

# Practical use of electronic health records among patients with diabetes in scientific research

Yun Shen<sup>1</sup>, Jian Zhou<sup>1</sup>, Gang Hu<sup>2</sup>

<sup>1</sup>Department of Endocrinology and Metabolism, Shanghai Clinical Center for Diabetes, Shanghai Diabetes Institute, Shanghai Key Laboratory of Diabetes Mellitus, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai 200233, China;

<sup>2</sup>Chronic Disease Epidemiology, Pennington Biomedical Research Center, Baton Rouge, LA, USA.

## Abstract

Electronic health (medical) records, which are also considered as patients' information that are routinely collected, provide a great chance for researchers to develop an epidemiological understanding of disease. Electronic health records systems cannot develop without the advance of computer industries. While conducting clinical trials that are always costly, feasible and reasonable analysis of routine patients' information is more cost-effective and reflective of clinical practice, which is also called real world study. Real world studies can be well supported by big data in healthcare industry. Real world studies become more and more focused and important with the development of evidence-based medicine. These big data will definitely help in making decisions, making policies and guidelines, monitoring of effectiveness and safety on new drugs or technologies. Extracting, cleaning, and analyzing such big data will be a great challenge for clinical researchers. Successful applications and developments of electronic health record in western countries (eg, disease registries, health insurance claims, etc) have provided a clear direction for Chinese researchers. However, it is still at primary stages in China. This review tries to provide a full perspective on how to translate the electronic health records into scientific achievements, for example, among patients with diabetes. As a summary in the end, resource sharing and collaborations are highly recommended among hospitals and healthcare groups.

**Keywords:** Electronic health records; Real world; Cohort study

## Introduction

In epidemiological studies, randomized controlled trials (RCTs) and cohort studies are commonly used to illustrate the associations between interventions or exposures and certain outcomes. While population-based cohorts are well recognized, hospital-based cohorts are usually considered less useful due to complicated data cleaning, data inaccuracy, and lack of behavioral variables. In the past decades, technical evolution and developments have gradually changed the paper health records to electronic health records,<sup>[1,2]</sup> bringing the possibility to simplify the data cleaning process by using structured templates. It is worth mentioning that, in terms of using RCTs and general population-based cohort studies to assess research questions, especially among patients with diabetes, several limitations should be considered. Firstly, the relatively small sample size, short follow-up time, and a low incidence of outcomes may limit the statistical power. Secondly, strict inclusion and exclusion criteria in RCTs

limit their applicability to the broader population of patients with type 2 diabetes seen in clinical practice. In addition, volunteers in RCTs and traditional cohort studies are typically healthier and more compliant with diabetes self-control recommendations than their counterparts who do not participate in any research, which may cause selection bias. Finally, most cohort studies only assess a single baseline measure of the exposure of interest (eg, glycated hemoglobin A1c [HbA1c], blood pressure, low-density lipoprotein [LDL] cholesterol, body mass index [BMI], estimated glomerular filtration rate [eGFR], plasma glucose, etc), which may also produce potential bias. However, hospital-based cohort studies, derived from administrative databases and electronic health records, especially for patients with diabetes who are seeking healthcare and not participating in a research study, may represent the population more accurately.<sup>[3]</sup> There may be some unique features included in this kind of cohort studies such as (1) reflecting everyday clinical practice; (2) having large study samples with different races and a range of

### Access this article online

Quick Response Code:



Website:

www.cmj.org

DOI:

10.1097/CM9.0000000000000784

**Correspondence to:** Jian Zhou, Department of Endocrinology and Metabolism, Shanghai Clinical Center for Diabetes, Shanghai Diabetes Institute, Shanghai Key Laboratory of Diabetes Mellitus, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, 600 Yishan Road, Shanghai 200233, China  
E-Mail: zhoujian@sjtu.edu.cn

Copyright © 2020 The Chinese Medical Association, produced by Wolters Kluwer, Inc. under the CC-BY-NC-ND license. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Chinese Medical Journal 2020;133(10)

Received: 14-11-2019 Edited by: Li-Shao Guo

health behaviors and outcomes; (3) with low cost; (4) linking to contextual data, which combined self-reported data to address questions and multiple measures of cardiovascular risk factors and treatment information captured over time.

An electronic health record refers to the systematized collection of patient and population electronically stored health information in a digital format. These records may include a range of data, such as demographics (age, race, and sex), health behavior (tobacco and alcohol consumption), medical history (past history of diseases), medication and allergies, immunization status, laboratory test results, radiology images, vital signs, disease codes, personal statistics such as height, weight, waist circumference and hip circumference, and billing information (insured or self-paid). The information is fully recorded in the system. As one of the advantages, patients with certain chronic disease in regular contacts with the healthcare system will definitely have multiple records so that the variability of measurements can be tracked. However, the format of records may vary among different data sets and different systems. In China, electronic health records started at early 2000s.<sup>[4]</sup> Technical workers in computer science along with healthcare provider are working hard to develop software that will be suitable and easy to use under the healthcare system in China. In other words, electronic health records are still at primary stage in China. For chronic diseases such as diabetes, long-term treatment, education, and prevention should be well-integrated. This comprehensive mode relies on a sophisticated content management system that both healthcare providers and patients can be interactively involved in. Tiered medical service networks are still not well covered in China, which will potentially affect the wide application of electronic health records.

