[redacted]

Boehringer Ingelheim

Non-Interventional Study (NIS) Protocol

Document Number:	*Not applicable*
BI Study Number:	1245-0203
BI Investigational Product(s):	None
Title:	Clinical characteristics and disease burden of patients with diabetes mellitus combined with cardiovascular or chronic kidney disease based on regional medical database in Tianjin China
Brief lay title:	Clinical characteristics and disease burden of diabetic patients based on Tianjin regional database
Protocol version identifier:	3.0
Date of last version of protocol:	08 May 2021
PASS:	No
EU PAS register number:	EUPAS38023
Active substance:	*Not applicable*
Medicinal product:	*Not applicable*
Product reference:	*Not applicable*
Procedure number:	*Not applicable*
Marketing authorisation holder(s):	*Not applicable*
Joint PASS:	*No*
Research question and objectives:	The study aims to investigate the clinical characteristics, treatment, and economic burden of disease of Chinese diabetic/non-diabetic patients with/without established cardiovascular disease, chronic kidney disease, or at high cardiovascular risk
Country(-ies) of study:	China
Author:	[redacted]

	[redacted]
Marketing authorisation holder(s):	*Not applicable*
In case of PASS, add: **MAH contact person:**	*Not applicable*
In case of PASS, add: **<EU-QPPV:>**	*Not applicable*
In case of PASS, add: **<Signature of EU-QPPV:>**	*Not applicable*
Date:	18 Jun 2021
Page 1 of 28	

1. TABLE OF CONTENTS

2. LIST OF ABBREVIATIONS

ADR	Adverse drug reaction
AE	Adverse Event
BI	Boehringer Ingelheim
CHD	Coronary heart disease
CKD	Chronic kidney disease
CVD	Cardiovascular disease
DMRP	Data management and review plan
eGFR	Estimated glomerular filtration rate
EMR	Electronic medical records
ENCEPP	European Network of Centers for Pharmacoepidemiology and Pharmacovigilance
ESRD	End stage renal disease
GPP	Good Pharmacoepidemiology Practice
HDL	High-density lipoprotein
ICD	International classification of disease
IDF	International Diabetes Federation
IEC	Independent Ethics Committee
IRB	Institutional Review Board
LDL	Low-density lipoprotein
LOS	Length of stay
MALB	Microalbumin
MI	Myocardial infarction
NIS	Non-interventional Study
PI	Principal Investigator
SEAP	Statistical and epidemiological analysis plan
SOP	Standard Operating Procedures
TC	Total cholesterol
TG	Triglyceride

3. RESPONSIBLE PARTIES

This study is sponsored by Boehringer Ingelheim (BI).

BI appointed NIS Leads who will be responsible for coordinating the activities required in this study. BI NIS Lead will manage the study in accordance with applicable regulations and internal standard operating procedures (SOPs).

BI NIS [redacted]:

[redacted]
Responsibilities: Operation related tasks

[redacted]
Responsibilities: scientific related tasks

A principal investigator and a co-principal investigator will be nominated to coordinate and implement the study according to this protocol. Tasks and responsibilities for the principal investigators will be defined in a contract filed before the initiation of the study.

Principal Investigator (PI)
[redacted]

Co-Principal Investigator (Co-PI)
[redacted]

4. ABSTRACT

Name of company: Boehringer Ingelheim			
Name of finished medicinal product: NA			
Name of active ingredient: NA			
Protocol date: 18 March 2021	**Study number:** 1245-0203	**Version/Revision:** 3.0	**Version/Revision date:** 18 Jun 2021
Title of study:	Clinical characteristics and disease burden of patients with diabetes mellitus combined with cardiovascular or chronic kidney disease based on regional medical database in Tianjin China		
Rationale and background:	Diabetes has been a major public health concern all over the world. According to the data from the International Diabetes Federation (IDF) Diabetes Atlas 2017, there are now around 425 million adults aged 20-79 worldwide living with diabetes. In China, the prevalence of diabetes among adults was 10.9%, representing an estimated 114.4 million Chinese adults with diabetes. The mortality among patient with diabetes was high, which estimated that 4 million adults died from diabetes worldwide in 2017. And diabetes has become a worrying burden upon the world's healthcare system recently. Studies have shown that the global health expenditure attributed to diabetes was estimated at USD 673 billion in 2015 and has risen to USD 727 billion 2017. The heavy disease burden is mainly due to diabetic complications. Diabetes is a major risk factor for cardiovascular disease (CVD) and chronic kidney disease (CKD). The prevalence of CVD and CKD among patients with diabetes is high. Study suggested that diabetic kidney disease developed in up to 50 percent of diabetic patients at 20 years, with 15 percent having progressed to ESRD by this time. In the INTERHEART study, diabetes accounted for 10 percent of the population-attributable risk of a first myocardial infarction (MI). Compared to non-diabetes, the risk of CVD and CKD increased in patients with diabetes. A meta-analysis including 102 studies showed that patients with diabetes had an overall risk of coronary heart disease (CHD) twice that of patients without diabetes. Except for the risk of CVD and CKD, clinical characteristics and treatment patterns may also differ among patients with or without diabetes. Studies showed diabetic patients may have a greater burden of other risk factors than nondiabetics, such as hypertension, hyperlipidemia, and obesity. Compared with individuals without diabetes, those with diabetes have a greater extent of coronary ischemia and are more likely to have a MI and silent myocardial ischemia. Moreover, the effect of treatment on patients with or without diabetes may differ, thus the treatment patterns may be different. China has been identified with the largest absolute disease burden of diabetes in the world recently. Diabetic patients with established CVD or CKD are bringing growing pressure upon our nation's healthcare expenditure.		

	However, the characteristic profile of Chinese diabetic patients who have CVD, CKD or are at high risk of CVD remains unclear thus is in urgent need for in-depth investigation. Several studies have shown that diabetes patients often suffer from multiple comorbidities, including the CVD and clearly described the characteristics and treatment patterns of such patients. However, patient characteristics and treatment patterns may differ across regions and hospitals. Lack of adequate understanding of such information may result in suboptimal management of patients and policy decisions. This is particularly true in China. In current China, however, the information regarding diabetes or non-diabetes patients who also had other comorbid conditions (e.g. established CV diseases, CKD or at high risk for such problem), is limited; the patient characteristics, treatment patterns and economic burden may not be fully understood. The treatment outcomes may be improved with better understanding about the current treatment pattern, patient characteristics, and potential treatment gap. Regional medical data has received a lot of attention for its advantages on larger sample size and better representation of the general population. Tianjin regional database, which cover individual-level health information, is a valuable data source for clinical studies. Through integrating EMR systems of hospitals by patient unique identify code, the database contains comprehensive information regarding clinical care, such as patients' basic information, medical advice, diagnosis, laboratory examination, medical records and medical costs. Till now, data from 40 tertiary hospitals has been collected, and there were approximately 600,000 patients diagnosed with diabetes in these 40 tertiary hospitals from 01/01/2015 to 31/12/2019. Of these, 100,000 patients were with comprehensive information through integrating EMR systems. Therefore, based on Tianjin regional database, we will describe the demographic, clinical characteristics, treatment, and economic burden of disease of Chinese diabetic/non-diabetic patients with/without established CVD, CKD, or at high CV risk including hypertension and hyperlipidemia. And we believe that the resulting findings will provide relevant evidence to achieve better healthcare for diabetes patients with established or at high risk of CVD or CKD.
Research question and objectives:	The study aims to investigate the clinical characteristics, treatment, and economic burden of disease of Chinese diabetic/non-diabetic patients with/without established cardiovascular disease, chronic kidney disease, or at high cardiovascular risk
Study design:	Non-interventional study based on regional medical database
Population:	We will include diabetic/non-diabetic patients with/without established cardiovascular disease, or chronic kidney disease, or at high cardiovascular risk in Tianjin regional database from 01/01/2015 to 31/12/2019. We will exclude patients with non-Chinese nationalities, or duplicated storage (records with same inpatient code).
Variables:	Diabetes or not, demographic characteristics, comorbidities, complications, treatment pattern and economic burden, outcome variables
Data sources:	Tianjin regional database from 01/01/2015 to 31/12/2019, it comprises of EMR system data from 40 tertiary hospitals.

