation period<sup>c</sup>.</p>	<p>Sex, age, year of cohort entry, hospitalization, first visit, emergency care (at the date of the first prescription), class number of concomitant medications, complications<sup>e</sup>, concomitant medication<sup>f</sup> (180 days prior to the date of the first prescription of any anti-diabetic drug)</p>	<p>Overall cohort = 3,430 (SGA users: 1,087, FGA users: 2,343)</p> <p>Complete cases of blood glucose = 2,990 (87.2%)</p>

**Table 3** (continued)

#	Scenario question/ Background	Study cohort	Target test item with missing	Outcome measure	Follow-up period	Patient-related factors <sup>d</sup>	Cohort size/ Number (%) of complete cases
2	<p>Whether users of rosuvastatin have a different risk of hepatic injury than atorvastatin users? Liver damage is a known risk of HMG-CoA reductase inhibitors (statins). Both rosuvastatin and atorvastatin are contraindicated in patients with impaired liver function and have similar levels of alertness for liver damage [30, 31]. Although the results of some observational studies support this [32], atorvastatin (especially at high doses) has been reported to have a higher risk of causing liver damage than other statins [33]. Therefore, the risk of liver damage in rosuvastatin is compared with that in atorvastatin.</p>	<p>Inclusion criteria:                      • Patients initiated with any rosuvastatin or atorvastatin during the study period (January 01, 2015, to December 31, 2017).                      • New users of monotherapy for statin as in Scenario 1.                      Exclusion criteria:                      Patients with a diagnosis of hepatic injury (ICD10B18, K70-K76, K770, K778) within the 180 days before the first prescription of any statin.</p>	<p>ALT, ALP, LDL-cholesterol, TG                      We selected the above four laboratory tests, thinking that the values of the liver function-related and lipid metabolism-related laboratory tests before prescribing the drug would be potential confounders.</p>	<p>Diagnosis of hepatic injury (ICD10: K71-K76, K770, K778).</p>	<p>From the date of the first prescription of rosuvastatin or atorvastatin to the first date of the following: occurrence of hepatic injury, end of exposure, end of study period (December 31, 2017), change to another statin, or addition of another statin, end of observation period.</p>	<p>Sex, age, year of cohort entry, hospitalization, first visit, emergency care (at the date of the first prescription), class number of concomitant medication, complications<sup>g</sup>, concomitant medication<sup>h</sup> (180 days prior to the date of the first prescription of any anti-diabetic drug)</p>	<p>Overall cohort = 4,119 (rosuvastatin users: 2,538; atorvastatin users: 1,581)                      Complete cases of ALT, ALP, LDL-cholesterol, and TG = 3,761 (91.3%), 2,996 (72.7%), 3,176 (77.1%), and 3,306 people (80.3%), respectively.</p>

**Abbreviations:** ALP Alkaline phosphatase, ALT Alanine transaminase, FGA First-generation antipsychotic, HbA1c Hemoglobin A1c, ICD International Classification of Diseases, JDS The Japan Diabetes Society, LDL-cholesterol Low-density lipoprotein cholesterol, MGSP The National Glycohemoglobin Standardization Program, SGA Second-generation antipsychotic, TG Triglyceride

<sup>a</sup> Aripiprazole, bronanserin, clozapine, olanzapine, perospirone, quetiapine

<sup>b</sup> Bromperidol, chlorpromazine, dlocapramine, fluphenazine, haloperidol, levomepromazine, mosapramine, nemonapride, perphenazine, pimozide, prochlorperazine, propellizazine, spiperone, sultopride, thimipterone, zotepine

<sup>c</sup> From the perspective of ensuring patient traceability, hospital visits are considered to be consecutive if the interval between visit dates is within a certain range, and we followed these visits to calculate "observation period" per patient

<sup>d</sup> Included in missing data method and confounding adjustment

<sup>e</sup> Hepatitis, liver cirrhosis, chronic pancreatitis, hypertension, hyperlipidemia, hyperthyroidism, Cushing's syndrome, primary aldosteronism, pancreatic cancer, liver cancer, pheochromocytoma, hemochromatosis, schizophrenia, mood disorder, neurotic disorder, or cancer other than liver and pancreatic

<sup>f</sup> Beta-blockers, thiazide diuretics, antidepressants, corticosteroids, interferon prepared, high-calorie transfusion agents, or immunosuppressants

<sup>g</sup> Chronic kidney disease, heart failure, acute myocardial infarction, hypertension, cerebrovascular diseases, diabetes mellitus, or peripheral vascular disease

<sup>h</sup> Antiepileptic drugs, fibrates, ezetimibe, anti-gout preparations, antithyroid agent, non-steroidal anti-inflammatory drug, antifungal drugs, antituberculosis agents, or therapeutic agents for chronic hepatitis B or C



Target test item with missing	Missing method	PS weighting		PS matching		Regression adjustment	
		N in model	aHR (95%CI)	N in model	aHR (95%CI)	N in model	aHR (95%CI)
Blood glucose	Exclusion	3,430	0.52 (0.34 - 0.81)	1,880	0.92 (0.62 - 1.37)	3,430	0.60 (0.41 - 0.87)
	CC method	2,990	0.78 (0.51 - 1.21)	1,550	0.88 (0.54 - 1.44)	2,990	0.78 (0.54 - 1.11)
	SRI	3,430	0.61 (0.39 - 0.96)	1,852	1.10 (0.72 - 1.66)	3,430	0.70 (0.49 - 1.01)
	MI	3,430	0.63 (0.42 - 0.93)	1,852†	1.01 (0.34 - 2.97)	3,430	0.69 (0.48 - 1.00)
	IPW method	2,990	0.76 (0.46 - 1.25)	1,568	1.31 (0.78 - 2.19)	2,990	0.77 (0.54 - 1.11)

**Fig. 2** Scenario1 (the risk of diabetes associated with SGA compared to FGA use). Hazard ratios from outcome models with and without baseline blood glucose (SRI, MI, and IPW method were applied by hospital-specific-approach). Abbreviations: aHR, adjusted hazard ratio; CC, complete cases; CI, confidence interval; IPW, inverse probability weighted; MI, multiple imputation; PS, propensity score; SRI, single regression imputation. †The sample size in PS matching with MI gives the mean of 10 matched samples

Target test item with missing	Missing method	Approach	PS weighting		PS matching		Regression adjustment	
			N in model	aHR (95%CI)	N in model	aHR (95%CI)	N in model	aHR (95%CI)
Blood glucose	SRI	Hospital-specific	3,430	0.61 (0.39 - 0.96)	1,852	1.10 (0.72 - 1.66)	3,430	0.70 (0.49 - 1.01)
		Overall	3,430	0.60 (0.38 - 0.94)	1,864	1.00 (0.67 - 1.50)	3,430	0.70 (0.49 - 1.00)
	MI	Hospital-specific	3,430	0.63 (0.42 - 0.93)	1,852†	1.01 (0.34 - 2.97)	3,430	0.69 (0.48 - 1.00)
		Overall	3,430	0.61 (0.41 - 0.90)	1,852†	0.94 (0.28 - 3.17)	3,430	0.69 (0.48 - 0.99)
	IPW method	Hospital-specific	2,990	0.76 (0.46 - 1.25)	1,568	1.31 (0.78 - 2.19)	2,990	0.77 (0.54 - 1.11)
		Overall	2,990	0.70 (0.41 - 1.20)	1,560	0.94 (0.59 - 1.51)	2,990	0.74 (0.52 - 1.07)

**Fig. 3** Scenario1 (the risk of diabetes associated with SGA use compared to FGA use). The difference in hazard ratios between approaches considering hospital variations. Abbreviations: aHR, adjusted hazard ratio; CI, confidence interval; IPW, inverse probability weighted; MI, multiple imputation; PS, propensity score; SRI, single regression imputation. †The sample size in PS matching with MI gives the mean of 10 matched samples