For patients with diabetes, several important relationships between intermediate outcomes and risk factors and final outcomes (death, loss of functional status) have been established. Those relationships have led to emphasis on achieving control of intermediate outcomes and risk factors, including glucose levels, lipid levels, blood pressure, renal function, retinopathy, peripheral neuropathy, obesity status, and tobacco use, so as to reduce the probability of onset (or early onset) of loss of functional status and death. These intermediate-to-final outcomes relationships also have led to an emphasis on appropriate monitoring of intermediate outcomes; that is, the need for routine periodic assessment. As a result it is now good clinical practice to monitor (measure) both levels of intermediate outcomes (via outcomes measures), and frequency at which those intermediate outcomes are periodically assessed (via process measures).

Further, a number of derivative measures are of clinical interest. Such derivative measures as some well-known risk factors can reflect the degree of intermediate outcomes to which patients would achieve. Clinicians are also interested to know whether patients can sustain improvement in those intermediate outcomes or risk factors or not. Also of interest are utilization-based measures that are indicative of potential disease control deficiencies, in particular,

emergency department visits and inpatient admissions for acute glycemic control-related events. In this review, we will describe several well-established cohorts of patients with diabetes and their major strategies on how to analyze the associations between risk factors and outcomes.

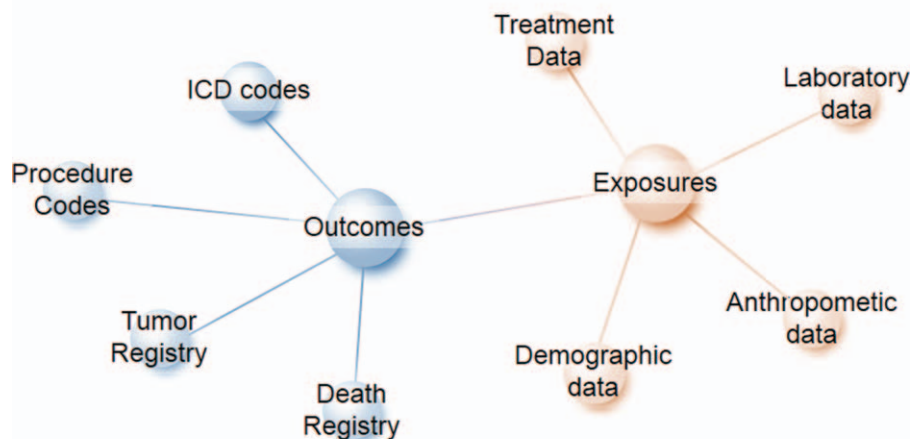
### Use of Electronic Health Records in Europe

There are some well-established national diabetes registries in some European countries. For example, there are about 40,000 patients with type 1 diabetes and more than 450,000 patients with type 2 diabetes included in the Swedish National Diabetes Register (NDR).<sup>[5]</sup> This registry, initiated in 1996, includes information on risk factors, complications of diabetes, and medications for patients 18 years of age or older.<sup>[6]</sup> Factors registered in the Swedish NDR include blood pressure, HbA1c, lipids, BMI, diabetes complications (eg, micro-albuminuria and retinopathy), and treatments.<sup>[6]</sup> This national registry allows researchers to establish retrospective cohorts of high quality. In addition, connections between national diabetes registry and national death registry can also be made. Associations between several risk factors and mortality or other outcomes were thus revealed and convinced. For example, by using the data from Swedish NDR, a cohort with 271,174 patients with type 2 diabetes was established.<sup>[7]</sup> They identified five strongest risk factors in terms of the risk of death among patients with type 2 diabetes including smoking, physical activity, marital status, HbA1c level, and use of statins. Associations between risk factors and other outcomes such as acute myocardial infarction, stroke, and hospitalization for heart failure were also analyzed. The authors concluded that patients with type 2 diabetes and five selected risk-factor variables within target range had, at most, marginally higher risks of death, stroke, and myocardial infarction than the general population. This retrospective cohort study using electronic health records provided a brief impression on the associations between risk factors and outcomes and supported the perspective that risk factors with target range are less likely associated with excess risk of death and cardiovascular outcomes when compared with the general population. However, the authors did not analyze the associations between variability and changes of risk factors with outcomes, which may also be of great interest in this kind of cohort study.

