

# Use of real world data to improve drug coverage decisions in China

**Wen Wang and colleagues** discuss the rationale and propose a framework for using real world evidence to support coverage decisions in Chinese setting

**M**edical expenditure has substantially increased in China in the past decade with the introduction of new drugs with higher prices. Spending on basic medical insurance increased from ¥486bn (£58bn; €66bn; \$70bn) in 2012 to ¥1760bn in 2018. The increased burdens posed great challenges for China's healthcare system. In response, the National Healthcare Security Administration (NHSA) was established in 2018, and a dynamic adjustment mechanism was implemented for the national formulary, known as the national reimbursement drug list (NRDL). The list is reassessed regularly to add new drugs, remove obsolete drugs, or change reimbursement restrictions.

Applications for inclusion in the reimbursement list are open once a year. Renewal of coverage of existing drugs, which includes pricing renegotiation through reassessment of value, is also considered at the same time. The decisions about whether drugs are covered are based on multiple factors, including unmet needs, clinical benefit, and economic value. Since 2018, 507 drugs have been added to the list and 391 drugs removed.<sup>1</sup>

In the past few years, China has implemented expedited procedures to ensure faster access to new drugs. Within

this context, the National Medical Products Administration (NMPA), which licenses products for use in China, adopted several policy amendments to speed up access to drugs newly approved for market. Upon approval, the NHSA would accelerate the inclusion of new drugs in the reimbursement list at the annual review. For instance, 74 drugs were included in the list in 2021, 66 of which were approved by NMPA in 2020.<sup>1</sup> In particular, several oncology drugs were added within six months of regulatory approval.<sup>2</sup> Orphan drugs are also prioritised for inclusion.<sup>3</sup>

## Gaps in evidence for coverage decisions

Substantial challenges arise in assessing the value of drugs to inform coverage decisions. One problem is inadequate clinical evidence. Traditionally, the assessment process largely relies on evidence derived from randomised controlled trials, which are typically designed with strict patient selection criteria, implemented with reinforced compliance to treatments, and have relatively short follow-up. The drug performance in randomised trials may differ from that in the real world. Furthermore, the trials are often multicentre international trials in populations that might not fully represent the Chinese population.

Obtaining clinical evidence is particularly challenging for oncology and orphan drugs because they are often approved through expedited review pathways and use data from single arm trials. For instance, 52.5% of new drugs were approved through expedited pathways from 2017 to 2020 in China.<sup>4</sup> Consequently, the effectiveness and comparative effectiveness was often uncertain at the time of approval. Moreover, these trials often had relatively short follow-up periods and evaluated short term treatment outcomes. This raises concerns about the clinical benefits of the drugs, especially in the long term.

Lack of clinical evidence is the tip of the iceberg. Some of the data required for economic modelling are not available from randomised trials, or the evidence is often weak or lacking. In many cases, for example, the data on disease burden and

treatment patterns are primarily based on data from populations outside China,<sup>5,6</sup> where the dose and course of treatments may differ from those in other countries.<sup>7,8</sup>

Economic models are commonly used to incorporate the data from multiple sources to produce cost effectiveness evidence to inform coverage decisions. However, the models are often criticised for lacking transparency and robustness, especially when locally relevant data are limited.<sup>9</sup> Indeed, the lack of locally relevant data can substantially compromise the economic models as well as affecting estimations of budget impact. The expected usage, duration, and dosage of drugs in highly controlled trial settings are likely to differ from those in routine practice.

## Use of real world data for coverage decisions

In the past several years, the extensive use of information technologies in healthcare and public health practices has boosted the generation of diverse real world data sources in China, including both national and local data sources.<sup>10</sup> Electronic medical records, patient registries, and claims databases are now well established across the country. Together with sound epidemiological designs and robust analyses, these data can provide locally applicable evidence for assessing drug value.<sup>11-16</sup> Typically, these data are observational in nature and are used in observational study designs. However, randomised trials that are built on real world settings and make use of the strengths of these data systems are more able to control for different sources of biases than observational studies. Consequently, these randomised trials may offer better real world evidence about the value of drugs.<sup>17</sup>

In China, real world data may provide an array of local evidence to support decisions about whether a drug should be included in the reimbursement list and whether it should be renewed.