Target test item with missing	Missing method	PS weighting		PS matching		Regression adjustment	
		N in model	aHR (95%CI)	N in model	aHR (95%CI) †	N in model	aHR (95%CI)
ALT	Exclusion	4,119	1.32 (0.83 - 2.08)	3,056	1.78 (0.98 - 3.24)	4,119	1.25 (0.80 - 1.96)
	CC method	3,761	1.29 (0.79 - 2.09)	2,772	1.45 (0.84 - 2.53)	3,761	1.27 (0.79 - 2.03)
	SRI	4,119	1.30 (0.82 - 2.05)	3,050	1.17 (0.68 - 2.02)	4,119	1.22 (0.78 - 1.90)
	MI	4,119	1.25 (0.87 - 1.79)	3,052‡	2.06 (0.36 - 11.69)	4,119	1.22 (0.78 - 1.92)
	IPW method	3,761	1.21 (0.73 - 2.00)	2,772	1.45 (0.84 - 2.53)	3,761	1.17 (0.72 - 1.90)
ALP	CC method	2,996	1.33 (0.76 - 2.32)	2,230	3.17 (1.52 - 6.58)	2,996	1.48 (0.88 - 2.50)
	SRI	4,119	1.32 (0.84 - 2.09)	3,060	2.25 (1.20 - 4.23)	4,119	1.28 (0.83 - 1.98)
	MI	4,119	1.27 (0.89 - 1.83)	3,057‡	2.38 (0.52 - 10.83)	4,119	1.27 (0.81 - 2.01)
	IPW method	2,996	0.98 (0.49 - 1.95)	2,268	7.50 (2.02 - 27.89)	2,996	1.12 (0.60 - 2.09)
	LDL-chol	CC method	3,176	1.08 (0.66 - 1.77)	2,164	0.89 (0.45 - 1.75)	3,176
TG	SRI	4,119	1.24 (0.78 - 1.96)	3,018	1.15 (0.68 - 1.95)	4,119	1.23 (0.78 - 1.92)
	MI	4,119	1.20 (0.84 - 1.71)	3,015‡	1.93 (0.31 - 11.9)	4,119	1.23 (0.78 - 1.93)
	IPW method	3,176	1.05 (0.63 - 1.75)	2,182	0.86 (0.50 - 1.48)	3,176	1.01 (0.61 - 1.66)
	CC method	3,306	1.25 (0.76 - 2.06)	2,372	2.00 (0.90 - 4.45)	3,306	1.12 (0.69 - 1.82)
	SRI	4,119	1.31 (0.83 - 2.06)	3,048	1.58 (0.94 - 2.67)	4,119	1.26 (0.80 - 1.96)
TG	MI	4,119	1.25 (0.88 - 1.80)	3,048‡	2.38 (0.59 - 9.67)	4,119	1.26 (0.80 - 1.97)
	IPW method	3,306	1.13 (0.67 - 1.91)	2,368	1.87 (0.99 - 3.54)	3,306	1.06 (0.64 - 1.76)

**Fig. 4** Scenario2 (the risk of hepatic injury associated with rosuvastatin use compared to atorvastatin use). Hazard ratios from outcome models with and without baseline ALT, ALP, LDL-chol, or TG (SRI, MI, and IPW method were applied by hospital-specific-approach). Abbreviations: aHR, adjusted hazard ratio; ALT, alanine transaminase; ALP, alkaline phosphatase; CC, complete cases; CI, confidence interval; IPW, inverse probability weighted; LDL-chol, low-density lipoprotein cholesterol; MI, multiple imputation; TG, triglyceride; PS, propensity score; SRI, single regression imputation. † Not shown in the forest plot if the aHR is 8.0 or more. ‡The sample size in PS matching with MI gives the mean of 10 matched samples

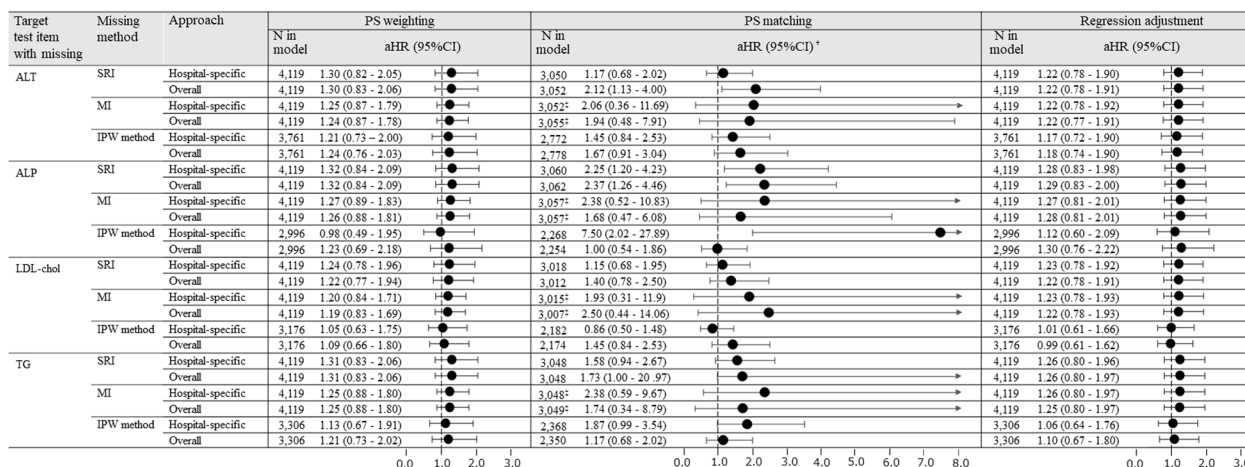
**Missing methods**

**SRI versus MI**

Although the setting is different, Marshall et al. [35] reported that the bias in SRI and MI, using the same covariates in imputation models, is the same when the missing proportion of a continuous variable is around 25%, and that there was no substantial difference in bias even if the missing proportion increased to 50%. In this study, we had no difference in the point estimates of

aHRs between SRI and MI. This might be due to the fact that we used the same covariates in imputation models.

However, for 95% CIs, we found some differences between SRI and MI in confounding adjustment with PS methods (PS weighting of Scenarios 1 and 2 and PS matching of Scenario 1). This will be further explained in **Confounding adjustment methods** section. Although SRI is generally described as underestimating SE [36], the reason that this study did not show a clear underestimation of SE regardless of the type of confounding



**Fig. 5** Scenario2 (the risk of hepatic injury associated with rosuvastatin use compared to atorvastatin use). The difference in hazard ratios between approaches considering the hospital variations. Abbreviations: aHR, adjusted hazard ratio; ALT, alanine transaminase; ALP, alkaline phosphatase; CI, confidence interval; IPW, inverse probability weighted; LDL-cholesterol, low-density lipoprotein cholesterol; MI, multiple imputation; TG, triglyceride; PS, propensity score; SRI, single regression imputation. † Not shown in the forest plot if the aHR is 8. or more. ‡The sample size in PS matching with MI gives the mean of 10 matched samples

adjustment method might be due to the targeted single laboratory test item.

**IPW method versus imputation methods (SRI and MI)**

The degree of difference in the point estimates of aHRs between the IPW method and the imputation methods varied depending on laboratory test items. In particular, there were large differences in Scenarios 1 and 2 adjusted by blood glucose level and ALP, respectively. This might be attributed to the large weights in the IPW method. In the analyses of blood glucose level and ALP, some hospital cohorts had IPWs of 20 or higher. Large weights can contribute to estimation instability. Therefore, it is important to check the stability of the effect estimation using the truncation of large weights [23] as a sensitivity analysis.

**CC and exclusion methods**

In the CC method, large differences in aHRs from those in the MAR-based methods were observed, especially in Scenario 1 adjusted by blood glucose level and in Scenario 2 adjusted by LDL-cholesterol. When the standardized mean differences (SMD) [37] between complete cases and overall cohorts were calculated for patient background factors, factors with SMD of 0.1 or higher existed in Scenarios 1 and 2 (complete cases for LDL-cholesterol) (Supplementary Tables S3 and S5). Such factors were considered to have non-negligible differences in patient backgrounds [38, 39], and the aHR difference between the CC method and the missing data methods was due to the fact that complete cases did not represent overall cohorts.

In the method excluding the laboratory test item in Scenario 1, a relatively large difference in aHR from those in the MAR-based methods occurred. However, in some Scenario 2 cases, the difference was relatively small and no noticeable difference in the analysis for ALT existed. This might be attributed to the fact that confounding by ALT was small because mean ALTs were the same for rosuvastatin and atorvastatin users.

**Assumption of missing data mechanism**

We considered the missing data mechanism in this study was a mixture of MAR and MNAR. Applying the missing data method based on the MNAR assumption requires extensive modeling of the missing data process. In such a case, the MAR-based method may be used as the main analysis method. Then, to evaluate the stability of the main results with the MAR assumption, the pre-planned sensitivity analysis should be considered [40].