Study size:	This study is a database research project using an established regional database in Tianjin. From 01/01/2015 to 31/12/2019, about 600,000 patients diagnosed as diabetes in Tianjin. Of which, 100,000 patients contained comprehensive information about patients' basic information, medical advice, diagnosis, laboratory examination, medical records and medical costs.
Data analysis:	1) Description of the patient demographic characteristics, comorbidities, complications, treatment pattern and economic burden. Characteristics presented as continuous data will be summarized through means, standard deviations, medians, minimum and maximum values; whereas, categorical data will be summarized as counts and proportions. 2) All tests will be two-sided unless otherwise specified
Milestones:	- EC approval 06/2021 - Start of data collection 07/2021 - End of data collection 08/2021 - Final results 10/2021 - Final NIS report 12/2021 **Publication Plan – Submission Date** - Abstract: 11/2021 - Oral presentation: 12/2021 - Full paper: 03/2022

5. AMENDMENTS AND UPDATES

Version Number	Revision Date	Rationale for Amendment
1.0	18 Mar 2021	NA
2.0	08 May 2021	1. Updated the organization of PI 2. Updated the milestones based on the current progress 3. Supplemented some information under the study sites 4. Separated Heart Failure from other CVDs under inclusion criteria according to the further outcome 5. Clarified the method how to protect patient privacy *6.* Specified the time period of data source
3.0	18 Jun 2021	1. Amended the study title according to the objectives 2. Specified the criteria for grouping and indicator details 3. Updated the milestones based on the current progress 4. Updated the version & date of DMRP and SEAP

Amendment 1:

CHANGE 1: Updated the organization of PI
Section of the Clinical Trial Protocol: 3
Principal Investigator (PI)

[redacted]
Was changed to:
Principal Investigator (PI) [redacted]
Reason(s) for change 1:
PI participates in this study as a member of [redacted] instead of [redacted]
CHANGE 2: Updated the milestones based on the current progress
Section of the Clinical Trial Protocol: 4 & 6
- EC approval 03/2021 - Start of data collection 04/2021 - End of data collection 05/2021 - Final results 10/2021 - Final NIS report 12/2021
Was changed to:
- EC approval 05/2021 - Start of data collection 06/2021 - End of data collection 07/2021 - Final results 10/2021 - Final NIS report 12/2021
Reason(s) for change 2:
Updated based on the current progress
CHANGE 3: Supplemented some information under the study sites
Section of the Clinical Trial Protocol: 9.2.1
This study is based on Tianjin regional database (managed by [redacted]), which covers individual-level health information from 15 million residents in Tianjin. Till 2020, data from 40 tertiary hospitals has been collected, and we will use data from these 40 tertiary hospitals.
Was changed to:
This study is based on Tianjin regional database which is approved by Tianjin Municipal Health Commission and managed by [redacted]. and authorized to [redacted] for this study use. It covers individual-level health information from 15 million residents in Tianjin. Till 2020, data from 40 tertiary hospitals has been collected, and we will use data from these 40 tertiary hospitals. The study was co-sponsored by [redacted] of the Health Data Special Committee of the [redacted], who was responsible for the research protocol design, as well as data cleaning, data analysis and final research report writing, and [redacted] of [redacted] as Co-PI who will be responsible for the research design.

Reason(s) for change 3:
Clarified the relationships between the organizations and the 2 Co-PIs.
CHANGE 4: Separated Heart Failure from other CVDs under inclusion criteria according to the further outcome
Section of the Clinical Trial Protocol: 9.2.2
Inclusion Criteria: 1) Patients in the Tianjin regional database from 2015 to 2019. 2) Group A: patients with diagnosis of diabetes, and with diagnosis of cardiovascular disease, or chronic kidney disease or at high cardiovascular risk; 3) Group B: patients with diagnosis of diabetes, but not with diagnosis of cardiovascular disease, chronic kidney disease or at high cardiovascular risk; 4) Group C: patients with diagnosis of cardiovascular disease, or chronic kidney disease or at high cardiovascular risk, but not with diagnosis of diabetes; 5) Group D: patients without diagnosis of cardiovascular disease, chronic kidney disease or at high cardiovascular risk, and without diagnosis of diabetes. We will randomly select a group of non-diabetic patients without any of the above diseases by matching on age and gender. Definition of diabetes, cardiovascular disease, chronic kidney disease and high cardiovascular risk: • Diabetes: patients with at least 1 discharged diagnosis or 2 outpatient diagnosis of diabetes (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) (International classification of disease (ICD)-10 E10-E14) • Cardiovascular disease: patients with at least 1 discharged diagnosis or 2 outpatient diagnosis of ischemic heart diseases (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) (ICD-10 I20~I25); or patients with at least 1 discharged diagnosis or 2 outpatient diagnosis of cerebrovascular diseases (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) (ICD-10 I60~I69); or patients with at least 1 discharged diagnosis or 2 outpatient diagnosis of ischemic peripheral artery disease (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) (ICD-10 E10.501, E11.603, E14.501, E14.606, E14.503, I73.9, I99.03, I99.04); or patients with at least 1 discharged diagnosis or 2 outpatient diagnosis of heart failure (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) (ICD-10 I50); • Chronic kidney disease: inpatients with at least once 1 discharged diagnosis CKD (ICD-10 N18), or inpatients with the last estimated glomerular filtration rate (eGFR, calculated by CKD-EPI equation) <60 mL/min/1.73 m2 or prescription of dialysis, but not with the diagnosis of acute kidney injury (ICD-10 N17); or outpatients with at least 2 diagnosis of CKD (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) or with two consecutive eGFR (calculated by CKD-EPI equation) <60 mL/min/1.73 m2 by 90 days or more. • High cardiovascular risk: patients with at least 1 discharged diagnosis or 2 outpatient diagnosis of hypertension (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) (ICD-10 I10~I15); or at least 1 discharged diagnosis or 2 outpatient diagnosis of hyperlipidemia (ICD-10 E78.001-E78.003, E78.101, E78.203, E78.301-E78.304, E78.306, E78.401, E78.501, E78.902)
Was changed to:

Inclusion Criteria:

1) Patients in the Tianjin regional database from 01/01/2015 to 31/12/2019.
2) Group A: patients with diagnosis of diabetes, and with diagnosis of cardiovascular disease, heart failure, chronic kidney disease or at high cardiovascular risk;
3) Group B: patients with diagnosis of diabetes, but not with diagnosis of cardiovascular disease, heart failure, chronic kidney disease or at high cardiovascular risk;
4) Group C: patients with diagnosis of cardiovascular disease, heart failure, chronic kidney disease or at high cardiovascular risk, but not with diagnosis of diabetes;
5) Group D: patients without diagnosis of cardiovascular disease, heart failure, chronic kidney disease or at high cardiovascular risk, and without diagnosis of diabetes. We will randomly select a group of non-diabetic patients without any of the above diseases by matching on age and gender.

Definition of diabetes, cardiovascular disease, chronic kidney disease and high cardiovascular risk:

- Diabetes: patients with at least 1 discharged diagnosis or 2 outpatient diagnosis of diabetes (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) (International classification of disease (ICD)-10 E10-E14)
- Cardiovascular disease: patients with at least 1 discharged diagnosis or 2 outpatient diagnosis of ischemic heart diseases (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) (ICD-10 I20~I25); or patients with at least 1 discharged diagnosis or 2 outpatient diagnosis of cerebrovascular diseases (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) (ICD-10 I60~I69); or patients with at least 1 discharged diagnosis or 2 outpatient diagnosis of ischemic peripheral artery disease (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) (ICD-10 E10.501, E11.603, E14.501, E14.606, E14.503, I73.9, I99.03, I99.04);
- Heart failure: patients with at least 1 discharged diagnosis or 2 outpatient diagnosis of heart failure (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) (ICD-10 I50);
- Chronic kidney disease: inpatients with at least once 1 discharged diagnosis CKD (ICD-10 N18), or inpatients with the last estimated glomerular filtration rate (eGFR, calculated by CKD-EPI equation) <60 mL/min/1.73 m2 or prescription of dialysis, but not with the diagnosis of acute kidney injury (ICD-10 N17); or outpatients with at least 2 diagnosis of CKD (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) or with two consecutive eGFR (calculated by CKD-EPI equation) <60 mL/min/1.73 m2 by 90 days or more.
- High cardiovascular risk: patients with at least 1 discharged diagnosis or 2 outpatient diagnosis of hypertension (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) (ICD-10 I10~I15); or at least 1 discharged diagnosis or 2 outpatient diagnosis of hyperlipidemia (ICD-10 E78.001-E78.003, E78.101, E78.203, E78.301-E78.304, E78.306, E78.401, E78.501, E78.902)

Reason(s) for change 4:

To be consistent with the further outcome

CHANGE 5: Clarified the method how to protect patient privacy
Section of the Clinical Trial Protocol: 10.2
The study will use de-identified data which contains a significant level of protection against the release of personal information to outside entities. Informed patient consent could be waived. BI will not access data at individual level, but only aggregate results will be obtained.
Was changed to:
The study will use MD-5 to de-identify personal information in the data including names, ID-numbers, visit numbers, home addresses, etc. against the release of personal information to outside entities. Informed patient consent could be waived. BI will not access data at individual level, but only aggregate results will be obtained.
Reason(s) for change 5:
Clarified the method how to protect patient privacy
CHANGE 6: Specified the time period of data source
Section of the Clinical Trial Protocol: full text
…Tianjin regional database from 2015 to 2019…
Was changed to:
…Tianjin regional database from 01/01/2015 to 31/12/2019…
Reason(s) for change 6:
Specified the time period of data source

Amendment 2:

CHANGE 1: Amended the study title according to the objectives
Section of the Clinical Trial Protocol: Cover page and 4
Title: Characteristics, treatment, and economic burden of disease of Chinese diabetic/non-diabetic patients with/without established cardiovascular disease, chronic kidney disease, or at high cardiovascular risk **Brief lay title:** Clinical Characteristics and economic burden of Chinese diabetic/non-diabetic patients with/without established cardiovascular disease, chronic kidney disease, or at high cardiovascular risk
Was changed to:
Title: Clinical characteristics and disease burden of patients with diabetes mellitus combined with cardiovascular or chronic kidney disease based on regional medical database in Tianjin China **Brief lay title:** Clinical characteristics and disease burden of diabetic patients based on Tianjin regional database
Reason(s) for change 1:
Amended the study title according to the objectives and the study database
CHANGE 2: Specified the criteria for grouping and indicator details
Section of the Clinical Trial Protocol: 9.3
None
Was changed to:
The following variables will be collected in this study. 1) Demographic characteristics of the patients: age, gender, type of health insurance, etc. 2) Status of the patient's disease diagnosis: outpatient diagnosis, discharge diagnosis. 3) Information on the patient's laboratory tests, glycated haemoglobin, blood glucose,

blood creatinine, estimated glomerular filtration rate, urine microalbumin, lipids, etc.
4) Patient medication and treatment: including CVD-related treatment (antiplatelet agents, statins, vasodilators), diabetes-related treatment (insulin, oral hypoglycaemic agents), hypertension-related treatment (e.g. calcium antagonists, angiotensin-converting enzyme antagonists, angiotensin receptor antagonists, beta-receptor antagonists, etc.), hyperlipidaemia-related treatment (e.g. lipid-lowering drugs), CKD-related treatment (e.g., dialysis, angiotensin-converting enzyme antagonists, angiotensin receptor antagonists, kidney transplantation), etc.
5) Number of days the patient was hospitalized
6) Patient's medical costs, including all costs during hospitalization (hospitalization, drugs, surgery, etc.)

Table 1. variables collected from database

...