## Use of real world data to support initial coverage decisions

Decisions about including a drug in the list for the first time require extensive and

### KEY MESSAGES

- Important gaps exist in evidence to support decisions on which drugs should be reimbursed in China
- Real world data collected in routine practice could provide useful information on the value of drugs both for initial coverage decisions and renewal
- Data from use of new drugs in the Boao Lecheng pilot zone could also provide early information on Chinese context
- Policy, academic, and technical forces are becoming available to improve the quality and relevance of data available to guide coverage decisions

**Table 1 | Framework for using real world data to supporting initial drug coverage decisions in China**

Evidence needed	Explanation	Likelihood of real world data	Data sources	Example
Disease epidemiology	Prevalence or incidence of disease in Chinese population	+++	Claims data, population based database, national epidemiological survey	A study used administrative claims database covering 14 million employees in Beijing, and found that the prevalence of prostate cancer with bone metastasis increased from 5.3 per 100 000 in 2011 and 8.3 per 100 000 in 2014 <sup>18</sup>
Treatment patterns	Routine practice, treatment duration and dose, adherence of treatment	+++	Claims data, EMRs, patient registries	A study evaluated EMRs from 15 healthcare institutions and showed that metformin-containing therapy continued to be the primary treatment option for Chinese patients. Through investigating treatment patterns in current use, decisions makers can assess financial effect of adding a new treatment to current practice <sup>19</sup>
Comparative effectiveness	Effects compared with alternative drugs	+	Registries	A cohort study using an established pregnancy registry showed that women receiving intramuscular motherwort plus intravenous oxytocin had decreased risk of postpartum haemorrhage after caesarean section compared with those receiving intramuscular plus intravenous oxytocin <sup>20</sup>
Safety	Adverse drug reaction, serious adverse events	+	Adverse event surveillance data, EMRs, population based database	A retrospective study using hospital data found the most common adverse event among patients receiving lenvatinib for hepatocellular carcinoma was raised levels of aminotransferases (46.2%). No grade 4 or 5 adverse events occurred <sup>21</sup>
Disease progression	Course of the disease, the incidence and onset time of progression	++	Patient registries, population based database, EMRs	A study in 214 patients in hospital with covid-19 pneumonia showed that 16.1% of moderately ill patients and 47.9% of severe ill patients experienced deterioration during hospital stays <sup>22</sup>
Resource use	Use of healthcare services	++	Claims data, population based database, EMRs	Analysis of claims data for 382 000 patients with lung cancer from China's urban basic medical insurance showed the annual number of outpatient visits per patient was 2.42 and inpatient admissions was 2.07. Western medicine (39.4%) accounted for the highest proportion of total cost; traditional Chinese medicine accounted for 11.9% <sup>23</sup>
Treatment cost	Direct cost and indirect cost associated with treatments	++	Claims data, population based database, EMRs	Using 5% random sample of claims data from China's urban basic medical insurance, a study showed that the average annual direct cost of a stroke was ¥10 637. Medication accounted for 50.6% of total costs, which was higher than other Asian countries. The study suggested an escalating burden on China's health system. <sup>24</sup>

EMR=electronic medical record.

broad local evidence, ranging from disease epidemiology to treatment costs (table 1). Some of this evidence may be readily available from real world data, such as on treatment patterns, healthcare costs, and disease courses. These data are often embedded in electronic medical records and healthcare claims, and studies have shown the potential usefulness of these data for generating evidence.<sup>19 25 26</sup> However, when new drugs are approved under accelerated coverage review, they have not been used in the routine care. Thus, the evidence on efficacy, comparative effectiveness, and safety is often limited, particularly if only single arm trials are available.

One solution is to generate evidence about comparative effectiveness before the initial coverage decision. This would require drug manufacturers to produce evidence from routine use of these drugs. However, this is unlikely to be feasible for all new drugs given the current healthcare security policy in China. We thus propose using real world evidence generated from the Hainan Boao Lecheng international medical tourism pilot zone, known as the Hainan model.<sup>27</sup>

In 2019, the Chinese government issued a policy to accelerate access to new and clinically demanding medical products.

Under this policy, Chinese patients can receive new medical products that have been licensed by major regulatory authorities overseas in Boao's healthcare institutions before the products are approved by the NMPA. The medicines are granted special approval for use in Boao for the indications approved overseas, and patients with a confirmed diagnosis are treated in the routine healthcare setting rather than as part of a trial. After treatment, they return home with discharge medications and are followed up at local medical institutions. Currently, the longest follow-up is three years. The data collected during routine use of the new medical products at Boao and follow-up information generated at local medical institutions are integrated through a patient registry.

Data from Boao can be used to assess the comparative effectiveness of new treatments through observational study designs or pragmatic randomised trials. Although randomised trials overcome the problem of confounding, observational designs may be more appropriate for rare diseases or when an urgent decision is needed.