### Use of Electronic Health Records in the United States

#### *Patient-centered outcomes research networks*

European countries such as Sweden are good at national registries mainly because of the limited national population density. However, in the United States, the use of electronic health records may be limited by the conflicts among different systems and medical claims. For example, the Patient-Centered Outcomes Research Institute (PCORI) is an independent, non-profit, non-governmental organization authorized under the Affordable Care Act of 2010 to help close gaps in research evidence that are needed to optimize patient and clinician decision-making and improve health outcomes. To enhance the nation's capacity to conduct comparative effectiveness research, PCORI has invested



**Figure 1:** Common data model of electronic health records. ICD: International Classification of Diseases.

more than \$100 million in the development of the National Patient-Centered Clinical Research Network (PCORnet).<sup>[8]</sup> The centerpiece of the PCORnet initiative is a distributed research network that combines clinical data from electronic health records and data contributed directly by patients from 11 clinical data research networks (CDRNs) and 18 patient-powered research networks located throughout the United States, covering about 100 million population nationwide. The PCORnet common data model is a specification that defines a standard organization and representation of data for the PCORnet distributed research network. Patients' data that can be extracted from this common data model for the present study include date of birth, race/ethnicity, sex, dates of follow-up visits, weight, height, BMI, blood pressure, tobacco use, diagnoses of various diseases and dates of diagnoses, laboratory test dates, total cholesterol, triglycerides, high-density lipoprotein (HDL) cholesterol, LDL cholesterol, HbA1c, eGFR, and medication prescriptions such as anti-hypertensive drugs, glucose-lowering drugs and lipid-lowering drugs.

By using the data from the Research Action for Health Network (REACHnet), one of the partner networks of PCORnet, we have established a cohort of 67,544 patients with type 2 diabetes.<sup>[9]</sup> Associations of several biomarkers including HDL cholesterol<sup>[10]</sup> and BMI<sup>[11]</sup> with stroke risk have been reported. We found an inverse association between HDL cholesterol and stroke risk among patients with type 2 diabetes. Patients with type 2 diabetes with lower HDL cholesterol would have a higher risk of both ischemic and hemorrhagic stroke. Interestingly, BMI also appeared to be inversely associated with stroke risk among patients with type 2 diabetes. This finding may support the hypothesis called obesity paradox.

### **Kaiser permanente**

Another large-scaled health coverage entity is the Kaiser Permanente, which currently serves 12.3 million members in 8 states and the District of Columbia in the United States.<sup>[12]</sup> Diabetes and cardiovascular diseases are active area of study for Kaiser Permanente research. Their rich and comprehensive longitudinal data have been used to

advance understanding of risk, improving outcomes and translating findings into policies and practice.<sup>[13]</sup> Kaiser Permanente was proud to be one of the study sites in the ACCORD study.<sup>[14]</sup> Kaiser Permanente also participated the SUPREME-DM cohort with more than 1 million patients with diabetes.<sup>[15]</sup> Electronic health records were widely used by Kaiser Permanente researchers. For example, they can report the annual rates of certain outcomes by using International Classification of Diseases (ICD) codes [Figure 1]. They can also look into the risk factors that can affect the outcomes by simply establishing a retrospective cohort.<sup>[16]</sup> By using longitudinal data, they can also assess the association between changes of risk factors and outcomes.<sup>[17-19]</sup> In addition, with the wide range of age, BMI, and other demographics, as well as rich data in race and ethnic, various subgroup analysis and sensitivity analysis can be conducted.<sup>[20,21]</sup> By using electronic health records from Kaiser Permanente, researchers were also able to establish a large cohort with patients receiving bariatric surgery.<sup>[22-25]</sup> Studies have shown that, particularly for people who are less severely obese, bariatric surgery can result in diabetes remission and a host of related benefits, including improved life expectancy.<sup>[25]</sup> Even for people who experience a relapse of diabetes after a period of remission, the remission has been linked to long-term health benefits, such as reduced risk of microvascular complications of diabetes.<sup>[22]</sup> Women with gestational diabetes and their offspring are also well studied within Kaiser Permanente healthcare systems. Women with prior gestational diabetes had a higher risk of neonatal macrosomia,<sup>[26,27]</sup> childhood obesity,<sup>[28,29]</sup> and development of autism.<sup>[30]</sup>

This fully integrated care and coverage model from Kaiser Permanente allows researchers to continuously contribute generalizable knowledge to diabetes and many other research areas.

### **Louisiana State University hospital-based longitudinal study**

Our group previously led the Louisiana State University (LSU) hospital-based longitudinal study.<sup>[31-34]</sup> LSU Health Care Services Division (LSUHCSD) operates seven public