Reason(s) for change 2:

Specified the criteria for grouping and indicator details

CHANGE 3: Updated the milestones based on the current progress

Section of the Clinical Trial Protocol: 4 & 6

- EC approval	05/2021
- Start of data collection	06/2021
- End of data collection	07/2021
- Final results	10/2021
- Final NIS report	12/2021

Was changed to:

- EC approval	06/2021
- Start of data collection	07/2021
- End of data collection	08/2021
- Final results	10/2021
- Final NIS report	12/2021

Reason(s) for change 3:

Updated based on the current progress

CHANGE 4: Updated the version & date of DMRP and SEAP

Section of the Clinical Trial Protocol: ANNEX 1

Number	Document Reference Number	Date	Title
1	*To be determined*	*To be determined*	Data Management Review Plan (DMRP)
2	*To be determined*	*To be determined*	Statistical and Epidemiological Analysis Plan (SEAP)

Was changed to:

	Number	Document Reference Number	Date	Title
	1	*Not applicable*	21 May 2021, Version 1.0	Data Management Review Plan (DMRP)
	2	*Not applicable*	13 May 2021, Version 1.0	Statistical and Epidemiological Analysis Plan (SEAP)

Reason(s) for change 4:

The DMRP and SEAP is now approved so have a valid version and date

6. MILESTONES

Milestone	Planned Date
IRB/IEC approval	Jun 2021
Start of data collection	Jul 2021
End of data collection	Aug 2021
Final report of study results:	Dec 2021

7. RATIONALE AND BACKGROUND

Diabetes has been a major public health concern all over the world. According to the data from the IDF Diabetes Atlas 2017, there are now around 425 million adults aged 20-79 worldwide living with diabetes. In China, the prevalence of diabetes among adults was 10.9%, representing an estimated 114.4million Chinese adults with diabetes. The mortality among patient with diabetes was high, which estimated that 4 million adults died from diabetes worldwide in 2017. And diabetes has become a worrying burden upon the world's healthcare system recently[2]. Studies have shown that the global health expenditure attributed to diabetes was estimated at USD 673 billion in 2015 and has risen to USD 727 billion 2017.

The heavy disease burden is mainly due to diabetic complications. Diabetes is a major risk factor for cardiovascular disease (CVD) and chronic kidney disease (CKD). The prevalence of CVD and CKD among patients with diabetes was high. Study suggested that diabetic kidney disease developed in up to 50 percent of diabetic patients at 20 years after a diabetes diagnosis, with 15 percent having progressed to end stage renal disease (ESRD) by this time. In the INTERHEART study, diabetes accounted for 10 percent of the population-attributable risk of a first myocardial infarction (MI). Compared to non-diabetes, the risk of CVD and CKD increased in patients with diabetes. A meta-analysis including 102 studies showed that patients with diabetes had an overall risk of coronary heart disease (CHD) twice that of patients without diabetes[8]. Except for the risk of CVD and CKD, clinical characteristics and treatment patterns may also differ among patients with or without diabetes. Studies showed diabetic patients may have a greater burden of other risk factors than nondiabetics, such as hypertension, hyperlipidemia, and obesity. Compared with individuals without diabetes, those with diabetes have a greater extent of coronary ischemia, and are more likely to have a MI and silent myocardial ischemia[10 11]. Moreover, the effect of treatment on patients with or without diabetes may differ, thus the treatment patterns may be different.

China has the largest absolute disease burden of diabetes in the world recently. Diabetic patients with established CVD or CKD are bringing growing pressure upon our nation's healthcare expenditure. However, the characteristic profile of Chinese diabetic patients who has CVD, CKD or at high risk of CVD remains unclear thus is in urgent need for in-depth investigation. Several studies have shown that diabetes patients often suffer from multiple comorbidities, including the CV diseases, and clearly described the characteristics and treatment patterns of such patients. However, patient characteristics and treatment patterns

may differ across regions and hospitals. Lack of adequate understanding of such information may result in suboptimal management of patients and policy decisions. This is particularly true in China. In current China, however, the information regarding diabetes or non-diabetes patients who also had other comorbid conditions (e.g. established CV diseases, CKD or at high risk for such problem), is limited; the patient characteristics, treatment patterns and economic burden may not be fully understood. The treatment outcomes may be improved with better understanding about the current treatment pattern, patient characteristics, and potential treatment gap.

Regional medical data has received a lot of attention for its advantages on larger sample size and better representation of the general population. Tianjin regional database, which cover individual-level health information, is a valuable data source for clinical studies. Through integrating EMR systems of hospitals by patient unique identify code, the database contains comprehensive information regarding clinical care, such as patients' basic information, medical advice, diagnosis, laboratory examination, medical records and medical costs. It could be a valuable big data resource for clinical studies. Till now, data from 40 tertiary hospitals has been collected, and there were approximately 600,000 patients diagnosed with diabetes in these 40 tertiary hospitals from 01/01/2015 to 31/12/2019. Of these, 100,000 patients were with comprehensive information through integrating EMR systems. And Tianjin has built a unified medical purchasing center for hospitals to undertake joint procurement of medical supplies and drugs.

Therefore, based on Tianjin regional database, we will describe the demographic, clinical characteristics, treatment, and economic burden of disease of Chinese diabetic/non-diabetic patients with/without established CV disease, CKD, or at high CV risk including hypertension and hyperlipidemia. And we believe that the resulting findings will inform a comprehensive group of evidence users to achieve better healthcare for diabetes patients with established or at high risk of CVD or CKD.

8. RESEARCH QUESTION AND OBJECTIVES

The study aims to investigate the clinical characteristics, treatment, and economic burden of disease of Chinese diabetic/non-diabetic patients with/without established cardiovascular disease, chronic kidney disease, or at high cardiovascular risk, including:

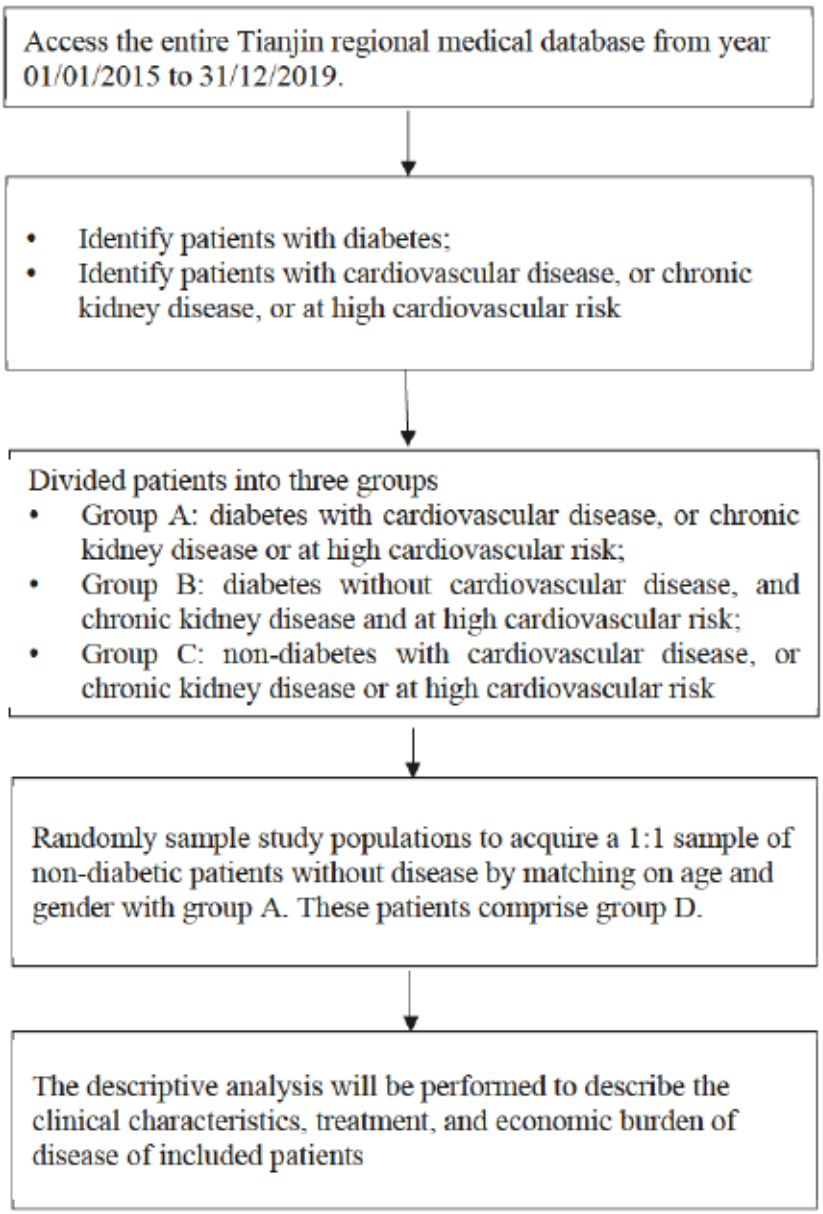
- Primary objectives: describe the proportion of Chinese diabetic/non-diabetic patients with established cardiovascular disease, CKD, or at high cardiovascular risk including hypertension and hyperlipidemia.
- Secondary objectives: describe the demographic characteristics of the last visit for all patients, and the demographic characteristics of inpatients over time; investigate the clinical characteristic for all patients.
- Further objectives: examine economic burden of disease of Chinese diabetic/non-diabetic patients with/without established cardiovascular disease, CKD, or at high cardiovascular risk; explore the trends of clinical characteristics, treatment patterns, economic burden of inpatients over time.

