Realizing the promise of big data: how Taiwan can help the world reduce medical errors and advance precision medicine

Kevin Wang
Vagelos College of Physicians and Surgeons, Columbia University, New York, New York, USA, and
Peter Alexander Muennig
Department of Health Policy and Management, Mailman School of Public Health, Columbia University, New York, New York, USA

Received 8 November 2021
Revised 6 April 2022
19 April 2022
Accepted 21 April 2022

Abstract

Purpose – The study explores how Taiwan's electronic health data systems can be used to build algorithms that reduce or eliminate medical errors and to advance precision medicine.
Design/methodology/approach – This study is a narrative review of the literature.
Findings – The body of medical knowledge has grown far too large for human clinicians to parse. In theory, electronic health records could augment clinical decision-making with electronic clinical decision support systems (CDSSs). However, computer scientists and clinicians have made remarkably little progress in building CDSSs, because health data tend to be siloed across many different systems that are not interoperable and cannot be linked using common identifiers. As a result, medicine in the USA is often practiced inconsistently with poor adherence to the best preventive and clinical practices. Poor information technology infrastructure contributes to medical errors and waste, resulting in suboptimal care and tens of thousands of premature deaths every year. Taiwan's national health system, in contrast, is underpinned by a coordinated system of electronic data systems but remains underutilized. In this paper, the authors present a theoretical path toward developing artificial intelligence (AI)-driven CDSS systems using Taiwan's National Health Insurance Research Database. Such a system could in theory not only optimize care and prevent clinical errors but also empower patients to track their progress in achieving their personal health goals.
Originality/value – While research teams have previously built AI systems with limited applications, this study provides a framework for building global AI-based CDSS systems using one of the world's few unified electronic health data systems.

Keywords Big data, Artificial intelligence, Electronic clinical decision support systems, Electronic medical records systems, Precision medicine

Paper type Viewpoint

Introduction

In the USA, it has long been recognized that clinicians practice medicine inconsistently [1]. Overt medical errors cause tens of thousands of premature deaths annually, a major preventable public health crisis [2, 3]. Not counted in these numbers are everyday deviations

Author Contributions: Kevin Wang conducted the core research that underpinned the paper and made substantial contributions to the writing of the manuscript. Peter Muennig is the senior author, mentored Kevin Wang and wrote the final version of the manuscript.

Applied Computing and Informatics
Emerald Publishing Limited
e-ISSN: 2210-8327
p-ISSN: 2634-1964
DOI 10.1108/ACI-11-2021-0298

from clinical practice guidelines, which plausibly contribute to even more suffering and death than overt medical errors.

Deviations from best practice likely arise because the body of medical knowledge is simply too large for the human brain to memorize and process. For example, the best practice involves using Bayesian math to guide laboratory testing, allowing clinicians to estimate the probability of making a correct diagnosis [4, 5]. In practice, physicians tend to order a slew of diagnostic tests, leading to false positive test results that sometimes require dangerous diagnostic interventions and treatments [3]. For example, if a test that correctly detects disease 99% of the time is administered to 1,000 people who do not have the disease, ten will have a positive test. Of these ten, some will require further workup, potentially including surgical intervention, which carries the risk of bleeding, infection and death. Such deadly mistakes could be reduced were electronic health record (EHR) systems able to nudge clinicians toward best clinical practices.

But even "best" clinical practices – those that are based on guidelines developed by a panel of experts – are subject to error. That is because the data used to develop guidelines is based on averages of normally distributed data. Treatment responses at the tails of a normal distribution can range from severe harm to better-than-expected outcomes. These outliers result from variations in genetics and environmental exposures. The result is that individuals on the tails of these distributions will not be receiving optimal care when clinical practice guidelines are applied.

Fortunately, the EHR is not only a "front-end" data system that can guide clinicians toward expert consensus on the best available care. It is also a "back-end" system that consists of hundreds of millions of different treatments and hundreds of millions of different responses over time. These variations can be exploited to optimize treatment plans for individuals rather than the "average" patient.

The big data challenge in health

Unfortunately, building a system that allows for either front-end prompts or back-end research is all but impossible in the USA because there are many different data systems, most of which are not interoperable [6]. The result is a fragmented system of "information silos" (multiple sources of small data that are not interconnected). These cannot be tapped for predictive analytics [6–8]. Despite ongoing efforts to standardize and integrate databases, technical, logistical and legal obstacles stand in the way [9, 10]. Technical obstacles include creating a standardized data format that can be used to analyze different EHR systems. It is logistically challenging in that private companies behind EHR systems must communicate with other companies to standardize databases for interoperability. It is legally challenging in that federal laws that require EHR companies to cooperate are difficult to enforce. Beyond the challenge of extracting data from EHR data silos, there is the need to integrate data from diverse billing systems, laboratories, consumer devices, imaging systems and, nearly as importantly, from nonmedical data sources, such as birth, death, income and housing registries.