When applying the MAR-based methods, the validity of the MAR assumption should be relatively increased by considering as many factors affecting the missing data as possible increases [41, 42]. In this study, all patient background characteristics were included in the missing data model. However, other missing data sources, such as the settings in the ordering system for laboratory tests, the measurement policy of tests, and the preferences of doctors were unobserved and could not be included in the missing data models.

**Approaches considering hospital variations**

The results of SRI and MI, and some of those in the IPW method showed no noticeable change in aHR due to

different approaches. In the overall approach, hospitals were included in the missing data model as a fixed variable and were expected to capture some of the effects of unobserved missing data sources at the hospital level. In the hospital-specific approach, interactions between patient backgrounds and hospitals can be considered. It is not known whether the hospital effect can be explained by the fixed effect. Additionally, the presence or absence of interaction has not been evaluated, but the results of this study indicate that the difference in approach may not significantly affect the difference in aHRs.

Some changes in aHRs due to the difference in approach in the IPW method may have been caused by the change in the distribution of IPW. In the analysis for ALP in Scenario 2 of the hospital-specific-approach, unlike the overall-approach, some patients with an IPW of 20 or higher existed.

### Confounding adjustment methods

Confounding adjustment methods affected the results in two ways. First, in PS methods (especially PS matching), there was a difference in 95% CIs between SRI and MI. For binary outcomes, Granger et al. [43] reported that coverage of 95% CI was too high in the PS matching compared to the PS weighting using SMRW due to the overestimation of variances. A combination of the PS matching and MI with the “within approach” may have affected the difference in 95% CIs in this study as well.

Second, in the analysis using PS matching, the degree of variation in aHRs and the ranges of 95% CI between missing data methods were large. One of the reasons was attributed to the decrease in the number of patients to be analyzed (Figs. 2 and 3). In Scenario 2, the incidence rate was low (Supplementary Table S5), and the impact of decreasing sample size was large. PS matching should be used carefully, considering the results from this study and its features (e.g., the target population will no longer be an exposed population if all patients in the exposed population do not have matched controls).

In scenario 1, unlike a previous study [44], the increased risk of diabetes associated with SGA use was not observed. In our study, although we did not confirm the frequency of laboratory measurements during the follow-up period, there were differences between groups in the missing proportion before drug prescription (SGA users: 20.6%, FGA users: 9.2%), suggesting that a detection bias may have existed and affected the results. In scenario 2, the point estimates of aHR were around or over 1, consistent with previous studies [32, 33].

When selecting scenarios and laboratory test items, we also considered the feasibility of the number of events; thus, it was not covered in this study, but the following case can also be a typical scenario under which missing laboratory values occur in the MID-NET<sup>®</sup> based on the knowledge from our previous study [9]; the scenario using laboratory tests with restrictions on the implementation interval based on the health insurance system (e.g., measurement can only be performed once every 3 months) that may be unique to the Japanese medical environment. It is important to considering the impact and characteristics of missing data is important when planning research.

There are two main limitations to this study. The first is that the scenarios and laboratory test items were limited. Since we used five missing methods, two approaches for considering hospital variations, and three confounder adjustment methods, we had to limit the scenarios and the number of laboratory test items. As more than one laboratory test item can be a confounding factor, further studies focusing on such situations are needed. The second is that we only examined the Tokushukai Database. When applying the missing data method to the entire MID-NET<sup>®</sup>, one must refer to the findings of this study considering the differences from the Tokushukai Database regarding the target population and missing data sources.

**Table 4** Main findings in this study

#	Brief description
1	Point estimates for aHR can be similar for SRI and MI, but there may be differences in 95% CI when adjusted with PS methods (especially PS matching)
2	In the IPW method and imputation methods, aHR may be similar, but differences can occur.
3	In the CC method and the method excluding the laboratory test item, aHR may be similar to that of the MAR-based method, but differences can occur.
4	It can be pointed out that the difference in approaches considering hospital variations may not significantly affect the difference in aHR.
5	It can be pointed out that the difference in confounding adjustment methods can be different for aHR, especially between PS matching and other methods.

*Abbreviations:* aHR Adjusted hazard ratio, CC Complete case, CI Confidence interval, IPW Inverse probability weighted, MAR Missing at random, MI Multiple imputation, PS Propensity score, SRI Single regression imputation

## Conclusions

Based on the five main findings of this study (Table 4), we concluded that, although the different missing methods may contribute to differences in parameter estimates of the outcome model, SRI and MI can provide similar point estimates, and two approaches considering hospital variations do not have a major impact on the results. Confounding adjustment by PS matching gave unstable point estimates and wide confidence intervals and should therefore be used carefully. Although we report findings based on a case study and cannot draw generalizable recommendation, our research results may help in the selection of missing data imputation methods and the interpretation of obtained results in the future utilization of MID-NET®.

## Abbreviations

aHR	Adjusted hazard ratio
ALP	Alkaline phosphatase
ALT	Alanine transaminase
CC	Complete case
CI	Confidence interval
EMR	Electronic medical record
FGA	First-generation antipsychotic
HbA1c	Hemoglobin A1c
ICD	International Classification of Diseases
IPW	Inverse probability weighted
JDS	The Japan Diabetes Society
LDL-chol	Low-density lipoprotein cholesterol
MAR	Missing at random
MCAR	Missing completely at random
MNAR	Missing not at random
MI	Multiple imputation
NGSP	The National Glycohemoglobin Standardization Program
PS	Propensity score
SGA	Second-generation antipsychotic
SMD	Standardized mean differences
SMRW	Standardized mortality ratio weighting
SRI	Single regression imputation
TG	Triglyceride

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12911-023-02345-7>.

**Additional file 1: Supplementary Table S1.** List of the 10 hospitals in the database system for MID-NET® collaborative organizations of Tokushukai Medical Group. **Supplementary Table S2.** Data items used in this study. **Supplementary Table S3.** Scenario 1. Patient backgrounds among the cohort, complete cases, and cases with missing data. **Supplementary Table S4.** Scenario 1. Patient backgrounds by hospital. **Supplementary Table S5.** Scenario 2. Patient backgrounds among the cohort, complete cases, and cases with missing data. **Supplementary Table S6.** Scenario 2. Patient backgrounds by hospital. **Supplementary Figure S1.** Sequence of steps from handling missing data to statistical analysis. **Supplementary Figure S2.** Study scenario selection flowchart. **Supplementary Figure S3.** Scenario 1. Number of patients in the study cohort: risk of diabetes associated with SGA. **Supplementary Figure S4.** Scenario 2. Number of patients in the study cohort: risk of hepatic injury associated with rosuvastatin.

## Acknowledgements

We thank Dr. Masaaki Doi, Dr. Yoshiaki Uyama, and Tokushukai Medical Group for their valuable assistance. Views expressed here are those of the authors and do not necessarily represent the official views and findings of the Pharmaceuticals and Medical Devices Agency. This research was funded by the Agency for Medical Research and Development (AMED) under Grant Number JP211k0201702.

## Authors' contributions

MK, MO, and TS conceptualized the study. MK analyzed the data. MK wrote the initial draft of the manuscript. MK, YF, MO, and TS contributed to the interpretation of findings and manuscript revisions. All authors have read and approved the final version of the manuscript.

## Funding

This research was supported by the Japan Agency for Medical Research and Development (AMED) under Grant Number JP211k0201702. The funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

## Availability of data and materials

The datasets generated and/or analyzed during the current study are not publicly available due to the terms of use for MID-NET® to which we adhered when conducting this study; the accessibility of the dataset used for this analysis is restricted to specific authors including the corresponding author in a predetermined secure environment. No outside researchers are allowed to access the dataset. This study used the database system for MID-NET®-collaborative organizations of the Tokushukai Medical Group, a part of MID-NET®, and not the entire MID-NET®. However, we followed the terms of use for MID-NET®, as the datasets were included in the entire MID-NET®.

## Declarations

### Ethics approval and consent to participate

The study was conducted in accordance with the Declaration of Helsinki and approved by the Kyoto University Graduate School and Faculty of Medicine Kyoto University Hospital Ethics Committee in November 2018 (R1793). Based on Japanese Ethical Guidelines for Medical and Health Research Involving Human Subjects, the need for informed consent from individual patients was waived by the Kyoto University Graduate School and Faculty of Medicine Kyoto University Hospital Ethics Committee. The use of extracted data for this study from the database system for MID-NET®-collaborative organizations of Tokushukai Medical Group was approved by the administrative board of General Incorporated Association Tokushukai, which is the data holder. The data used in this study were anonymized before their use (refer to the previous study by Yamaguchi et al. [2] for details of the data anonymization).