hospitals and affiliated clinics in different areas in Louisiana, which provide quality medical care to residents of Louisiana, regardless of their income or insurance coverage. Overall, LSUHCSD facilities have served about 1.6 million unique patients (35% of the Louisiana population) since 1997, and approximately 500,000 patients annually. About 46% of LSUHCSD patients qualify for free care, about 10% of patients are self-pay, about 20% of patients are covered by Medicaid, about 14% of patients have Medicare, and about 10% of patients are covered by commercial insurance. The LSU Health Care Services Division has access to the administrative data, the anthropometric data, the laboratory data, the treatment data, and the clinical diagnosis data of these facilities. All these data are available in electronic form since 1997 for both inpatients and outpatients. A retrospective cohort of patients with type 2 diabetes was thus established to follow all patients with major chronic diseases. With this rich data set, we identified 35,406 patients with type 2 diabetes aged over 30 years old with complete information. We have reported the prevalence of diabetes,<sup>[31]</sup> gestational diabetes,<sup>[32]</sup> and hypertension<sup>[33]</sup> over a 10-year follow-up period. Racial and ethnic disparities in the control status of cardiovascular risk factors were also revealed.<sup>[34]</sup> BMI,<sup>[35-38]</sup> blood pressure,<sup>[39-42]</sup> and HbA1c<sup>[43-47]</sup> were reported to be either positively or negatively correlated with outcomes such as all-cause mortality, and incident coronary heart diseases, heart failure, and stroke. For example, BMI was proved to have a U-shaped association with all-cause mortality risk among black and white patients with type 2 diabetes. A significantly increased risk of all-cause mortality was observed among blacks with BMI <30.0 kg/m<sup>2</sup> and ≥35.0 kg/m<sup>2</sup> compared with patients with BMI of 30.0 to 34.9 kg/m<sup>2</sup>. Similar results were seen among whites with BMI <25.0 kg/m<sup>2</sup> and ≥40.0 kg/m<sup>2</sup> compared with patients with BMI of 30.0 to 34.9 kg/m<sup>2</sup>. BMI was also likely to have a positive association with risks of both coronary diseases and heart failure, especially among men patients with type 2 diabetes. Findings from real world studies are robust. The large sample size allows multiple sensitivity analysis to convince the results. Epidemiology studies using electronic health records can definitely reflect the dose-response associations between risk factors and outcomes.

The electronic health records from LSUHCSD provided researchers a great opportunity to assess the healthcare burden of south Louisiana. The relatively large proportion of African American population in Louisiana also allowed researchers to reveal the racial or ethnic disparities among minorities. State registries on death, tumor, and other outcomes further confirmed the diagnosis of outcomes. The integrated linkage across data sets was particularly essential in establishing cohorts using electronic health records.

### Use of Electronic Health Records in China

Use of electronic health records in China is limited due to a variety of healthcare systems and different criteria. In addition, with the rapid economic development in China, the mobility of population in China within provinces is

extremely high. This phenomenon may cause a rapid flow of patients among different hospitals and it is hard to follow up the outcomes. However, researchers in China also did a good job in the establishment of cohorts using electronic health records. For example, there is a cohort of patients with coronary heart disease established by using the ICD-10 through the Department of Cardiology in three hospitals in Guangzhou between October 2008 and December 2011.<sup>[48]</sup> In this cohort, fasting and 2-h glucose,<sup>[49]</sup> several cardiovascular diseases risk factors and biomarkers such as BMI,<sup>[50]</sup> lipids,<sup>[51]</sup> metabolic syndrome,<sup>[52]</sup> fibroblast growth factor 21,<sup>[53]</sup> C-reactive protein,<sup>[50]</sup> serum microelements,<sup>[54]</sup> and serum vitamin D level<sup>[48]</sup> were reported to be either positively or negatively correlated with mortality among patients with coronary heart diseases.

The electronic health records systems in China were still under development. China has the largest number of patients with diabetes worldwide. There are still no national networks across provinces in China. Collaborations among hospitals are encouraged so that electronic health records can be shared in a proper way.

### Using Electronic Health Records to Design a Target Trial

Electronic health records had certain advantages in illustrating questions in the clinical practice. Obviously, limitation should also be fully considered. In the past 5 years, researches using electronic health records have been focused on the cohort studies. However, cohort studies cannot determine the comparative effectiveness of a certain clinical intervention such as drugs or any new techniques, while electronic health records with a large sample size, a long follow-up, and including individuals at low to moderate risk have the capability to demonstrating such questions, which are often considered as causal effects. These causal effects are rather important for patients, clinicians, and other stakeholders to develop clinical guidelines and support decisions making. While conducting a trial is costly, using electronic health records to emulate the trial design becomes more and more feasible and reasonable.

In general, like a proposal of a clinical trial, several components should be fully considered before we can precede our analysis, including eligibility criteria, treatment strategies, treatment assignment, start and end of follow-up, outcomes, causal contrasts, and statistical analysis plan.<sup>[55]</sup> Firstly, to identify the individuals in the data set who will meet the eligibility criteria, information on age, baseline date when a disease was firstly diagnosed, prescriptions, hospital encounters, lab measurements, at least 1 year of pre-baseline data, and pregnancy status should be extracted. Secondly, determination of initial assignment of the clinical intervention is needed. Unfortunately, randomizations cannot be realized with electronic health records. To tackle the bias caused by the un-randomizations, all important confounding factors should be adjusted and matched usually by using propensity score matching, which will to a large extent minimize the difference between the intervention group and the control group. Thirdly, the definition of follow-up period is important. Follow-ups usually starts at the first

prescription or launch of the clinic intervention and ends at either death, the diagnosis of the outcome, the loss of the follow-up or the end date of the whole data set. Outcomes are usually defined by ICD codes. One important thing that should be pointed out is how to handle the missing data. Individuals with missing data at baseline are not allowed into the analysis, while missing data during follow-up can be handled by multivariate imputation.