Using data generated from Boao may have limitations, particularly that the travelling population receiving these new

products may differ from general residents. As a result, the data may not be fully representative of the Chinese population. However, the special policy measures at Boao have created supportive conditions for new medical products to speed up their launch in the China market, and tens of thousands of people have received these new medical products at Boao.

So far real world data from Boao have been used to support regulatory approval without conducting a trial in China. This type of evidence could also be used to support reimbursement decisions to minimise delay to market entry. When no local trial is conducted, the real world data can be used to support regulatory approval and coverage decisions concurrently. In such a scenario, the main contribution of real world evidence is to provide evidence on effectiveness and safety of these drugs. If the registry contains information on clinical care and treatment outcomes of patients receiving standard care as well as those receiving the new drugs, statistical methods can be used to compare the two approaches. Additionally, these data can provide evidence on treatment patterns and costs.

When a local trial has been conducted for regulatory approval, real world data from

Table 2 | Framework for using real world data to support decisions to renew drug coverage in China

Evidence needed	Explanation	Likelihood of real world data	Data sources	Examples
Change of treatment patterns	Prescription behaviour and treatment adherence after approval	+++	Population based database, Claims data, EMRs	A study of 473 patients with acromegaly using patient data reported through an online nationwide platform found the proportion treated with long acting release pasireotide was significantly lower among patients diagnosed from 2015 onwards compared with those diagnosed before 2015. <sup>28</sup>
Updated evidence on comparative effectiveness	Local data related to the use and treatment outcomes accumulated after approval	+	Population based database, patient registries, EMRs	A prospective observational study on a large registry was conducted to evaluate the effectiveness of different basal insulin therapies. Compared with intermediate acting insulin, long acting insulin analogues were associated with better glucose control and decreased risk of hypoglycaemia and weight gain in people with type 2 diabetes. <sup>29</sup>
Serious or long term adverse events	Incidence of rare but serious adverse drug reaction among Chinese population	++	Population based database, patient registries, EMRs, adverse drug reaction monitoring data	A retrospective cohort study using a large EMR database and involving 23 242 patients found that cefoperazone-sulbactam was associated with an increased risk of coagulation disorders. The risk of bleeding among treated patients was 4.2%. <sup>30</sup>
Updated resource use	Changes of use and market diffusion	+++	Claims data, population based database, patient registries, EMRs	Analysis of prescription data from 804 sample hospitals suggested that use of NRDL drugs increased by 25.7% from 2015 to 2021. Oncology drugs were the fastest growing drugs. <sup>2</sup>
Updated treatment cost	Direct cost and indirect cost in routine care	+++	Claims data, population based database, EMRs	An analysis using prescription data from 804 sample hospitals showed that amount spent on oncology drugs increased from 7% of all spending on NRDL drugs in 2015 to 13.9% in 2021. <sup>2</sup>

EMR=electronic medical records.

Boao would be used primarily to support coverage decisions. In this case, real world data from Boao can provide an array of evidence for coverage decisions, including on real world effectiveness and safety, treatment patterns, and costs, which would inform economic modelling.

#### Use of real world data to support renewal

The reimbursement of new drugs often results in important changes in clinical practice, patient outcomes, and use of healthcare resources. In such cases, real world data can provide critical information on decisions about renewals (table 2). As new drugs are increasingly used over time a broader array of evidence may be accumulated. Various data sources are readily available under such circumstances and may cover all aspects of information for renewal. In particular, claims databases may offer unequivocal advantages over electronic medical records in the reassessment of treatment patterns, resources use, and budget impact. Other databases such as electronic medical records and patient registries can be important sources for assessing real world effectiveness and safety.<sup>2 30 31</sup> Notably, pragmatic randomised trials may offer stronger evidence on comparative effectiveness than observational studies, as well as providing data on economic impact to support coverage renewal.

#### Show me the data: improving real world data

Despite the great potential of real world data in informing coverage decisions in China, important challenges remain. Data relevance and quality are probably the pri-

mary obstacles. Available data may not be able to answer questions of interest for decisions, and data inaccuracy and incompleteness are not uncommon.<sup>32-34</sup> Databases from single institutions often have incomplete outpatient data and lack information on patient follow-up, limiting their usefulness for assessing long term effectiveness and safety. In addition, data sharing and access are also limited, making linkage of data from multiple medical institutions and subsequent use less likely. As a result, research questions for coverage decisions may not be well addressed.