9. RESEARCH METHODS

9.1 STUDY DESIGN

This study is an observational, cross-sectional, study based on routinely-collected-data. We will use Tianjin regional medical database. It is conducted to investigate the clinical characteristics, treatment, and economic burden of disease of Chinese diabetic/non-diabetic patients with/without established cardiovascular disease, chronic kidney disease, or at high cardiovascular risk.

The study process is illustrated in figure 1.



9.2 SETTING

9.2.1 Study sites

This study is based on Tianjin regional database which is approved by Tianjin Municipal Health Commission and managed by [redacted]. and authorized to [redacted] for this study use. It covers individual-level health information from 15 million residents in Tianjin. Till 2020, data from 40 tertiary hospitals has been collected, and we will use data from these 40 tertiary hospitals. The study was co-sponsored by [redacted] of the Health Data Special Committee of the [redacted], who was responsible for the research protocol design, as well as data cleaning, data analysis and final research report writing, and [redacted] of [redacted] as Co-PI who will be responsible for the research design.

9.2.2 Study population

Diabetic/non-diabetic patients with/without established cardiovascular disease, chronic kidney disease, or at high cardiovascular risk.

Inclusion Criteria:

1) Patients in the Tianjin regional database from 01/01/2015 to 31/12/2019.
2) Group A: patients with diagnosis of diabetes, and with diagnosis of cardiovascular disease, heart failure, chronic kidney disease or at high cardiovascular risk.
3) Group B: patients with diagnosis of diabetes, but not with diagnosis of cardiovascular disease, heart failure, chronic kidney disease or at high cardiovascular risk.
4) Group C: patients with diagnosis of cardiovascular disease, heart failure, chronic kidney disease or at high cardiovascular risk, but not with diagnosis of diabetes.
5) Group D: patients without diagnosis of cardiovascular disease, heart failure, chronic kidney disease or at high cardiovascular risk, and without diagnosis of diabetes. We will randomly select a group of non-diabetic patients without any of the above diseases by matching on age and gender.

Definition of diabetes, cardiovascular disease, chronic kidney disease and high cardiovascular risk:

- Diabetes: patients with at least 1 discharged diagnosis or 2 outpatient diagnosis of diabetes (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) (International classification of disease (ICD)-10 E10-E14)
- Cardiovascular disease: patients with at least 1 discharged diagnosis or 2 outpatient diagnosis of ischemic heart diseases (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) (ICD-10 I20~I25); or patients with at least 1 discharged diagnosis or 2 outpatient diagnosis of cerebrovascular diseases (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) (ICD-10 I60~I69); or patients with at least 1 discharged diagnosis or 2 outpatient diagnosis of ischemic peripheral artery disease (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) (ICD-10 E10.501, E11.603, E14.501, E14.606, E14.503, I73.9, I99.03, I99.04);

- Heart failure: patients with at least 1 discharged diagnosis or 2 outpatient diagnosis of heart failure (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) (ICD-10 I50).
- Chronic kidney disease: inpatients with at least once 1 discharged diagnosis CKD (ICD-10 N18), or inpatients with the last estimated glomerular filtration rate (eGFR, calculated by CKD-EPI equation) <60 mL/min/1.73 m2 or prescription of dialysis, but not with the diagnosis of acute kidney injury (ICD-10 N17); or outpatients with at least 2 diagnosis of CKD (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) or with two consecutive eGFR (calculated by CKD-EPI equation) <60 mL/min/1.73 m2 by 90 days or more.
- High cardiovascular risk: patients with at least 1 discharged diagnosis or 2 outpatient diagnosis of hypertension (the first diagnosis will be the index diagnosis, and the time interval between diagnoses is not restricted) (ICD-10 I10~I15); or at least 1 discharged diagnosis or 2 outpatient diagnosis of hyperlipidemia (ICD-10 E78.001-E78.003, E78.101, E78.203, E78.301-E78.304, E78.306, E78.401, E78.501, E78.902)

Exclusion Criteria:

1) Patients with non-Chinese nationalities.
2) Duplicated storage (records with same inpatient code).

9.2.3 Study visits

Not applicable

9.2.4 Study discontinuation

Boehringer Ingelheim reserves the right to discontinue the study at any time for the following reasons:

1. Failure to reach eligible analyzable population/data
2. Violation of Good Pharmacoepidemiology Practice (GPP), the study protocol, or the contract by a study site, investigator or research collaborator, disturbing the appropriate conduct of the study

The investigator /research collaborator will be reimbursed for reasonable expenses incurred in case of study termination defined in a contract filed before the initiation of the study (except in case of the third reason).

9.3 VARIABLES

The following variables will be collected in this study.

1) Demographic characteristics of the patients: age, gender, type of health insurance, etc.
2) Status of the patient's disease diagnosis: outpatient diagnosis, discharge diagnosis.
3) Information on the patient's laboratory tests, glycated haemoglobin, blood glucose, blood creatinine, estimated glomerular filtration rate, urine microalbumin, lipids, etc.
4) Patient medication and treatment: including CVD-related treatment (antiplatelet agents, statins, vasodilators), diabetes-related treatment (insulin, oral hypoglycaemic agents),

hypertension-related treatment (e.g. calcium antagonists, angiotensin-converting enzyme antagonists, angiotensin receptor antagonists, beta-receptor antagonists, etc.), hyperlipidaemia-related treatment (e.g. lipid-lowering drugs), CKD-related treatment (e.g., dialysis, angiotensin-converting enzyme antagonists, angiotensin receptor antagonists, kidney transplantation), etc.