Using electronic health records, several famous trials have been emulated such as CAROLINA trial,<sup>[56]</sup> EMPA-REG OUTCOME trial,<sup>[57]</sup> and LEADER trial,<sup>[57]</sup> Electronic health records containing all the key elements above are rare to find. Many Chinese healthcare systems have high rates of loss of follow-up because of the high mobility of the population. They often switch visiting doctors across different hospitals. An integrated system that combines all information including insurance claims, electronic medical records, and pharmacy records seems to be the best option for emulation of a target trial on clinic interventions and outcomes.

### Limitations of Electronic Health Records

Although there are obvious advantages and strength in using electronic health records, several limitations should also be addressed. Firstly, data cleaning is the first step and the most important step when electronic health records are used. Inaccurate records and misinput of anthropometric and lab measurements are commonly seen within the data sets. Inconsistent records within individuals should also be considered seriously. Secondly, validations of diagnosis or chart review are always necessary because misdiagnosis sometimes occurs. In addition, it is very usual, that in routine clinical visits, some patients are not available for every measure. Still it is very important to address the problem of missing data. During analysis, either choosing a statistical procedure that can handle missing (eg, a mixed model approach), checking for possible biases (eg, by repeating the analysis with imputed data sets) or at least discussing why this was not done and was not necessary are usually applied. Finally, major confounding factors are often incorporated into the electronic health records, while there are still some unmeasured factors such as family history of diabetes, dietary factors, and physical activity status which are seldomly recorded in the system. Meanwhile, patients are always switching among hospitals and systems. A unique identifier is also necessary for duplicate patients.

### Summary

Electronic health records have great potential as a data source for clinical researchers to use. Development in computer science will definitely help improving the current healthcare databases. Resource sharing and collaborations are highly recommended among hospitals and healthcare groups. With the rich longitudinal data in electronic health records, linking changes or variability of risk factors with outcomes will be of great interests as a future direction. Especially for chronic disease such as diabetes, registries at national level in China will help healthcare professionals with their routine management. Findings from the real

world data will also help policy makers optimizing the structure and bringing benefits to patients. Notably, a project called the National Metabolic Management Center launched by the Chinese Association of Clinical Endocrinologists<sup>[58]</sup> is leading the exploration in China. Consistent data format within healthcare networks as well as data sharing policies are important during the development of electronic medical systems. It is also necessary and recommended to establish a national center supporting the management and routine maintenance of the data. The absolutely rich population resources will definitely provide huge and considerable evidence on clinical practice. Researches using electronic health records will become a more cost-effective way to address clinical questions and will be widely applied and accepted in the near future.

### Conflicts of interest

None.

### References

- Kern LM, Edwards A, Kaushal R. The patient-centered medical home, electronic health records, and quality of care. *Ann Intern Med* 2014;160:741–749. doi: 10.7326/m13-1798.
- Eggleston EM, Klompas M. Rational use of electronic health records for diabetes population management. *Curr Diab Rep* 2014;14:479. doi: 10.1007/s11892-014-0479-z.
- Ali SM, Giordano R, Lakhani S, Walker DM. A review of randomized controlled trials of medical record powered clinical decision support system to improve quality of diabetes care. *Int J Med Inform* 2016;87:91–100. doi: 10.1016/j.ijmedinf.2015.12.017.
- Wu T, Xu K, Li P, Li XF, Xu WG. The model of “taking electronic medical records as the core for information construction in hospitals”. *Chin Med J* 2013;126:373–377. doi: 10.3760/cma.j.issn.0366-6999.20121787.
- Rawshani A, Rawshani A, Franzen S, Eliasson B, Svensson AM, Miftaraj M, *et al.* Mortality and cardiovascular disease in type 1 and type 2 diabetes. *N Engl J Med* 2017;376:1407–1418. doi: 10.1056/NEJMoa1608664.
- Lind M, Bounias I, Olsson M, Gudbjornsdottir S, Svensson AM, Rosengren A. Glycaemic control and incidence of heart failure in 20,985 patients with type 1 diabetes: an observational study. *Lancet* 2011;378:140–146. doi: 10.1016/s0140-6736(11)60471-6.
- Rawshani A, Rawshani A, Franzen S, Sattar N, Eliasson B, Svensson AM, *et al.* Risk factors, mortality, and cardiovascular outcomes in patients with type 2 diabetes. *N Engl J Med* 2018;379:633–644. doi: 10.1056/NEJMoa1800256.
- Fleurence RL, Curtis LH, Califf RM, Platt R, Selby JV, Brown JS. Launching PCORnet, a national patient-centered clinical research network. *J Am Med Inform Assoc* 2014;21:578–582. doi: 10.1136/amiajnl-2014-002747.
- Shen Y, Shi L, Nauman E, Katzmarzyk PT, Price-Haywood EG, Yin P, *et al.* Race and sex differences in rates of diabetic complications. *J Diabetes* 2019;11:449–456. doi: 10.1111/1753-0407.12869.
- Shen Y, Shi L, Nauman E, Katzmarzyk PT, Price-Haywood EG, Bazzano AN, *et al.* Inverse association between HDL (high-density lipoprotein) cholesterol and stroke risk among patients with type 2 diabetes mellitus. *Stroke* 2019;50:291–297. doi: 10.1161/strokeaha.118.023682.
- Shen Y, Shi L, Nauman E, Katzmarzyk PT, Price-Haywood EG, Bazzano AN, *et al.* Association between body mass index and stroke risk among patients with type 2 diabetes. *J Clin Endocrinol Metab* 2020;105:96–105. doi: 10.1210/clinem/dgz032.
- Reed M, Huang J, Graetz I, Brand R, Hsu J, Fireman B, *et al.* Outpatient electronic health records and the clinical care and outcomes of patients with diabetes mellitus. *Ann Intern Med* 2012;157:482–489. doi: 10.7326/0003-4819-157-7-201210020-00004.
- Krieger N. Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. *Am J Public Health* 1992;82:703–710. doi: 10.2105/ajph.82.5.703.