Another challenge lies in the limited research capacity to produce trustworthy real world evidence. Data quality is often suboptimal because of lack of transparent and appropriate data collection. These are issues not only about data collection, but also about appropriate use of data for research purposes. Although several advanced methods have been proposed, selecting an appropriate method to adequately disentangle confounding is a challenge.<sup>35 36</sup> Currently, few research groups have received training in methodology. The process of handling data remains obscure. Development of quality standards is a high priority for the production of real world evidence, and transparency also needs to be improved.

A few important initiatives to address these problems have emerged in China. National policies on the development and application of “big” healthcare data have been issued, facilitating the production and application of real world data.<sup>37</sup> Population based healthcare data systems

have been increasingly established in large municipal cities, with data linkage put in place.<sup>25 38</sup> These linked data can provide important information for assessing the value of drugs. A working panel consisting of experts from academic institutions and authorities has also been formed to develop technical guidance documents to improve the transparency and quality of real world evidence. In the future, a real world evidence community should be developed that consists of policy makers, healthcare providers, patients, academic institutions, and industry. Such a community would be likely to result in better use of real world evidence to support coverage decisions in China.

We thank John Cairns for helpful suggestions and comments.

**Contributors and sources:** WW's research focuses on real world data studies and clinical evaluation of drugs. JT has extensive expertise in real world data studies. SYH focuses on health policy. YXH researches real world data and health policy. FX's research interests include health technology assessment and health economics. JW has expertise in pharmacoconomics and health policy. XS is an expert in evidence based medicine and real world data research and is the guarantor of the article. WW drafted and critically revised the manuscript. JT, SYH, YXH, FX, and JW offered critical suggestions and revisions. XS conceptualised, drafted, and critically revised the manuscript.

**Competing interests:** We have read and understood BMJ policy on declaration of interests and declare that the study was supported by the National Key R&D Program of China (Grant No 2017YFC1700406 and 2017YFC1700400), the National Natural Science Foundation of China (Grant No 82225049, 72104155), China Medical Broad (Grant No CMB 19-324), and 1.3.5 project for disciplines of excellence, West China Hospital, Sichuan University (Grant No ZYYC08003). These funders had no role in the study design, collection, analysis, or interpretation of the

data, writing the manuscript, or the decision to submit the paper for publication.

**Provenance and peer review:** Commissioned; externally peer reviewed.

This article is part of a series (info to come)

**Wen Wang**, associate professor<sup>1,2,3</sup>

**Jing Tan**, associate professor<sup>1,2,3</sup>

**Jing Wu**, professor<sup>4</sup>

**Shiyao Huang**, research associate<sup>1,2,3</sup>

**Yunxiang Huang**, PhD candidate<sup>1,2,3</sup>

**Feng Xie**, professor of health economics<sup>5</sup>

**Xin Sun**, professor<sup>1,2,3</sup>

<sup>1</sup>Chinese Evidence-based Medicine Centre, and National Clinical Research Centre for Geriatrics, West China Hospital, Sichuan University, Chengdu, Sichuan, China

<sup>2</sup>NMPA Key Laboratory for Real World Data Research and Evaluation in Hainan, Chengdu, China

<sup>3</sup>Sichuan Center of Technology Innovation for Real World Data, Chengdu, China

<sup>4</sup>School of Pharmacology Science and Technology, Tianjin University, Tianjin, China

<sup>5</sup>Department of Clinical Epidemiology and Biostatistics, University of Macmaster, Hamilton, Canada

**Correspondence to:** Xin Sun  
sunx79@hotmail.com



OPEN ACCESS

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.



- 1 Chinese State Council. 74 drugs to be added to national insurance list. 2021. [http://english.www.gov.cn/statecouncil/ministries/202112/07/content\\_WS61aeb805c6d09c94e48a1cbd.html](http://english.www.gov.cn/statecouncil/ministries/202112/07/content_WS61aeb805c6d09c94e48a1cbd.html)
- 2 Technology Development Center of Chinese Pharmaceutical Association. *Blue book on progress and achievement of medical insurance reform in China*. CPA, 2021.
- 3 Chinese State Council. Over 60 rare-disease drugs approved for Chinese market. 2021. [https://english.www.gov.cn/statecouncil/ministries/202112/19/content\\_WS61be7078c6d09c94e48a26d9.html](https://english.www.gov.cn/statecouncil/ministries/202112/19/content_WS61be7078c6d09c94e48a26d9.html)
- 4 Tang Z, Guo J, Tang J, Ying X. Analysis on the current situation of listed innovative drugs in China from 2017 to 2020. *World Clinical Drugs* 2022;43:294-9.
- 5 Zhou T, Sheng Y, Guan H. Cost-effectiveness of vedolizumab in the treatment of moderate-to-severe Crohn's disease in China. *Adv Ther* 2021;38:4233-45. doi:10.1007/s12325-021-01806-7
- 6 Hu Z, Tao H, Shi A, Pan J. The efficacy and economic evaluation of roxadustat treatment for anemia in patients with kidney disease not receiving dialysis. *Expert Rev Pharmacoecon Outcomes Res* 2020;20:411-8. doi:10.1080/14737167.2020.1747436
- 7 Ni J, Cheng X, Zhao Q, et al. The efficacy and safety of niraparib for ovarian cancer: a single-center observational study from China. *J Ovarian Res* 2021;14:68. doi:10.1186/s13048-021-00803-2