### Consent for publication

Not applicable.

### Competing interests

Maki Komamine is employed by the Pharmaceuticals and Medical Devices Agency and has no financial or personal relationships with other people or organizations that could inappropriately influence or bias the contents of this paper. Other authors have no financial or personal relationships with other people or organizations that could inappropriately influence or bias the contents of this paper.

### Author details

<sup>1</sup>Department of Biostatistics, Kyoto University School of Public Health, Yoshida-konocho, Sakyo-ku, Kyoto 606-8501, Japan. <sup>2</sup>Office of Medical Informatics and Epidemiology, Pharmaceuticals and Medical Devices Agency, Tokyo, Japan. <sup>3</sup>Head Office, Tokushukai Information System Incorporated, Osaka, Japan. <sup>4</sup>Department of Clinical Biostatistics, Graduate School of Medicine, Kyoto University, Kyoto, Japan.

Received: 30 November 2022 Accepted: 19 October 2023  
Published online: 30 October 2023

## References

- Yamada K, Itoh M, Fujimura Y, et al. The utilization and challenges of Japan's MID-NET<sup>®</sup> medical information database network in postmarketing drug safety assessments: a summary of pilot pharmacoepidemiological studies. *Pharmacoepidemiol Drug Saf.* 2019;28(5):601–8.
- Yamaguchi M, Inomata S, Harada S, et al. Establishment of the MID-NET<sup>®</sup> medical information database network as a reliable and valuable database for drug safety assessments in Japan. *Pharmacoepidemiol Drug Saf.* 2019;28(10):1395–404.
- Pharmaceuticals and Medical Devices Agency. Summary of MID-NET<sup>®</sup> study: no. 2018-001. 2020. <https://www.pmda.go.jp/files/000233987.pdf>. Accessed 7 May 2022.
- Pharmaceuticals and Medical Devices Agency. Summary of MID-NET<sup>®</sup> study: no. 2018-002. 2020. <https://www.pmda.go.jp/files/000234446.pdf>. Accessed 7 May 2022.
- Schneeweiss S, Rassen JA, Glynn RJ, et al. Supplementing claims data with outpatient laboratory test results to improve confounding adjustment in effectiveness studies of lipid-lowering treatments. *BMC Med Res Methodol.* 2012;12:180.
- Raebel MA, Shetterly S, Lu CY, et al. Methods for using clinical laboratory test results as baseline confounders in multi-site observational database studies when missing data are expected. *Pharmacoepidemiol Drug Saf.* 2016;25(7):798–814.
- Wells BJ, Chagin KM, Nowacki AS, Kattan MW. Strategies for handling missing data in electronic health record-derived data. *EGEMS (Wash DC).* 2013;1(3):1035.
- Eekhout I, de Boer RM, Twisk JW, de Vet HC, Heymans MW. Missing data: a systematic review of how they are reported and handled. *Epidemiology.* 2012;23(5):729–32.
- Komamine M, Fujimura Y, Nitta Y, Omiya M, Doi M, Sato T. Characteristics of hospital differences in missing of clinical laboratory test results in a multi-hospital observational database contributing to MID-NET<sup>®</sup> in Japan. *BMC Med Inform Decis Mak.* 2021;21(1):181.
- Molenberghs G, Fitzmaurice G, Kenward MG, Tsiatis A, Verbeke G. *Handbook of missing data methodology.* Boca Raton: CRC Press; 2014.
- Herring AH, Ibrahim JG. Likelihood-based methods for missing covariates in the Cox proportional hazards model. *J Am Stat Assoc.* 2001;96(453):292–302.
- Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. *J Clin Epidemiol.* 2006;59(10):1087–91.
- Perkins NJ, Cole SR, Harel O, et al. Principled approaches to missing data in epidemiologic studies. *Am J Epidemiol.* 2018;187(3):568–75.
- Dziura JD, Post LA, Zhao Q, Fu Z, Peduzzi P. Strategies for dealing with missing data in clinical trials: from design to analysis. *Yale J Biol Med.* 2013;86(3):343–58.
- Little RJA, Rubin DB. *Statistical analysis with missing data.* 2nd ed. Hoboken: Wiley; 2002.
- Baraldi AN, Enders CK. An introduction to modern missing data analyses. *J Sch Psychol.* 2010;48(1):5–37.
- Rubin DB. Inference and missing data. *Biometrika.* 1976;63:581–92.
- Rubin DB. Multiple imputation for nonresponse in surveys. New York: Wiley; 1987.
- Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prev Sci.* 2007;8(3):206–13.
- Bounthavong M, Watanabe JH, Sullivan KM. Approach to addressing missing data for electronic medical records and pharmacy claims data research. *Pharmacotherapy.* 2015;35(4):380–7.
- van Buuren S. *Flexible imputation of missing data.* CRC Press; 2018.
- Wang CY, Chen HY. Augmented inverse probability weighted estimator for Cox missing covariate regression. *Biometrics.* 2001;57(2):414–9.
- Seaman SR, White IR. Review of inverse probability weighting for dealing with missing data. *Stat Methods Med Res.* 2013;22(3):278–95.
- Von Elm E, Altman DG, Egger M, et al. The strengthening of reporting of observational studies in epidemiology (STROBE) Statement: guidelines for reporting observational studies. *Ann Intern Med.* 2007;147(8):573–7.
- White IR, Royston P. Imputing missing covariate values for the Cox model. *Stat Med.* 2009;28(15):1982–98.
- Xu Q, Paik MC, Rundek T, Elkind MS, Sacco RL. Reweighting estimators for Cox regression with missing covariate data: analysis of insulin resistance and risk of stroke in the Northern Manhattan Study. *Stat Med.* 2011;30(28):3328–40.
- Sato T, Matsuyama Y. Marginal structural models as a tool for standardization. *Epidemiology.* 2003;14(6):680–6.
- Newcomer JW, Haupt DW, Fucetola R, et al. Abnormalities in glucose regulation during antipsychotic treatment of schizophrenia. *Arch Gen Psychiatry.* 2002;59(4):337–45.
- Lindenmayer JP, Nathan AM, Smith RC. Hyperglycemia associated with the use of atypical antipsychotics. *J Clin Psychiatry.* 2001;62Suppl23:30–8.
- Crestor (rosuvastatin) [package insert]. Osaka: AstraZeneca K.K.; 2022. [https://www.pmda.go.jp/PmdaSearch/IyakuDetail/ResultDataSetPDF/670227\\_2189017F1022\\_1\\_28](https://www.pmda.go.jp/PmdaSearch/IyakuDetail/ResultDataSetPDF/670227_2189017F1022_1_28). Accessed 7 May 2022.
- Lipitor (atorvastatin) [package insert]. Tokyo: Viatris Inc; 2021. [https://www.pmda.go.jp/PmdaSearch/IyakuDetail/ResultDataSetPDF/671450\\_2189015F1023\\_2\\_01](https://www.pmda.go.jp/PmdaSearch/IyakuDetail/ResultDataSetPDF/671450_2189015F1023_2_01). Accessed 7 May 2022.
- Clarke AT, Johnson PC, Hall GC, Ford J, Mills PR. High dose atorvastatin associated with increased risk of significant hepatotoxicity in comparison to simvastatin in UK GPRD cohort. *PLoS One.* 2016;11(3):e0151587.
- Chang CH, Chang YC, Lee YC, Liu YC, Chuang LM, Lin JW. Severe hepatic injury associated with different statins in patients with chronic liver disease: a nationwide population-based cohort study. *J Gastroenterol Hepatol.* 2015;30(1):155–62.
- Leyrat C, Seaman SR, White IR, et al. Propensity score analysis with partially observed covariates: how should multiple imputation be used? *Stat Methods Med Res.* 2019;28(1):3–19.
- Marshall A, Altman DG, Holder RL. Comparison of imputation methods for handling missing covariate data when fitting a Cox proportional hazards model: a resampling study. *BMC Med Res Methodol.* 2010;10:112.
- Little RJA, Rubin DB. *Statistical analysis with missing data.* Wiley; 1987, pp.44–47.
- Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med.* 2009;28(25):3083–107.
- Normand ST, Landrum MB, Guadagnoli E, et al. Validating recommendations for coronary angiography following an acute myocardial infarction in the elderly: a matched analysis using propensity scores. *J Clin Epidemiol.* 2001;54(4):387–98.
- Cappabianca G, Mariscalco G, Biancari F, et al. Safety and efficacy of prothrombin complex concentrate as first-line treatment in bleeding after cardiac surgery. *Crit Care.* 2016;20:5.
- E9 (R1): addendum to statistical principles for clinical trials on choosing appropriate estimands and defining sensitivity analyses in clinical trials. [https://database.ich.org/sites/default/files/E9-R1\\_Step4\\_Guideline\\_2019\\_1203.pdf](https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf). Accessed 7 May 2022.
- Beaulieu-Jones BK, Lavage DR, Snyder JW, Moore JH, Pendergrass SA, Bauer CR. Characterizing and managing missing structured data in electronic health records: data analysis. *JMIR Med Inform.* 2018;6(1):e11.
- van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. *Stat Med.* 1999;18(6):681–94.
- Granger E, Sergeant JC, Lunt M. Avoiding pitfalls when combining multiple imputation and propensity scores. *Stat Med.* 2019;38(26):5120–32.
- Smith M, Hopkins D, Peveler RC, Holt RI, Woodward M, Ismail K. First- v. second-generation antipsychotics and risk for diabetes in schizophrenia: systematic review and meta-analysis. *Br J Psychiatry.* 2008;192(6):406–11.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Supplementary Information