14. Gerstein HC, Miller ME, Byington RP, Goff DC Jr, Bigger JT, Buse JB, *et al.* Effects of intensive glucose lowering in type 2 diabetes. *N Engl J Med* 2008;358:2545–2559. doi: 10.1056/NEJMoa0802743.
15. Schmittديل JA, Raebel MA, Dyer W, Xu S, Goodrich GK, Schroeder EB, *et al.* Prescription medication burden in patients with newly diagnosed diabetes: a SURveillance, PREvention, and ManagEMENT of Diabetes Mellitus (SUPREME-DM) study. *J Am Pharm Assoc* (2003) 2014;54:374–382. doi: 10.1331/JAPhA.2014.13195.
16. Lawrence JM, Lukacz ES, Liu IL, Nager CW, Lubner KM. Pelvic floor disorders, diabetes, and obesity in women: findings from the Kaiser Permanente Continence Associated Risk Epidemiology Study. *Diabetes Care* 2007;30:2536–2541. doi: 10.2337/dc07-0262.
17. Nichols GA, Hillier TA, Brown JB. Normal fasting plasma glucose and risk of type 2 diabetes diagnosis. *Am J Med* 2008;121:519–524. doi: 10.1016/j.amjmed.2008.02.026.
18. Cooper-DeHoff RM, Bird ST, Nichols GA, Delaney JA, Winterstein AG. Antihypertensive drug class interactions and risk for incident diabetes: a nested case-control study. *J Am Heart Assoc* 2013;2:e000125. doi: 10.1161/jaha.113.000125.
19. Tinker LF, Bonds DE, Margolis KL, Manson JE, Howard BV, Larson J, *et al.* Low-fat dietary pattern and risk of treated diabetes mellitus in postmenopausal women: the Women's Health Initiative randomized controlled dietary modification trial. *Arch Intern Med* 2008; 168:1500–1511. doi: 10.1001/archinte.168.14.1500.
20. Huang ES, Karter AJ, Danielson KK, Warton EM, Ahmed AT. The association between the number of prescription medications and incident falls in a multi-ethnic population of adult type-2 diabetes patients: the diabetes and aging study. *J Gen Intern Med* 2010;25:141–146. doi: 10.1007/s11606-009-1179-2.
21. Lawrence JM, Contreras R, Chen W, Sacks DA. Trends in the prevalence of preexisting diabetes and gestational diabetes mellitus among a racially/ethnically diverse population of pregnant women, 1999–2005. *Diabetes Care* 2008;31:899–904. doi: 10.2337/dc07-2345.
22. Coleman KJ, Haneuse S, Johnson E, Bogart A, Fisher D, O'Connor PJ, *et al.* Long-term microvascular disease outcomes in patients with type 2 diabetes after bariatric surgery: evidence for the legacy effect of surgery. *Diabetes Care* 2016;39:1400–1407. doi: 10.2337/dc16-0194.
23. Arterburn DE, Bogart A, Sherwood NE, Sidney S, Coleman KJ, Haneuse S, *et al.* A multisite study of long-term remission and relapse of type 2 diabetes mellitus following gastric bypass. *Obes Surg* 2013;23:93–102. doi: 10.1007/s11695-012-0802-1.
24. Arterburn D, Bogart A, Coleman KJ, Haneuse S, Selby JV, Sherwood NE, *et al.* Comparative effectiveness of bariatric surgery vs. nonsurgical treatment of type 2 diabetes among severely obese adults. *Obes Res Clin Pract* 2013;7:e258–e268. doi: 10.1016/j.orcp.2012.08.196.
25. Schauer DP, Arterburn DE, Livingston EH, Coleman KJ, Sidney S, Fisher D, *et al.* Impact of bariatric surgery on life expectancy in severely obese patients with diabetes: a decision analysis. *Ann Surg* 2015;261:914–919. doi: 10.1097/sla.0000000000000907.
26. Black MH, Sacks DA, Xiang AH, Lawrence JM. The relative contribution of prepregnancy overweight and obesity, gestational weight gain, and IADPSG-defined gestational diabetes mellitus to fetal overgrowth. *Diabetes Care* 2013;36:56–62. doi: 10.2337/dc12-0741.
27. Sacks DA, Black MH, Li X, Montoro MN, Lawrence JM. Adverse pregnancy outcomes using The International Association of the Diabetes and Pregnancy Study Groups Criteria: glycemic thresholds and associated risks. *Obstet Gynecol* 2015;126:67–73. doi: 10.1097/aog.0000000000000865.
28. Ehrlich SF, Rosas LG, Ferrara A, King JC, Abrams B, Harley KG, *et al.* Pregnancy glucose levels in women without diabetes or gestational diabetes and childhood cardiometabolic risk at 7 years of age. *J Pediatr* 2012;161:1016–1021. doi: 10.1016/j.jpeds.2012.05.049.
29. Ferrara A, Weiss NS, Hedderston MM, Quesenberry CP Jr, Selby JV, Ergas IJ, *et al.* Pregnancy plasma glucose levels exceeding the American Diabetes Association thresholds, but below the National Diabetes Data Group thresholds for gestational diabetes mellitus, are related to the risk of neonatal macrosomia, hypoglycaemia and hyperbilirubinaemia. *Diabetologia* 2007;50:298–306. doi: 10.1007/s00125-006-0517-8.