In contrast, EHRs in Taiwan are compiled into centralized and standardized data systems and integrated with a wide array of additional health and sociological data on patients [11–13]. Taiwan's data systems could, in theory, be used both to encourage better decisions and high-value care in the form of clinical decision support systems (CDSSs), while also advancing precision medicine by developing algorithms that allow us to understand how different individuals respond to different treatments for any given condition when deviations from best practices occur. However, Taiwan is home to only a small fraction of the world's data scientists, and there are fewer commercial interests in exploiting these systems in an island of 23 million people.

Building a better front-end

Clinical decision support systems seek to "improve healthcare delivery by enhancing medical decisions with targeted clinical knowledge, patient information, and other health information" [14, 15]. This concept was first conceived in the mid-20th century [8], and the first computer-based CDSSs emerged in the 1970s [14]. In an era of pen-and-paper medical records, these early systems were burdensome and rarely used despite their potential to save lives [2, 3].

The potential of CDSSs is salient not only because it can store an individual's medical records on servers than can readily be accessed without pulling up paper charts, but also because they can use natural language processing to provide clinicians with feedback on their diagnosis and treatment plans.

CDSSs in the USA can be conceptualized as "small data" because they tend to be siloed. Even some provider networks use multiple EHR systems that are not interoperable, meaning that a patient who was seen by the ophthalmology clinic might not have records that can be accessed when admitted to the emergency department. Nevertheless, given the resources, arduously developed clinical practice guidelines can be integrated into small data networks [8, 14, 15]. These are known as knowledge-based systems [8, 14, 15]. Where this has been achieved, it reduces medication errors and encourages high-value care, but we are far from widespread use of such systems, let alone using them to develop precision medicine algorithms [6–9, 14, 16].

Knowledge-based systems have been used in Taiwan to make provider reimbursement contingent on some aspects of the quality of care [17]. However, these systems require regular manual updating as guidelines change and rely on average responses rather than tailored medicine. On the other hand, non-knowledge-based systems use artificial intelligence (AI) technology to scour through large datasets, identifying medical decisions correlated with clinical success [10]. To date, applications of such systems have largely been relegated to somewhat simple tasks, such as detecting disease from pattern recognition in diagnostics (e.g., retinal scanning or examining radiographs for pathology) [18, 19].

Taiwan as a test case

One of the largest population databases in the world is Taiwan's National Health Insurance Research Database (NHIRD). Covering approximately 23 million individuals – nearly the entire population of Taiwan – the claims-based NHIRD contains comprehensive, longitudinal, real-world health data, integrable with other national datasets with additional information [12, 20]. This uniquely interoperable, accessible and centralized database could serve as an ideal starting point to develop real-world data-driven CDSSs, improving clinician decision-making and advancing precision medicine.

The NHIRD is rooted in Taiwan's single-payer healthcare system, the National Health Insurance (NHI), which was established in 1995 and now administered by the National Health Insurance Administration (NHIA). It covers virtually all of the Taiwanese population [21].

The NHI allows extensive patient choice and short waiting times, in part due to the absence of gatekeepers. Gatekeepers are less important when any clinician in any specialty can access patient records from anywhere in Taiwan as is the case with the NHI [13, 21, 22]. The system has multiple cost-containment strategies, including global budgeting, cost-sharing and the use of diagnosis-related groups [21]. As a result, the NHI operates efficiently and with relatively low administrative costs and rates of medical inflation; in 2018, healthcare expenditure equaled 6.6% of total gross domestic product, lower than the Organisation for Economic Co-operation and Development average of 8.8% [23, 24]. The NHI's affordability, accessibility, equity and lack of red tape likely explain the roughly 90% user satisfaction rate with the system in 2019 [22, 25].

The efficient centralization of insurance through the NHI has paved the way for a digitized healthcare system ("eHealth"). Digitization efforts included implementing an electronic,

automated insurance claims system for rapid review and processing [21, 26]. The foundational database, the NHIRD, was first released in 2002 to compile claims data from NHI beneficiaries for public research use, and has been continuously updated alongside new eHealth initiatives [11]. In 2004, the NHIA began issuing NHI integrated circuit (IC) cards to replace paper-based insurance cards – a pivotal centerpiece of eHealth. The IC card provides an additional level of security and a backup of recent patient health records [21, 22].