- 8 Food and Drug Administration (FDA). FDA approves niraparib for first-line maintenance of advanced ovarian cancer. 2022. <https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-niraparib-first-line-maintenance-advanced-ovarian-cancer>
- 9 Zhang F, He X, Xiang W, Li SC. Assessment of the quality of pharmacoeconomic evaluation literature in China. *J Med Econ* 2017;20:510-7. doi:10.1080/13696998.2017.1280501
- 10 Chinese State Council. Outline for the national IT application development strategy. 2016. [http://www.gov.cn/zhengce/2016-07/27/content\\_5095336.htm](http://www.gov.cn/zhengce/2016-07/27/content_5095336.htm)
- 11 Justo N, Espinoza MA, Ratto B, et al. Real-world evidence in healthcare decision making: global trends and case studies from Latin America. *Value Health* 2019;22:739-49. doi:10.1016/j.jval.2019.01.014
- 12 Makady A, Ham RT, de Boer A, Hillege H, Klungel O, Goettsch W, GetReal Workpackage 1. Policies for use of real-world data in health technology assessment (HTA): a comparative study of six HTA agencies. *Value Health* 2017;20:520-32. doi:10.1016/j.jval.2016.12.003
- 13 George E. How real-world data compensate for scarce evidence in HTA. *Z Evid Fortbild Qual Gesundheitsw* 2016;112(Suppl 1):S23-6. doi:10.1016/j.zefq.2016.04.012
- 14 Sun X, Tan J, Tang L, Guo JJ, Li X. Real world evidence: experience and lessons from China. *BMJ* 2018;360:j5262. doi:10.1136/bmj.j5262
- 15 Kang J, Cairns J. Protocol for data extraction: how real-world data have been used in the National Institute for Health and Care Excellence appraisals of cancer therapy. *BMJ Open* 2022;12:e055985. doi:10.1136/bmjopen-2021-055985
- 16 Bullement A, Podkonjak T, Robinson MJ, et al. Real-world evidence use in assessments of cancer drugs by NICE. *Int J Technol Assess Health Care* 2020. doi:10.1017/S0266462320000434
- 17 Corrigan-Curay J, Sacks L, Woodcock J. Real-world evidence and real-world data for evaluating drug safety and effectiveness. *JAMA* 2018;320:867-8. doi:10.1001/jama.2018.10136
- 18 Zhuo L, Cheng Y, Pan Y, et al. Prostate cancer with bone metastasis in Beijing: an observational study of prevalence, hospital visits and treatment costs using data from an administrative claims database. *BMJ Open* 2019;9:e028214. doi:10.1136/bmjopen-2018-028214
- 19 Wang C, Gao Y, Zhu L, Huang M, Wu Y, Xuan J. Treatment patterns in patients with newly diagnosed type 2 diabetes in China: a retrospective, longitudinal database study. *Clin Ther* 2019;41:1440-52. doi:10.1016/j.clinthera.2019.05.003
- 20 Li MX, Liu CR, Chen M, et al. Effects of motherwort injection versus intramuscular oxytocin for preventing postpartum hemorrhage among women who underwent cesarean section. *Front Pharmacol* 2022;13:859495. doi:10.3389/fphar.2022.859495
- 21 Zhu Y, Sun P, Wang K, et al. Efficacy and safety of lenvatinib monotherapy and lenvatinib-based combination therapy for patients with unresectable hepatocellular carcinoma: a retrospective, real-world study in China. *Cancer Cell Int* 2021;21:503. doi:10.1186/s12935-021-02200-7
- 22 Liu J, Liu Z, Jiang W, et al. Clinical predictors of COVID-19 disease progression and death: analysis of 214 hospitalised patients from Wuhan, China. *Clin Respir J* 2021;15:293-309. doi:10.1111/crj.13296
- 23 Zhu D, Shi X, Nicholas S, Ma Y, He P. Estimated annual prevalence, medical service utilization and direct costs of lung cancer in urban China. *Cancer Med* 2021;10:2914-23. doi:10.1002/cam4.3845