30. Xiang AH, Wang X, Martinez MP, Walthall JC, Curry ES, Page K, *et al.* Association of maternal diabetes with autism in offspring. *JAMA* 2015;313:1425–1434. doi: 10.1001/jama.2015.2707.
31. Li W, Wang Y, Chen L, Horswell R, Xiao K, Besse J, *et al.* Increasing prevalence of diabetes in middle or low income residents in Louisiana from 2000 to 2009. *Diabetes Res Clin Pract* 2011;94:262–268. doi: 10.1016/j.diabres.2011.08.010.
32. Wang Y, Chen L, Xiao K, Horswell R, Besse J, Johnson J, *et al.* Increasing incidence of gestational diabetes mellitus in Louisiana, 1997–2009. *J Womens Health (Larchmt)* 2012;21:319–325. doi: 10.1089/jwh.2011.2838.
33. Zhang Y, Li W, Wang Y, Chen L, Horswell R, Xiao K, *et al.* Increasing prevalence of hypertension in low income residents within Louisiana State University Health Care Services Division Hospital System. *Eur J Intern Med* 2012;23:e179–e184. doi: 10.1016/j.ejim.2012.08.015.
34. Wang Y, Katzmarzyk PT, Horswell R, Zhao W, Li W, Johnson J, *et al.* Racial disparities in cardiovascular risk factor control in an underinsured population with Type 2 diabetes. *Diabet Med* 2014;31:1230–1236. doi: 10.1111/dme.12470.
35. Zhao W, Katzmarzyk PT, Horswell R, Wang Y, Li W, Johnson J, *et al.* Body mass index and the risk of all-cause mortality among patients with type 2 diabetes mellitus. *Circulation* 2014;130:2143–2151. doi: 10.1161/circulationaha.114.009098.
36. Li N, Katzmarzyk PT, Horswell R, Zhang Y, Li W, Zhao W, *et al.* BMI and coronary heart disease risk among low-income and underinsured diabetic patients. *Diabetes Care* 2014;37:3204–3212. doi: 10.2337/dc14-1091.
37. Li W, Katzmarzyk PT, Horswell R, Zhang Y, Zhao W, Wang Y, *et al.* Body mass index and stroke risk among patients with type 2 diabetes mellitus. *Stroke* 2015;46:164–169. doi: 10.1161/strokeaha.114.006718.
38. Li W, Katzmarzyk PT, Horswell R, Zhang Y, Wang Y, Johnson J, *et al.* Body mass index and heart failure among patients with type 2 diabetes mellitus. *Circ Heart Fail* 2015;8:455–463. doi: 10.1161/circheartfailure.114.001837.
39. Zhao W, Katzmarzyk PT, Horswell R, Wang Y, Johnson J, Cefalu WT, *et al.* Blood pressure and stroke risk among diabetic patients. *J Clin Endocrinol Metab* 2013;98:3653–3662. doi: 10.1210/jc.2013-1757.
40. Zhao W, Katzmarzyk PT, Horswell R, Wang Y, Li W, Johnson J, *et al.* Aggressive blood pressure control increases coronary heart disease risk among diabetic patients. *Diabetes Care* 2013;36:3287–3296. doi: 10.2337/dc13-0189.
41. Zhao W, Katzmarzyk PT, Horswell R, Li W, Wang Y, Johnson J, *et al.* Blood pressure and heart failure risk among diabetic patients. *Int J Cardiol* 2014;176:125–132. doi: 10.1016/j.ijcard.2014.06.051.
42. Li W, Katzmarzyk PT, Horswell R, Wang Y, Johnson J, Hu G. Blood pressure and all-cause mortality among patients with type 2 diabetes. *Int J Cardiol* 2016;206:116–121. doi: 10.1016/j.ijcard.2016.01.009.
43. Zhao W, Katzmarzyk PT, Horswell R, Wang Y, Johnson J, Heymsfield SB, *et al.* HbA1c and lower-extremity amputation risk in low-income patients with diabetes. *Diabetes Care* 2013;36:3591–3598. doi: 10.2337/dc13-0437.
44. Zhao W, Katzmarzyk PT, Horswell R, Wang Y, Johnson J, Hu G. HbA1c and heart failure risk among diabetic patients. *J Clin Endocrinol Metab* 2014;99:E263–E267. doi: 10.1210/jc.2013-3325.
45. Zhao W, Katzmarzyk PT, Horswell R, Wang Y, Johnson J, Hu G. HbA1c and coronary heart disease risk among diabetic patients. *Diabetes Care* 2014;37:428–435. doi: 10.2337/dc13-1525.
46. Zhao W, Katzmarzyk PT, Horswell R, Wang Y, Johnson J, Hu G. Sex differences in the risk of stroke and HbA(1c) among diabetic patients. *Diabetologia* 2014;57:918–926. doi: 10.1007/s00125-014-3190-3.
47. Li W, Katzmarzyk PT, Horswell R, Wang Y, Johnson J, Hu G. HbA1c and all-cause mortality risk among patients with type 2 diabetes. *Int J Cardiol* 2016;202:490–496. doi: 10.1016/j.ijcard.2015.09.070.
48. Yu C, Xue H, Wang L, Chen Q, Chen X, Zhang Y, *et al.* Serum bioavailable and free 25-hydroxyvitamin D levels, but not its total level, are associated with the risk of mortality in patients with coronary artery disease. *Circ Res* 2018;123:996–1007. doi: 10.1161/circresaha.118.313558.
49. Ding D, Qiu J, Li X, Li D, Xia M, Li Z, *et al.* Hyperglycemia and mortality among patients with coronary artery disease. *Diabetes Care* 2014;37:546–554. doi: 10.2337/dc13-1387.