5) Number of days the patient was hospitalized
6) Patient's medical costs, including all costs during hospitalization (hospitalization, drugs, surgery, etc.)

Table 1. variables collected from database

Table Name	Variables
Basic information	medical organization
	unique ID of patients
	birthday
	sex
	ethic
	type of insurance
Laboratory	unique ID of patients
	report number
	test date
	test time
	item test name
	item test result
	item test unit
	item test sample
	item test range
Diagnosis	unique ID of patients
	diagnosis
	admission status
	discharge status
	diagnosis date
	department
Advice	unique ID of patients
	advice item name
	dose per time

	total dose
	unit of dosage
	frequency
	advice start date
	advice stop date
Front sheet	unique ID of patients
	admission department
	admission date
	discharge department
	discharge date
	length of hospital stay
	discharge way
Fee information	unique ID of patients
	fee item name
	unit price
	quantity
	total price
	charge date

9.3.1 Exposures

Not applicable.
We defined four groups (group A-D) according to the diagnosis and laboratory measurements. The group A and group B included all diabetic patients with/without established CV disease, CKD, or at high CV risk, respectively. The group C included all non-diabetic patients with established cardiovascular disease, chronic kidney disease, or at high cardiovascular risk. The group D included non-diabetic patients without established cardiovascular disease, chronic kidney disease, and not at high cardiovascular risk. For group D, we will randomly sample study populations to acquire a 1:1 sample of non-diabetic patients without disease by matching on age and gender with group A.

9.3.2 Outcomes

Our primary outcomes will be the proportion of diabetic/non-diabetic patients with disease/risk. Secondary outcomes include demographic characteristics (e.g. age, gender, insurance payment) and clinical characteristics (e.g. HbA1c, random blood glucose, serum creatine) of patients. Further outcomes include economic burden of disease of diabetic/non-diabetic patients with or without disease/risk (e.g. length of stay, hospital cost).

9.3.2.1 Primary outcomes

1. The proportion of diabetic/non-diabetic patients with established cardiovascular disease, or chronic kidney disease, or at high cardiovascular risk.
2. The proportion of diabetic patients with established cardiovascular disease, or chronic kidney disease, or at high cardiovascular risk among all diabetic patients.
3. The proportion of non-diabetic patients with established cardiovascular disease, or chronic kidney disease, or at high cardiovascular risk among all non-diabetic patients.

9.3.2.2 Secondary outcomes

1. The demographic characteristics of the latest visit for all patients, as well as the trends for inpatients over time (2015, 2017, and 2019 respectively):
   1) Age at admission
   2) Gender
   3) Insurance payment

2. The clinical characteristics of the studied population of the latest visit for all patients, as well as the trends for inpatients over time (2015, 2017, and 2019 respectively):
   1) Discharge department
   2) Death
   3) Laboratory test results:
      a) HbA1c
      b) random blood glucose
      c) serum creatine

[redacted]

9.3.3 Covariates

No Applicable.

9.4 DATA SOURCES

Tianjin regional database cover individual-level health information. There were 40 tertiary hospitals, 32 secondary hospitals and 276 community hospitals Tianjin. It has approximately 200,000 discharges annually. Through integrating EMR systems of hospitals by patient unique identify code, the database contains comprehensive information regarding clinical care, such as patients' basic information, medical advice, diagnosis, laboratory examination, medical records and medical costs. It could be a valuable big data resource for clinical studies. Till 2020, data from 40 tertiary hospitals has been collected, and there were approximately 600,000 patients diagnosed with diabetes in these 40 tertiary hospitals from 01/01/2015 to 31/12/2019. Of these, 100,000 patients were with comprehensive information through integrating EMR systems.

9.5 STUDY SIZE

This study is a database research project using an established regional database in Tianjin. From 01/01/2015 to 31/12/2019, about 600,000 patients diagnosed as diabetes in Tianjin. Of which, 100,000 patients contained comprehensive information about patients' basic information, medical advice, diagnosis, laboratory examination, medical records and medical costs.

9.6 DATA MANAGEMENT

The data management plan is summarized below. Full details of the data management plan are documented in a separate NIS-Data Management and Review Plan (NIS-DMRP).
This study will use de-identified data which contains a significant level of protection against the release of personal information to outside entities. The investigator or qualified designee is responsible for recording and verifying the accuracy of subject data.

9.7 DATA ANALYSIS

The statistical analysis plan for the study is summarized below. Full details of the statistical analysis will be documented in the SEAP, which will be finalized before the end of data collection.

9.7.1 Main analysis

Before analyses start, 200 randomly selected medical chart will be fully reviewed by well-trained experts to validate the accuracy and sensibility of code used to identify patients with cardiovascular disease or at high cardiovascular risk, chronic kidney disease and type 2 diabetes. This QC procedure for database study would be performed by [redacted].
The analyses will mainly describe the patient demographic characteristics, comorbidities, treatment pattern and economic burden based on 40 hospitals EMRs of Tianjin regional database. Descriptive data will be provided. The statistical methods and analyses will be described in detail in a Statistical and Epidemiological Analysis Plan (SEAP) prior to the commencement of any analyses.

Characteristics presented as continuous data will be summarized through means, standard deviations, medians, minimum and maximum values, whereas, categorical data will be summarized as counts and proportions.
All tests will be two-sided unless otherwise specified

[redacted]

9.7.3 Safety Analysis

Not applicable

9.8 QUALITY CONTROL

The quality control and review are summarized below. Greater details are documented in the NIS-DMRP.

The investigator is responsible for implementing and maintaining a quality management system with written development procedures and functional area standard operating procedures (SOPs) to ensure that studies are conducted and data are generated, documented, and reported in compliance with the protocol, accepted standards of Good Pharmacoepidemiology Practice, and all laws, rules and regulations relating to the conduct of the study.

9.9 LIMITATIONS OF THE RESEARCH METHODS

There is no unique ICD code for cardiovascular disease, CKD or at high cardiovascular risk. Each of those diseases has more than one ICD code in the medical charts. The subjects came from different hospitals; the understanding of ICD code may be different according to different hospitals.

The study is limited in representativeness and generalizability. Of 600,000 patients, only 100,000 patients contained comprehensive information about patients' basic information, medical advice, diagnosis, laboratory examination, medical records and medical costs.

9.10 OTHER ASPECTS

9.10.1 Data quality assurance

A quality assurance audit/inspection of this study may be conducted by the sponsor or sponsor's designees or by regulatory authorities. The quality assurance auditor will have access to all results of the study, the investigator's study-related files and correspondence, and the informed consent documentation of this study.

9.10.2 Study records

None

10. PROTECTION OF HUMAN SUBJECTS

The section is not applicable since this is a non-interventional study based on existing database.