- File name: Supplement.pdf
- File format: PDF
- Title of data: Supplementary Tables and Figures
- Description of data:
  - Supplementary Table S1. List of the 10 hospitals in the database system for MID-NET® collaborative organizations of Tokushukai Medical Group
  - Supplementary Table S2. Data items used in this study
  - Supplementary Table S3. Scenario 1. Patient backgrounds among the cohort, complete cases, and cases with missing data
  - Supplementary Table S4. Scenario 1. Patient backgrounds by hospital
  - Supplementary Table S5. Scenario 2. Patient backgrounds among the cohort, complete cases, and cases with missing data
  - Supplementary Table S6. Scenario 2. Patient backgrounds by hospital
  - Supplementary Figure S1. Sequence of steps from handling missing data to statistical analysis
  - Supplementary Figure S2. Study scenario selection flowchart
  - Supplementary Figure S3. Scenario 1. Number of patients in the study cohort: risk of diabetes associated with SGA
  - Supplementary Figure S4. Scenario 2. Number of patients in the study cohort: risk of hepatic injury associated with rosuvastatin

## SUPPLEMENTARY TABLES

### Supplementary Table S1. List of the 10 hospitals† in the database system for MID-NET® collaborative organizations of Tokushukai Medical Group

Uji Tokushukai Hospital
Kishiwada Tokushukai Hospital
Sapporo Tokushukai Hospital
Shonan Fujisawa Tokushukai Hospital
Tokyo-Nishi Tokushukai Hospital
Nagoya Tokushukai General Hospital
Nozaki Tokushukai Hospital
Fukuoka Tokushukai Hospital
Matsubara Tokushukai Hospital
Yao Tokushukai General Hospital

MID-NET® is a distributed closed network system in which a database system is built for each collaborative organization. The 10 collaborative organizations include three group hospitals and seven national university hospitals. The data of collaborative hospitals of the group is collected in the database system of the group hospital.

† These are diverse in terms of scale and regionality, but they are all general hospitals with an emergency room, providing core medical care for each region

### Supplementary Table S2. Data items used in this study

Data sources	Data items
EMR data	Patient identifiers
	Diagnostic orders data
	Discharge summary data
	Prescription orders/execution data
	Injection orders/execution data
Laboratory test data	
Claim data	Medical procedure data†

† Used to confirm whether it was their first visit and emergency care.

**Supplementary Table S3. Scenario 1. Patient backgrounds among the cohort, complete cases, and cases with missing data**

	Cohort			CC	Missing
	FGA N=2,343	SGA N=1,087	Overall N=3,430	N=2,990	N=440
<b>Number of events</b>	131	55	186	179	NA
<b>Total follow-up periods (person-years)</b>	90.1	159.5	249.6	138.0	111.6
<b>Incidence rate(/one person-year)</b>	1.45	0.34	0.75	1.30	NA
<b>Sex, male<sup>†</sup> (%)</b>	61.2	53.2	58.7	59.8	50.7
<b>Age<sup>†</sup>, mean (SD)</b>	72.5 (13.4)	68.3 (24.4)	71.2 (17.7)	73.6 (14.2)	54.8 (28.0)
<b>Year of cohort entry<sup>†</sup></b>					
2015	19.3	18.5	19.0	19.1	18.4
2016	41.6	42.6	41.9	42.2	39.5
2017	39.1	38.9	39.1	38.6	42.0
<b>Hospitalization<sup>†</sup> (%)</b>	76.7	69.5	74.4	83.3 <sup>§</sup>	13.9
<b>Class number of concomitant medications<sup>‡</sup>, mean (SD)</b>	0.8 (0.9)	0.6 (0.8)	0.7 (0.8)	0.8 (0.9)	0.3(0.6)
<b>Complications<sup>‡</sup> (%)</b>					
Hepatitis	4.7	2.6	4.0	4.5	NA
Liver cirrhosis	2.0	1.7	1.9	2.1	NA
Chronic pancreatitis	0.4	NA	0.5	0.5	0
Hypertension	11.8	11.8	11.8	12.8	4.5
Hyperlipidemia	6.5	7.2	6.7	7.4	2.5
Hyperthyroidism	2.7	3.6	3.0	3.2	NA
Cushing's syndrome	0.7	1.7	1.0	1.2	NA
Primary aldosteronism	NA	0	NA	NA	NA
Pancreatic cancer	3.4	1.2	2.7	3.0	NA
Liver cancer	6.2	2.3	5.0	5.5	NA
Pheochromocytoma	NA	0	NA	NA	NA
Hemochromatosis	NA	0	NA	NA	NA
Schizophrenia	7.3	23.0	12.2	10.0	27.5
Mood disorder	3.6	6.2	4.4	4.2	5.9
Neurotic disorder	5.7	3.9	5.1	5.1	5.5
Cancer other than liver and pancreatic	49.8	25.9	42.2	45.4	20.5
<b>Concomitant medication<sup>‡</sup> (%)</b>					
Beta-blockers	20.3	19	19.9	22.0	5.2
Thiazide diuretics	3.8	3.8	3.8	4.0	2.5
Antidepressants	6.4	12.9	8.5	7.8	13
Corticosteroids	31.6	15.9	26.6	28.6	13
Interferon prepared	NA	0	NA	NA	NA
High-calorie transfusion agents	7.4	2.2	5.7	6.6	0
Immunosuppressant	0.9	NA	0.8	0.9	NA
<b>First visit<sup>†</sup> (%)</b>	1.1	1.7	1.3	1.1	2.5



<b>Emergency care<sup>†</sup> (%)</b>	27.7	28.0	27.8	31.5	NA
---------------------------------------	------	------	------	------	----

---

Some categories in which the patient numbers did not exceed 10 were not shown due to privacy.

NA is the case when the number of patients was less than 10.

There was no missing data for sex, age, and year of cohort entry.

Abbreviations: CC, complete cases; FGA, first-generation antipsychotic; SGA, second-generation antipsychotic.

† At the date of the first prescription of any antidiabetic drug. "First visit" and "emergency care" were defined by whether the associated codes were recorded on the same day as the first prescription of the drug of interest.

‡ 180 days prior to the date of the first prescription of any antidiabetic drug.

§ Standardized differences between cohort and CC with greater than 0.1.