The development of EHRs in Taiwan also began in 2004, followed by a national EHR exchange system in 2008. The exchange was built to improve cross-institution continuity of care and now exists as a standard for data aggregation [26]. This was further enhanced in 2015 with the release of the NHI MediCloud – a cloud-based platform for retrieval of the patient's complete medical history, including high-resolution medical images [22].

Today, the NHIRD aggregates health data from MediCloud and ties them together with information on the patient's family relationships [11, 20]. Using unique identifiers representing anonymized individuals, the NHIRD is also linked to 70 other national databases, including birth and death registries, immunization records, cancer registries, reportable infectious diseases, contraception surveys, low-income household registries and family violence/sexual assault data [11, 20].

To access the NHIRD and all the data sets to which it has been linked, researchers must first access the Health and Welfare Data Center. Researchers must apply for on-site access to the data center to access the data, but once permissions are established, all the datasets can be accessed, linked and analyzed.

Test case: how the NHIRD was leveraged to combat COVID-19

Other countries that have made great progress at integrating health data systems include Estonia, Sweden, Denmark, Japan, the United Kingdom and South Korea [20]. The comprehensiveness, interoperability and size of these systems vary from nation-to-nation. Taiwan's NHIRD and related databases comprise one of the more comprehensive systems that have been made accessible for researchers, and its depth and flexibility can be seen in Taiwan's response to the COVID-19 pandemic.

In early 2020, the government linked its databases with immigration and customs information, alerting physicians about patients' travel history. The database also helped to identify suspected COVID-19 cases, focusing on individuals with respiratory symptoms who had tested negative for influenza. The real-time nature of the database allowed timely testing and quarantining [27]. Later, the government developed a name-based mask rationing system to avoid a mask shortage, using the NHI IC cards for identity verification to ensure equal distribution [28]. In 2021, the NHI's My Health Bank smartphone application was also updated to include COVID-19 test results and vaccination records [29].

Taiwan's pandemic response exemplifies part of what unified and integrated health databases can do, but it is just part of the bigger picture. In developing CDSSs powered by Taiwan's comprehensive, real-world data, the NHIRD is capable of revolutionizing healthcare and precision medicine in other nations.

CDSS development and its future potential

Many CDSSs exist in the USA and around the world serving medication alerts, contraindication warnings, assessment tools and more [7, 14]. However, more complex CDSSs, such as those aiming to offer diagnostic support through big data analysis, have seen slower rates of development and uptake owing largely to gaps in data availability and integration that hinder their performance [14]. However, the most revolutionary benefits of CDSSs – predictive diagnostics and precision medicine – are also the most data-dependent.

Powering AI-driven CDSSs with large, integrated datasets would massively expand their potential utility in guiding clinicians in day-to-day clinical practice and in advancing AI-based diagnostics and precision medicine. For example, one could imagine incorporating multiple data points from medical history, family history, diagnoses, treatments, laboratory results, demographics, social determinants of health in the community context and -omic data (such as the epigenome, proteome and metabolome). Outside of data from point-of-care encounters, wearable technology further allows continuous monitoring of data that can be harnessed as well [30]. Further integration with databases containing sociological data could also facilitate the inclusion of public health-based determinants in individual patient care.

Big data-driven CDSSs can then recognize patterns and match patient profiles as references for diagnosis and treatment, at a level exponentially better than humans. For example, one's medical history, laboratory tests and social context may be used to improve predictions of the likelihood of heart disease [31]. Likewise, the epigenome might greatly improve the sensitivity of existing tests, while the metabolome and prior laboratory tests may improve specificity, thereby reducing the false positive and false negative error rates of current approaches.

Such tools may eventually allow CDSSs and AI-based diagnostics and treatments to improve what is now a potentially dangerous trend of screening for diseases using laboratory tests or even wearable devices that have not been thoroughly vetted. This trend is complicated by both physician ignorance of the threat posed by false positive laboratory tests and industry's desire to market their products. One prime example of this is the single-lead Apple Watch, which was promoted by the American Heart Association, even as experts warn that even a full 12-lead ECG will produce more harm than benefit [32]. Taiwan's NHIRD provides one example of how this vision might be realized.

Big data experiments using human subjects

However, to realize its full potential, predictive analytics must be coupled with experimentation. Any user on Google or TikTok is all too familiar with the errors that AI entities make daily. As in associational research, such errors arise from spurious correlations, reverse causality or other data errors. When a human clinician is the diagnostician, the resulting scenario is analogous to self-driving cars – the machines can reduce errors, but it is easy for clinicians to fall asleep at the wheel and accept poor recommendations made by machines.