The procedures set out in this study protocol are designed to ensure that the sponsor and investigator abide by the principles of the GPP guidelines. The study also will be carried out in keeping with local legal requirements and BI Standard Operating Procedures (SOPs).

10.1 STUDY APPROVAL, PATIENT INFORMATION, AND INFORMED CONSENT

This NIS will be initiated only after all required legal documentation has been reviewed and approved by the respective Institutional Review Board (IRB)/Independent Ethics Committee (IEC) and Competent Authority (CA) according to national and international regulations. The same applies for the implementation of changes introduced by amendments.

10.2 STATEMENT OF CONFIDENTIALITY

The study will use MD-5 to de-identify personal information in the data including names, ID-numbers, visit numbers, home addresses, etc. against the release of personal information to outside entities. Informed patient consent could be waived. BI will not access data at individual level, but only aggregate results will be obtained.

11. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

The current study is an analysis of the existing data collected from the Tianjin regional database. Boehringer Ingelheim or third party acting on behalf of BI will not have access to individual patient data during the study. Data will be provided to BI in aggregated manner. Thus, collection and reporting of adverse events (AEs)/adverse drug reactions (ADRs) is not applicable.

12. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

The rights of the investigator and of the sponsor with regard to publication of the results of this study are described in the investigator contract. As a general rule, no study results should be published prior to finalization of the NIS Report.

The final publication will be written after the final NIS report is completed.

BI intends to use data from this study to prepare peer-reviewed publications and other scientific communications, study protocol and final study report planned for American Diabetes Association Scientific Sessions 2022.
The final report plan to be published in Diabetes Research and Clinical Practice.

Results of this NIS will be disclosed on clinicaltrials.gov and encepp.eu once available.

13. REFERENCES

13.1 PUBLISHED REFERENCES

R Federation ID. IDF Diabetes Atlas. 8th edn ed, 2017.
J-2016-316(24) Dieleman JL, Baral R, Birger M, et al. US Spending on Personal Health Care and Public Health, 1996-2013. JAMA 2016;316(24):2627-46. [published Online First: 2016/12/28]
J-1998-339(4) Haffner SM, Lehto S, Ronnemaa T, et al. Mortality from coronary heart disease in subjects with type 2 diabetes and in nondiabetic subjects with and without prior myocardial infarction. The New England journal of medicine 1998;339(4):229-34. [published Online First: 1998/07/23]
J-2005-28(12) Juutilainen A, Lehto S, Ronnemaa T, et al. Type 2 diabetes as a "coronary heart disease equivalent": an 18-year prospective population-based study in Finnish subjects. Diabetes care 2005;28(12):2901-7. [published Online First: 2005/11/25]
J-2008-358(24) Patel A, MacMahon S, Chalmers J, et al. Intensive blood glucose control and vascular outcomes in patients with type 2 diabetes. The New England journal of medicine 2008;358(24):2560-72. [published Online First: 2008/06/10]
J-1993-16(1) Nelson RG, Knowler WC, Pettitt DJ, et al. Diabetic kidney disease in Pima Indians. Diabetes care 1993;16(1):335-41. [published Online First: 1993/01/01]
J-2004-364(9438) Yusuf S, Hawken S, Ounpuu S, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. Lancet (London, England) 2004;364(9438):937-52. [published Online First: 2004/09/15]
J-2010-375(9733) Sarwar N, Gao P, Seshasai SR, et al. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. Lancet (London, England) 2010;375(9733):2215-22. [published Online First: 2010/07/09]
J-1999-100(10) Grundy SM, Benjamin IJ, Burke GL, et al. Diabetes and cardiovascular disease: a statement for healthcare professionals from the American Heart Association. Circulation 1999;100(10):1134-46. [published Online First: 1999/09/08]
J-1973-32(1) Margolis JR, Kannel WS, Feinleib M, et al. Clinical features of unrecognized myocardial infarction--silent and symptomatic. Eighteen years follow-up: the Framingham study. The American journal of cardiology 1973;32(1):1-7. [published Online First: 1973/07/01]
J-2006-47(1) Scognamiglio R, Negut C, Ramondo A, et al. Detection of coronary artery disease in asymptomatic patients with type 2 diabetes mellitus. Journal of the American College of Cardiology 2006;47(1):65-71. [published Online First: 2006/01/03]

13.2 UNPUBLISHED REFERENCES

None

ANNEX 1. LIST OF STAND-ALONE DOCUMENTS

Number	Document Reference Number	Date	Title
1	*Not applicable*	21 May 2021, Version 1.0	Data Management Review Plan (DMRP)
2	*Not applicable*	13 May 2021, Version 1.0	Statistical and Epidemiological Analysis Plan (SEAP)

ANNEX 2. ENCEPP CHECKLIST FOR STUDY PROTOCOLS

Study title: ***Clinical characteristics and disease burden of patients with diabetes mellitus combined with cardiovascular or chronic kidney disease based on regional medical database in Tianjin China***

EU PAS Register® number:38023
Study reference number (if applicable): 1245-0203

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[1]	☒	☐	☐	6
1.1.2 End of data collection[2]	☒	☐	☐	6
1.1.3 Progress report(s)	☒	☐	☐	6
1.1.4 Interim report(s)	☐	☐	☒	6
1.1.5 Registration in the EU PAS Register®	☒	☐	☐	6
1.1.6 Final report of study results.	☒	☐	☐	6

Comments:

Section 2: Research question	**Yes**	**No**	**N/A**	**Section Number**
2.1 Does the formulation of the research question and objectives clearly explain:	☒	☐	☐	7& 8& 9

[1] Date from which information on the first study is first recorded in the study dataset or, in the case of secondary use of data, the date from which data extraction starts.
[2] Date from which the analytical dataset is completely available.

Section 2: Research question	**Yes**	**No**	**N/A**	**Section Number**
2.1.1 Why the study is conducted? (e.g. to address an important public health concern, a risk identified in the risk management plan, an emerging safety issue)	☒	☐	☐	7
2.1.2 The objective(s) of the study?	☒	☐	☐	8
2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	9.2.2
2.1.4 Which hypothesis(-es) is (are) to be tested?	☐	☐	☒	
2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☐	☒	

Comments:

Section 3: Study design	**Yes**	**No**	**N/A**	**Section Number**
3.1 Is the study design described? (e.g. cohort, case-control, cross-sectional, other design)	☒	☐	☐	9.1
3.2 Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	9.4
3.3 Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	☒	☐	☐	9.3.2
3.4 Does the protocol specify measure(s) of association? (e.g. risk, odds ratio, excess risk, rate ratio, hazard ratio, risk/rate difference, number needed to harm (NNH))	☐	☐	☒	
3.5 Does the protocol describe the approach for the collection and reporting of adverse events/adverse reactions? (e.g. adverse events that will not be collected in case of primary data collection)	☐	☐	☒	

Comments:

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Section Number**
4.1 Is the source population described?	☒	☐	☐	9.2
4.2 Is the planned study population defined in terms of:				
4.2.1 Study time period	☒	☐	☐	9.2
4.2.2 Age and sex	☐	☒	☐	9.2
4.2.3 Country of origin	☒	☐	☐	9.2
4.2.4 Disease/indication	☒	☐	☐	9.2
4.2.5 Duration of follow-up	☐	☐	☒	9.2

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Section Number**
4.3 Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	9.2

Comments:

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
5.1 Does the protocol describe how the study exposure is defined and measured? (e.g. operational details for defining and categorising exposure, measurement of dose and duration of drug exposure)	☐	☐	☒	
5.2 Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)	☐	☐	☒	
5.3 Is exposure categorised according to time windows?	☐	☐	☒	
5.4 Is intensity of exposure addressed? (e.g. dose, duration)	☐	☐	☒	
5.5 Is exposure categorised based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☐	☐	☒	
5.6 Is (are) (an) appropriate comparator(s) identified?	☐	☐	☒	

Comments:

Section 6: Outcome definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
6.1 Does the protocol specify the primary and secondary (if applicable) outcome(s) to be investigated?	☒	☐	☐	9.3.2
6.2 Does the protocol describe how the outcomes are defined and measured?	☒	☐	☐	9.3.2
6.3 Does the protocol address the validity of outcome measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, use of validation sub-study)	☐	☒	☐	
6.4 Does the protocol describe specific outcomes relevant for Health Technology Assessment? (e.g. HRQoL, QALYs, DALYS, health care services utilisation, burden of disease or treatment, compliance, disease management)	☐	☒	☐	

Comments:

Section 7: Bias	**Yes**	**No**	**N/A**	**Section Number**
7.1 Does the protocol address ways to measure confounding? (e.g. confounding by indication)	☐	☐	☒	
7.2 Does the protocol address selection bias? (e.g. healthy user/adherer bias)	☒	☐	☐	9.9
7.3 Does the protocol address information bias? (e.g. misclassification of exposure and outcomes, time-related bias)	☐	☒	☐	

Comments:

Section 8: Effect measure modification	**Yes**	**No**	**N/A**	**Section Number**
8.1 Does the protocol address effect modifiers? (e.g. collection of data on known effect modifiers, sub-group analyses, anticipated direction of effect)	☐	☐	☒	

Comments:

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
9.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	☐	☐	☒	
9.1.2 Outcomes? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	☒	☐	☐	9.4
9.1.3 Covariates and other characteristics?	☐	☐	☒	
9.2 Does the protocol describe the information available from the data source(s) on:				
9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☐	☐	☒	
9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	9.4
9.2.3 Covariates and other characteristics? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☐	☐	☒	
9.3 Is a coding system described for:				
9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☐	☐	☒	

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☐	☒	☐	
9.3.3 Covariates and other characteristics?	☐	☐	☒	
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☒	☐	☐	9.4

Comments:

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.1 Are the statistical methods and the reason for their choice described?	☒	☐	☐	9.7.1
10.2 Is study size and/or statistical precision estimated?	☒	☐	☐	9.5
10.3 Are descriptive analyses included?	☒	☐	☐	9.7.1
10.4 Are stratified analyses included?	☐	☒	☐	
10.5 Does the plan describe methods for analytic control of confounding?	☐	☒	☐	
10.6 Does the plan describe methods for analytic control of outcome misclassification?	☐	☒	☐	
10.7 Does the plan describe methods for handling missing data?	☐	☒	☐	
10.8 Are relevant sensitivity analyses described?	☒	☐	☐	9.7.1

Comments:

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Section Number**
11.1 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☐	☒	☐	
11.2 Are methods of quality assurance described?	☒	☐	☐	9.10.1
11.3 Is there a system in place for independent review of study results?	☐	☒	☐	

Comments:

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1 Does the protocol discuss the impact on the study results of:				
12.1.1 Selection bias?	☒	☐	☐	9.9

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1.2 Information bias?	☒	☐	☐	9.9
12.1.3 Residual/unmeasured confounding? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods).	☐	☐	☒	
12.2 Does the protocol discuss study feasibility? (e.g. study size, anticipated exposure uptake, duration of follow-up in a cohort study, patient recruitment, precision of the estimates)	☐	☐	☒	

Comments:

Section 13: Ethical/data protection issues	**Yes**	**No**	**N/A**	**Section Number**
13.1 Have requirements of Ethics Committee/ Institutional Review Board been described?	☒	☐	☐	9.10.1
13.2 Has any outcome of an ethical review procedure been addressed?	☐	☒	☐	
13.3 Have data protection requirements been described?	☒	☐	☐	10.2

Comments:

Section 14: Amendments and deviations	**Yes**	**No**	**N/A**	**Section Number**
14.1 Does the protocol include a section to document amendments and deviations?	☒	☐	☐	5

Comments:

Section 15: Plans for communication of study results	**Yes**	**No**	**N/A**	**Section Number**
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	☒	☐	☐	12
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	12

Comments:

[redacted]

Study Title: Clinical characteristics and disease burden of patients with diabetes mellitus combined with cardiovascular or chronic kidney disease based on regional medical database in Tianjin China

Study Number: 1245-0203

Protocol Version: 3.0

I herewith certify that I agree to the content of the study protocol and to all documents referenced in the study protocol.

Position: PI	Name/Date [redacted]	Signature:
Position: Co-PI	Name/Date	Signature: DocuSigned by: [redacted] 签名者名称： [redacted] 签名原因： 我是此文档的作者 [redacted]
Position: Author	Name/Date	Signature: [redacted] 21 | 08:51 CEST [redacted]
Position: NIS [redacted]	Name/Date	Signature: 签名者名称： [redacted] 签名原因： 我批准此文档 签名时间： 22-6-2021 | 08:49 CEST
Position: NIS [redacted]	Name/Date	Signature: DocuSigned by: [redacted] 签名者名称： [redacted] 签名原因： 我批准此文档 签名时间： 22-6-2021 | 08:49 CEST
Position: RWE CoE	Name/Date: [redacted] or [redacted]	Signature: DocuSigned by: [redacted] Signer Name: [redacted] Signing Reason: I approve this docum Signing Time: 22-Jun-2021 | 09:31 CE [redacted]
Position: Global TA [redacted] Epi	Name/Date: [redacted]	Signature: DocuSigned by: [redacted] Name des Unterzeichners: [redacted] Signiergrund: Ich genehmige dieses Dokument Signierzeit: 24-Jun-2021 | 13:01 MESZ [redacted]
Position: Global TA [redacted] MA	Name/Date: [redacted] on behalf of Global TA [redacted] MA	Signature: DocuSigned by: [redacted] Signer Name: [redacted] Signing Reason: I approve this docum Signing Time: 24-Jun-2021 | 07:50 CE [redacted]
Position: Global TA [redacted] Market Access	Name/Date: [redacted] on behalf of Global TA [redacted] of MAcc	Signature: DocuSigned by: [redacted] Signer Name: [redacted] Signing Reason: I have reviewed this documen Signing Time: 22-Jun-2021 | 08:43 CEST [redacted]