**Supplementary Table S4. Scenario 1. Patient backgrounds by hospital**

	Hospital #									
	1 N=174	2 N=121	3 N=463	4 N=505	5 N=388	6 N=531	7 N=563	8 N=208	9 N=276	10 N=201
<b>Sex, male<sup>†</sup> (%)</b>	54.0	58.7	64.8	59.2	57.7	58.2	60.0	43.3	56.5	65.2
<b>Age<sup>†</sup>, mean (SD)</b>	70.7 (15.5)	75.9 (11.1)	71.3 (13.9)	70.2 (19.5)	70.1 (17.7)	71.4 (15.4)	74.2 (12.6)	75.1 (17.3)	76.1 (13.6)	52.8 (32.0)
<b>Year of cohort entry<sup>†</sup></b>										
2015	16.7	19.0	19.7	19.8	17.5	18.5	19.9	25.5	17.8	14.9
2016	35.6	32.2	43.2	46.7	40.7	42.6	43.0	35.1	44.6	38.8
2017	47.7	48.8	37.1	33.5	41.8	39.0	37.1	39.4	37.7	46.3
<b>Hospitalization<sup>†</sup> (%)</b>	66.7	75.2	79.5	72.5	73.7	81.4	66.3	77.4	92.8	51.7
<b>Class number of concomitant medications<sup>‡</sup>, mean (SD)</b>					0.7 (0.8)					
<b>Exposure (%)</b>										
FGAs	72.4	74.4	90.5	65.5	54	68.4	73.5	66.8	55.4	48.8
SGAs	27.6	25.6	9.5	34.5	45.9	31.6	26.5	33.2	44.6	51.2
<b>Baseline result available (%)</b>										
Blood glucose	86.2	86.8	94.8	86.7	87.9	92.3	84.2	89.4	90.2	58.7
<b>Complications<sup>‡</sup> (%)</b>										
Hepatitis	NA	NA	5.4	5.1	4.1	3.2	2.1	4.8	5.4	NA
Liver cirrhosis	NA	NA	NA	2.6	3.6	1.9	NA	NA	NA	NA
Chronic pancreatitis	NA	0.0	NA	NA	0.0	NA	NA	NA	NA	NA
Hypertension	11.5	13.2	12.7	11.5	8.0	14.5	12.1	12.5	11.2	9.0
Hyperlipidemia	NA	NA	4.8	4.0	8.0	5.8	12.1	6.3	7.2	6.0
Hyperthyroidism	NA	NA	NA	2.4	2.6	6.4	NA	4.8	5.1	NA
Cushing's syndrome	0.0	0.0	NA	0.0	7.5	NA	NA	0.0	0.0	0.0
Primary aldosteronism	0.0	0.0	0.0	0.0	0.0	0.0	NA	0.0	0.0	0.0
Pancreatic cancer	NA	NA	5.0	3.4	NA	2.3	2.3	NA	NA	NA

Liver cancer	7.5	NA	4.3	6.3	5.9	3.4	5.7	10.6	NA	NA
Pheochromocytoma	0.0	0.0	0.0	0.0	NA	0.0	0.0	0.0	0.0	0.0
Hemochromatosis	0.0	NA	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Schizophrenia	6.3	24.0	5.8	15.4	8.5	10.9	19.7	11.1	5.4	17.4
Mood disorder	NA	NA	4.8	3.6	7.0	4.3	4.4	8.2	NA	NA
Neurotic disorder	NA	8.3	4.5	4.2	2.8	4.0	9.8	7.7	NA	NA
Cancer other than liver and pancreatic	25.9	42.1	46.4	48.9	34.3	45.2	46.2	43.3	38.8	29.9
<b>Concomitant medication<sup>‡</sup> (%)</b>										
Beta-blockers	22.4	20.7	26.6	19.8	24.0	19.8	13.0	15.4	25.0	11.4
Thiazide diuretics	NA	NA	3.0	7.1	4.9	2.8	2.1	5.3	4.3	NA
Antidepressants	13.2	8.3	8.9	5.9	15.5	10.0	6.6	7.7	NA	7.0
Corticosteroids	18.4	21.5	32.6	31.1	18.3	29.6	28.8	21.2	29.3	15.9
Interferon prepared	NA	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
High-calorie transfusion agents	NA	9.1	7.1	9.1	4.6	6.0	3.0	7.7	4.7	NA
Immunosuppressant	0.0	0.0	NA	NA	NA	NA	NA	NA	NA	NA
<b>First visit<sup>†</sup> (%)</b>	NA	NA	NA	NA	NA	NA	1.8	NA	NA	NA
<b>Emergency care<sup>†</sup> (%)</b>	24.7	34.7	34.1	43.4	18.8	25.2	27.0	23.1	17.8	16.9

Some categories in which the patient numbers did not exceed 10 were not shown due to privacy.

NA is the case when the number of patients was less than 10.

There was no missing data for sex, age, and year of cohort entry.

Abbreviations: CC, complete cases; FGA, first-generation antipsychotic; SGA, second-generation antipsychotic.

<sup>†</sup> At the date of the first prescription of any antidiabetic drug. "First visit" and "emergency care" were defined by whether the associated codes were recorded on the same day as the first prescription of the drug of interest.

<sup>‡</sup> 180 days prior to the date of the first prescription of any antidiabetic drug.

**Supplementary Table S5. Scenario 2. Patient backgrounds among the cohort, complete cases, and cases with missing data**

	Cohort			ALT		ALP		LDL		TG	
	Atorva statin N=1,581	Rosuva statin N=2,538	Overall N=4,119	CC	Missing	CC	Missing	CC	Missing	CC	Missing
				N=3,761	N=358	N=2,996	N=1,123	N=3,176	N=943	N=3,306	N=813
<b>Number of events (n)</b>	28	71	99	90	NA	75	24	87	12	83	16
<b>Total follow-up periods (person-years)</b>	817.6	1452.7	2,270.3	2,029.0	241.3	1,574.4	695.9	1,974.4	295.9	1,969.3	301.0
<b>Incidence rate/(one person-year)</b>	0.03	0.05	0.04	0.04	0.04	0.05	0.03	0.04	0.28	0.04	0.05
<b>Sex, male<sup>†</sup> (%)</b>	52.8	56.6	55.1	56.0	45.8	56.2	52.4	57.6	46.8	57.0	47.6
<b>Age<sup>†</sup>, mean (SD)</b>	70.3 (12.3)	68.5 (12.3)	69.2 (12.3)	69.5 (12.2)	66.0 (12.7)	69.8 (12.2)	67.5 (12.3)	68.5 (12.5)	71.3 (11.5)	68.6 (12.4)	71.4 (11.6)
<b>Year of cohort entry<sup>†</sup>(%)</b>											
2015	19.7	21.2	20.6	19.9	27.7	19.9	22.4	20.5	20.9	20.1	22.5
2016	39.1	39.0	39.0	39.1	38.3	38.6	40.2	39.3	38.2	39.4	37.3
2017	41.2	39.9	40.4	41.0	34.1	41.6	37.3	40.2	40.9	40.4	40.2
<b>Hospitalization<sup>†</sup> (%)</b>	43.5	37.6	39.9	42.6	10.6	47.1 <sup>§</sup>	20.7	32.7 <sup>§</sup>	63.9	35.5	57.7
<b>Class number of concomitant medications<sup>‡</sup>, mean (SD)</b>	0.5 (0.7)	0.5 (0.7)	0.5 (0.7)	0.5 (0.7)	0.4 (0.6)	0.5 (0.7)	0.4 (0.6)	0.4 (0.7)	0.6(0.7)	0.5 (0.7)	0.5 (0.7)
<b>Complications<sup>‡</sup> (%)</b>											
Chronic kidney disease	2.8	2.9	2.8	3.1	NA	3.5	1.2	3.1	1.8	3.1	1.7
Heart failure	21.9	27.3	25.2	26.8	8.7	27.5	19.1	29.2	11.7	28.4	12.1
Acute myocardial infarction	10.2	13.4	12.1	12.7	5.9	13.8	7.7	13.6	7.3	12.9	9.1
Hypertension	21.3	19.9	20.4	20.9	15.4	22.1	16.1	19.4	24	20.1	22
Cerebrovascular diseases	9.5	12.2	11.2	11.8	4.7	12.4	7.9	12.6	6.3	12.3	6.4
Diabetes mellitus	43.8	46.8	45.7	48.0	20.7	48.0	39.4	50.8	28.3	49.7	29.2

Peripheral vascular disease	NA	0.6	0.6	0.6	0.0	0.6	NA	0.6	NA	0.6	NA
<b>Concomitant medication<sup>‡</sup> (%)</b>											
Antiepileptic drugs	0.7	0.6	0.6	0.6	NA	0.7	NA	0.7	NA	0.7	NA
Fibrates	2.5	3.7	3.2	3.2	3.9	3.1	3.7	3.7	1.6	3.6	1.6
Ezetimibe	2.5	4.3	3.6	3.7	NA	3.8	3.0	3.6	3.4	3.5	3.8
Anti-gout preparations	15.3	12.6	13.6	13.7	13.1	14.0	12.6	13.5	14	13.8	13.2
Antithyroid Agent	NA	1.0	0.7	0.7	NA	0.7	NA	0.8	NA	0.8	NA
NSAIDs	21.4	22.2	21.9	22.4	16.8	23.3	18.2	19.0	31.4	20.1	29.2
Antifungal drugs	1.4	1.8	1.7	1.8	NA	1.9	0.9	1.6	1.9	1.6	1.7
Antituberculosis agents	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Therapeutic agents for chronic hepatitis B or C	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
<b>First visit<sup>†</sup> (%)</b>	3.7	3.0	3.3	2.8	8.7	2.5	5.3	3.0	4.1	3.0	4.6
<b>Emergency care<sup>†</sup>(%)</b>	14.5	16.1	15.5	16.9	NA	19.3 <sup>§</sup>	5.3	15.3	16.0	15.5	15.3

Some categories in which the patient numbers did not exceed 10 were not shown due to privacy.