- 24 Zhu D, Shi X, Nicholas S, et al. Medical service utilization and direct medical cost of stroke in urban China. *Int J Health Policy Manag* 2022;11:277-86.
- 25 Tan J, Xiong Y, Qi Y, et al. Data resource profile: Xiamen registry of pregnant women and offspring (Represent): a population-based, long-term follow-up database linking four major healthcare data platforms. *Int J Epidemiol* 2021;50:27-8. doi:10.1093/ije/dyaa161
- 26 Li C, Bian X, Liu Z, et al. Effectiveness and safety of pyrotinib-based therapy in patients with HER2-positive metastatic breast cancer: a real-world retrospective study. *Cancer Med* 2021;10:8352-64. doi:10.1002/cam4.4335
- 27 Hainan Boao Lecheng international medical tourism pilot zone. About us. 2021. <http://en.lecityhn.com/index.html>
- 28 Guo X, Wang K, Yu S, et al. Patient characteristics, diagnostic delays, treatment patterns, treatment outcomes, comorbidities, and treatment costs of acromegaly in China: a nationwide study. *Front Endocrinol (Lausanne)* 2020;11:610519. doi:10.3389/fendo.2020.610519
- 29 Ji L, Zhang P, Zhu D, et al. ORBIT study group. Comparative effectiveness and safety of different basal insulins in a real-world setting. *Diabetes Obes Metab* 2017;19:1116-26. doi:10.1111/dom.12920
- 30 Wang W, Liu Y, Yu C, et al. Cefoperazone-sulbactam and risk of coagulation disorders or bleeding: a retrospective cohort study. *Expert Opin Drug Saf* 2020;19:339-47. doi:10.1080/14740338.2020.1713090
- 31 Xu R, Wang W, Zhu B, et al. Disease characteristics and treatment patterns of Chinese patients with metastatic colorectal cancer: a retrospective study using medical records from China. *BMC Cancer* 2020;20:131. doi:10.1186/s12885-020-6557-5
- 32 Wallach JD, Zhang AD, Skydel JJ, et al. Feasibility of using real-world data to emulate postapproval confirmatory clinical trials of therapeutic agents granted US Food and Drug Administration accelerated approval. *JAMA Netw Open* 2021;4:e2133667. doi:10.1001/jamanetworkopen.2021.33667
- 33 Bartlett VL, Dhruva SS, Shah ND, Ryan P, Ross JS. Feasibility of using real-world data to replicate clinical trial evidence. *JAMA Netw Open* 2019;2:e1912869. doi:10.1001/jamanetworkopen.2019.12869
- 34 Chiang HY, Liang LY, Lin CC, et al. Electronic medical record-based deep data cleaning and phenotyping improve the diagnostic validity and mortality assessment of infective endocarditis: medical big data initiative of CMUH. *Biomedicine (Taipei)* 2021;11:59-67. doi:10.37796/2211-8039.1267
- 35 Dahabreh IJ, Sheldrick RC, Paulus JK, et al. Do observational studies using propensity score methods agree with randomized trials? A systematic comparison of studies on acute coronary syndromes. *Eur Heart J* 2012;33:1893-901. doi:10.1093/eurheartj/ehs114
- 36 Austin PC, Schuster T. The performance of different propensity score methods for estimating absolute effects of treatments on survival outcomes: a simulation study. *Stat Methods Med Res* 2016;25:2214-37. doi:10.1177/0962280213519716
- 37 Chinese State Council. Guidance on promoting and regulating of the development of health big data. 2016. [http://www.gov.cn/zhengce/content/2016-06/24/content\\_5085091.htm](http://www.gov.cn/zhengce/content/2016-06/24/content_5085091.htm)
- 38 Lin H, Tang X, Shen P, et al. Using big data to improve cardiovascular care and outcomes in China: a protocol for the Chinese electronic health records research in Yinzhou (CHERRY) Study. *BMJ Open* 2018;8:e019698. doi:10.1136/bmjopen-2017-019698

Cite this as: *BMJ* 2023;381:e068911  
<http://dx.doi.org/10.1136/bmj-2021-068911>