50. Ding D, Wang M, Su D, Hong C, Li X, Yang Y, *et al.* Body mass index, high-sensitivity C-reactive protein and mortality in Chinese with coronary artery disease. *PLoS One* 2015;10:e0135713. doi: 10.1371/journal.pone.0135713.
  51. Ding D, Li X, Qiu J, Li R, Zhang Y, Su D, *et al.* Serum lipids, apolipoproteins, and mortality among coronary artery disease patients. *Biomed Res Int* 2014;2014:709756. doi: 10.1155/2014/709756.
  52. Chen Q, Zhang Y, Ding D, Li D, Xia M, Li X, *et al.* Metabolic syndrome and its individual components with mortality among patients with coronary heart disease. *Int J Cardiol* 2016;224:8–14. doi: 10.1016/j.ijcard.2016.08.324.
  53. Li Q, Zhang Y, Ding D, Yang Y, Chen Q, Su D, *et al.* Association between serum fibroblast growth factor 21 and mortality among patients with coronary artery disease. *J Clin Endocrinol Metab* 2016;101:4886–4894. doi: 10.1210/jc.2016-2308.
  54. Chen Q, Zhang Y, Ding D, Li D, Yang Y, Li Q, *et al.* Associations between serum calcium, phosphorus and mortality among patients with coronary heart disease. *Eur J Nutr* 2018;57:2457–2467. doi: 10.1007/s00394-017-1518-8.
  55. Hernan MA, Robins JM. Per-protocol analyses of pragmatic trials. *N Engl J Med* 2017;377:1391–1398. doi: 10.1056/NEJMs1605385.
  56. Patorno E, Schneeweiss S, Gopalakrishnan C, Martin D, Franklin JM. Using real-world data to predict findings of an ongoing phase IV cardiovascular outcome trial: cardiovascular safety of linagliptin versus glimepiride. *Diabetes Care* 2019;42:2204–2210. doi: 10.2337/dc19-0069.
  57. Arnold SV, Inzucchi SE, Tang F, McGuire DK, Mehta SN, Maddox TM, *et al.* Real-world use and modeled impact of glucose-lowering therapies evaluated in recent cardiovascular outcomes trials: an NCDR(R) research to practice project. *Eur J Prev Cardiol* 2017;24:1637–1645. doi: 10.1177/2047487317729252.
  58. Zhang Y, Wang W, Ning G. Metabolic Management Center: an innovation project for the management of metabolic diseases and complications in China. *J Diabetes* 2019;11:11–13. doi: 10.1111/1753-0407.12847.
- 
- How to cite this article:** Shen Y, Zhou J, Hu G. Practical use of electronic health records among patients with diabetes in scientific research. *Chin Med J* 2020;133:1224–1230. doi: 10.1097/CM9.0000000000000784