NA is the case when the number of patients was less than 10.

There was no missing data for sex, age, and year of cohort entry.

Abbreviations: ALP, alkaline phosphatase; ALT, alanine transaminase CC, complete cases; LDL-chol, low-density lipoprotein cholesterol; NSAID, non-steroidal anti-inflammatory drug; TG, triglyceride.

† At the date of the first prescription of any antidiabetic drug. "First visit" and "emergency care" were defined by whether the associated codes were recorded on the same day as the first prescription of the drug of interest.

‡ 180 days prior to the date of the first prescription of any antidiabetic drug.

§ Standardized differences between cohort and CC with greater than 0.1.

**Supplementary Table S6. Scenario 2. Patient backgrounds by hospital**

	Hospital #								
	1 N=241	3 N=529	4 N=631	5 N=518	6 N=662	7 N=661	8 N=282	9 N=281	10 N=314
<b>Sex, male<sup>†</sup> (%)</b>	51.5	59.5	53.4	57.5	56.3	55.7	45.4	55.5	54.8
<b>Age<sup>†</sup>, mean (SD)</b>	68.3 (13.2)	70.0 (10.6)	68.5 (13.0)	67.7 (12.4)	69.0 (12.0)	71.1 (12.2)	67.1 (12.7)	70.6 (12.2)	69.1 (12.5)
<b>Year of cohort entry<sup>†</sup>(%)</b>									
2015	25.7	16.8	26.1	20.7	22.4	21.2	18.8	12.8	15.3
2016	36.1	37.8	39.6	41.3	34	38.9	39.4	47.3	41.4
2017	38.2	45.4	34.2	38	43.7	39.9	41.8	39.9	43.3
<b>Hospitalization<sup>‡</sup> (%)</b>	17.0	43.1	35.3	46.9	44.3	54.0	23.0	30.2	34.1
<b>Class number of concomitant medications<sup>§</sup>, mean (SD)</b>	0.6 (0.7)	0.5 (0.7)	0.5 (0.7)	0.4 (0.6)	0.5 (0.7)	0.4 (0.6)	0.5 (0.7)	0.5 (0.7)	0.5 (0.7)
<b>Exposure (%)</b>									
Atorvastatin	33.2	36.3	32.3	31.1	51.5	38.1	46.5	36.7	37.3
Rosuvastatin	66.8	63.7	67.7	68.9	48.5	61.9	53.5	63.3	62.7
<b>Baseline result available (%)</b>									
ALT	87.6	93.6	88.4	95.0	94.3	92.9	89.7	94.0	79.6
ALP	62.7	58.0	65.5	92.1	92.1	70.8	58.9	71.5	64.6
LDL-chol	76.3	75.4	74.2	81.7	81.1	73.7	78.0	80.8	73.6
TG	73.4	90.7	71.2	89.8	80.7	74.9	79.8	86.1	76.1
<b>Complications<sup>†</sup> (%)</b>									
Chronic kidney disease	NA	3.6	NA	6.8	2.3	2.6	3.5	NA	NA
Heart failure	27.0	36.5	22.2	25.5	22.8	18.0	27.0	33.5	21.7
Acute myocardial infarction	17.4	12.3	9.2	14.3	9.5	12.0	14.2	16.0	10.8
Hypertension	18.7	20.0	21.4	17.6	22.2	24.1	12.1	23.5	18.8
Cerebrovascular diseases	12.9	9.3	13.9	12.2	14.0	13.0	NA	6.8	7.3
Diabetes mellitus	49.4	46.7	39.1	45.9	45.3	49.9	53.2	43.4	40.8

Peripheral vascular disease	NA	NA	NA	NA	NA	NA	NA	NA	NA
<b>Concomitant medication<sup>‡</sup> (%)</b>									
Antiepileptic drugs	NA	NA	NA	NA	NA	NA	NA	NA	NA
Fibrates	4.1	3.4	2.4	NA	2.7	3.0	5.0	4.3	6.7
Ezetimibe	NA	2.3	2.7	8.5	2.3	2.9	NA	5.3	3.2
Anti-gout preparations	14.9	13.8	14.1	9.5	13.0	13.9	14.9	12.8	18.8
Antithyroid Agent	NA	NA	NA	NA	NA	NA	NA	NA	NA
NSAIDs	29.9	25.1	26.0	16.8	22.7	15.7	24.5	23.1	18.2
Antifungal drugs	NA	NA	1.7	NA	1.8	3.0	NA	NA	NA
Antituberculosis agents	NA	NA	NA	NA	NA	NA	NA	NA	NA
Therapeutic agents for chronic hepatitis B or C	NA	NA	NA	NA	NA	NA	NA	NA	NA
<b>First visit<sup>†</sup> (%)</b>	5.0	NA	2.4	2.3	3.8	2.7	7.4	6.8	3.2
<b>Emergency care<sup>†</sup>(%)</b>	6.2	19.1	20.6	12.4	17.8	20.0	7.8	7.5	11.1

Some categories in which the patient numbers did not exceed 10 were not shown due to privacy.

NA is the case when the number of patients was less than 10.

There was no missing data for sex, age, and year of cohort entry.

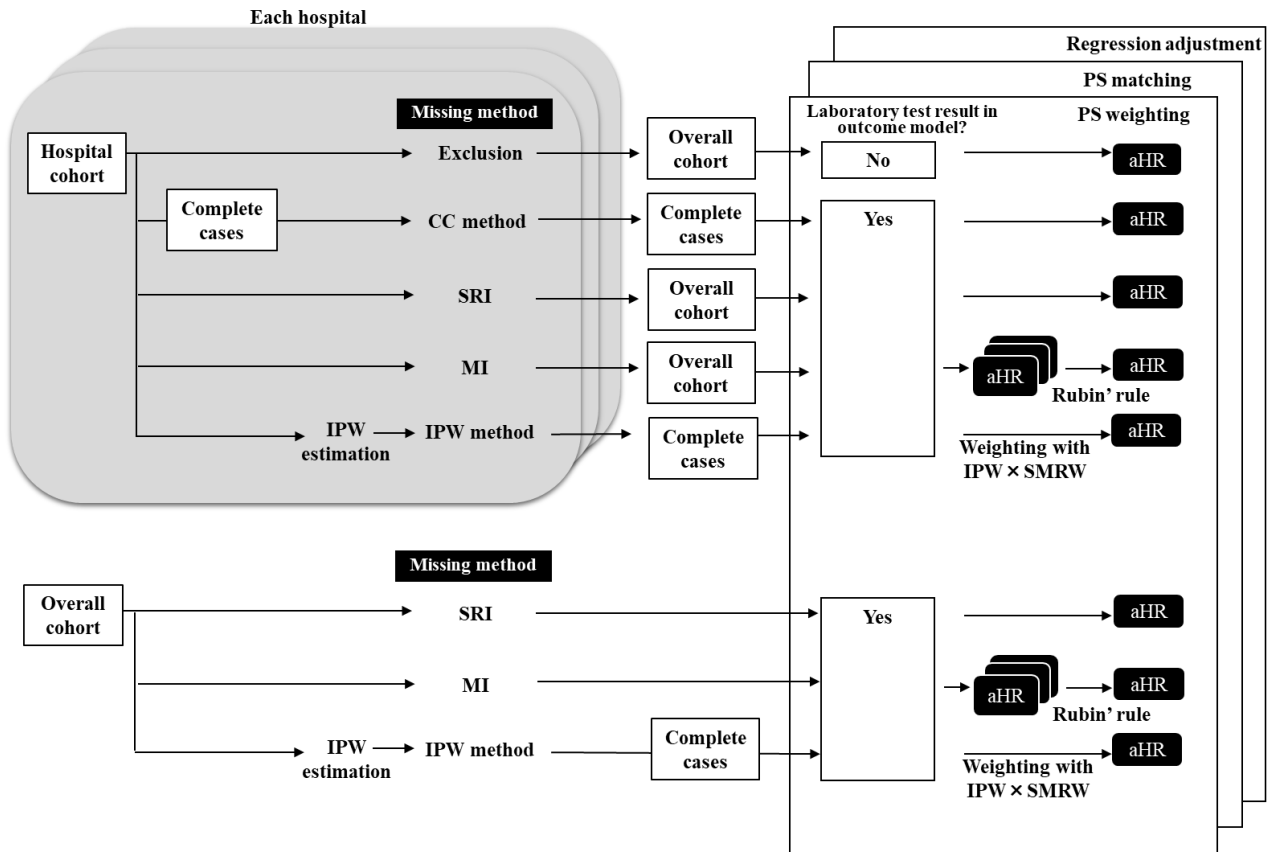
Abbreviations: ALP, alkaline phosphatase; ALT, alanine transaminase CC, complete cases; LDL-chol, low-density lipoprotein cholesterol; NSAID, non-steroidal anti-inflammatory drug;

TG, triglyceride.

†At the date of the first prescription of any antidiabetic drug. "First visit" and "emergency care" were defined by whether the associated codes were recorded on the same day as the first prescription of the drug of interest.

‡ 180 days prior to the date of the first prescription of any antidiabetic drug.

## SUPPLEMENTARY FIGURES

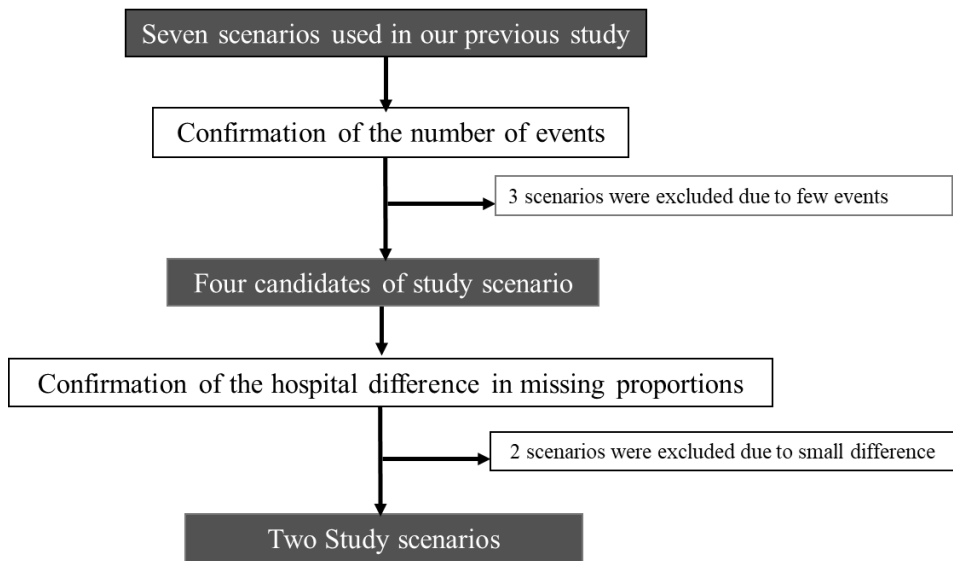


**Supplementary Figure S1. Sequence of steps from handling missing data to statistical analysis**

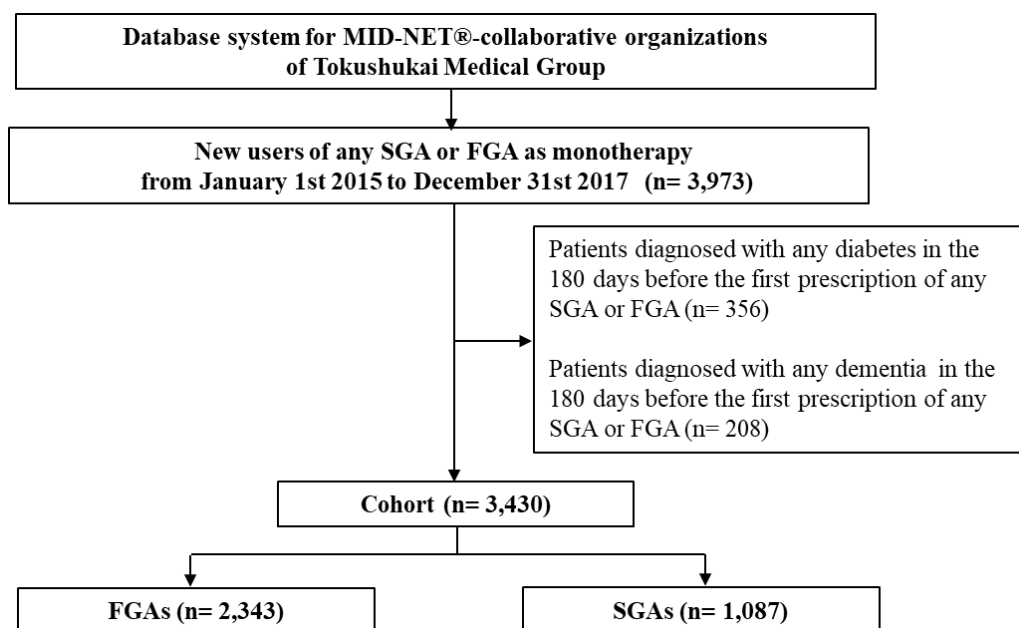
PS was estimated by a logistic regression model with confounding factors as covariates. In PS matching, one-to-one greedy matching was performed, and the caliper at the time of matching was 0.2 times the standard deviation of the Logit conversion value of PS. Stratified Cox proportional-hazard model by matched pairs was used as the outcome model in PS matching.

Abbreviations: aHR, adjusted hazard ratio; CC, complete cases; IPW, inverse probability weighted; MI, multiple imputation; PS, propensity score; SMRW, standardized mortality ratio weighting; SRI, single regression imputation.



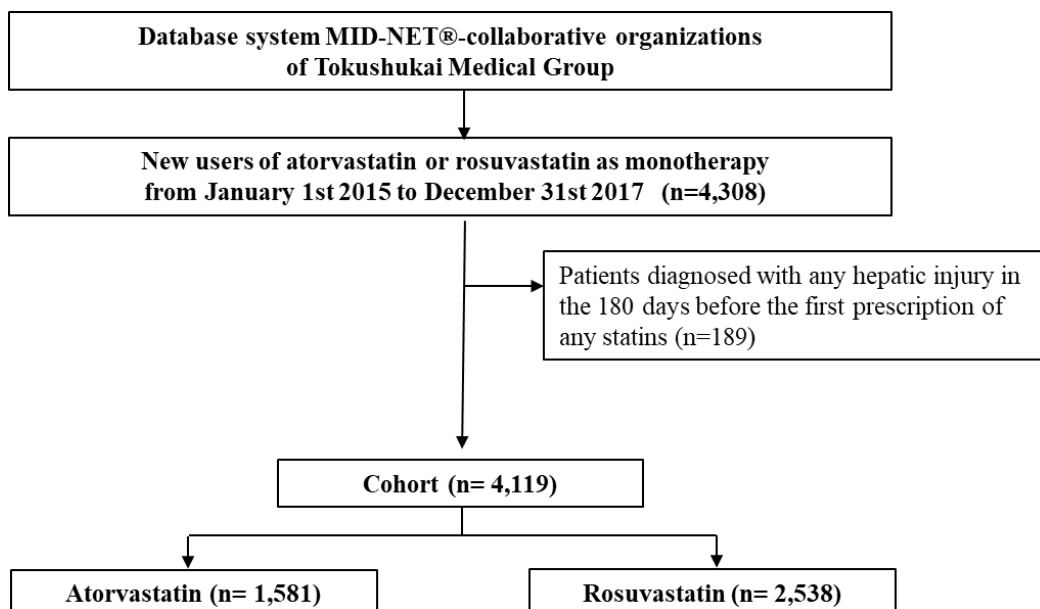


**Supplementary Figure S2. Study scenario selection flowchart**



**Supplementary Figure S3. Scenario 1. Number of patients in the study cohort: risk of diabetes associated with SGA**

Abbreviations: FGA, first-generation antipsychotic; SGA, second-generation antipsychotic.



**Supplementary Figure S4. Scenario 2. Number of patients in the study cohort: risk of hepatic injury associated with rosuvastatin**