For these reasons, it is necessary to test patterns recognized by machines with experimental testing. This requires a great deal of caution and careful ethics oversight by knowledgeable experts. For example, a clinical diagnosis assigned a 98% chance of being correct using Bayesian statistics could be compared with the existing clinical practice guidelines provided that both alternatives are deemed to be reasonable in the eyes of expert clinicians and institutional review boards. Once this equivalence standard is met, it is possible to randomly assign thousands or even millions of people to one approach or the other, thereby proving that the algorithm is superior (or inferior) to the human-generated clinical practice guidelines. These "reality checks" need not be conducted on every diagnosis and treatment. As with self-driving cars, a human must be at the wheel. Nevertheless, the experiments provide valuable causal insights as to whether AI (non-knowledge-based) diagnostic and treatment algorithms are working.

Current applications of the NHIRD

The NHIRD and its interoperable databases currently serve important roles in academia and healthcare. Primarily, its provision of real-world data has been a powerful resource for

research across many topics including clinical medicine, pharmacology and healthcare utilization [20]. The databases have also supported initiatives such as MediCloud and pay-for-performance programs [13] to reduce health expenditures.

In recent years, the NHIA's increased focus on AI and big data aims to "use the NHIRD for 'intelligent' healthcare services, technology, and decision support services", [25] such as alert systems at point of care and enhancing information security. To streamline AI projects, the NHIA is also compiling databases of structured (e.g. claims) and unstructured data (e.g. written reports, imaging) to build a platform for AI development. One major project focuses on training AI with de-identified computed tomography and magnetic resonance imaging databases, to bolster predictive diagnostics, imaging interpretation and postoperative risk prediction. This draws on nearly 2.4 billion medical images compiled as of September 2020. Its initial efforts have been applied to COVID-19 diagnosis using chest x-rays as well [25, 33]. This represents an essential step in building AI-driven technology for health data interpretation, which could then be applied to CDSS development.

Academic research in Taiwan has contributed to CDSS development including AI-based CDSSs to screen for contraindications [34], predict medical conditions [35, 36], recommend medications [37] and support diagnosis [15].

Using deep learning-based systems and retrospective NHIRD data for a given condition, one team developed automated laboratory test recommendation systems drawing on EHR data. This approach could reduce both over- and underutilization of tests to improve patient care and prevent unnecessary costs [38, 39]. Similarly, private corporations have joined the effort. For instance, ASUS Intelligent Cloud Services (AICS) is collaborating with Changhua Christian Hospital to develop an AI-driven and precision medicine-focused CDSS by integrating clinical guidelines, laboratory test interpretation and patient-specific data [40].

Applying the NHIRD to integrative, big data-driven CDSS development

While the utility of systems like the NHIRD is enormous, the effort to build on such systems should be global if meaningful progress is to be made. One example is AI-based mammography, which was developed using the United Kingdom's integrated dataset alongside the dataset of a smaller hospital system in the USA and an international team of scientists [41].

A much broader and sophisticated set of algorithms could also be developed using an integrated national dataset first, and then tested and modified for local EHR systems; optimal clinical practice patterns should not vary much from one nation to the next.

Nevertheless, such a large-scale project should be conducted in collaboration with both the public and private sectors to support the diversity of real-world clinical scenarios. Therefore, significantly bolstering this momentum calls for a concerted, cross-institutional and cross-national push for CDSS development drawing from already mature databases such as the NHIRD.

Toward precision medicine

The final product – a big data-driven CDSS and AI-based precision medicine research platform based on 23 million cases over many years of follow-up – can be used to build predictive algorithms alongside the supervision of established guidelines. This allows for the capture of an individual patient's nuanced biological and social circumstances. Potential applications include:

(1) Pattern recognition to identify novel correlations previously not detectable in smaller databases;

(2) Analyzing patient profiles to develop predictive algorithms to generate novel medical hypotheses and test these hypotheses using automated randomized trials;

(3) Develop AI-based diagnostics using genomic, methylation, proteomic and metabolomic profiles;

(4) Improving AI-based diagnostics using data from wearable devices;

(5) Promoting high-value care, using past data to reduce both under- and overutilization of interventions, such as tests with low pre-test probabilities or low-yield treatment options, which helps optimize care and reduce costs;

(6) Identifying patients in need of preventive care through risk prediction based on health and sociological data [6, 14, 16].

Challenges

First, inherent in any data initiative is the challenge of preserving data privacy, especially health and personal data that are highly sensitive by nature. Unlike the Health Insurance Portability and Accountability Act in the USA, Taiwan instead maintains protections through a combination of laws that are applied to data more generally. This is a feature, rather than a flaw, for a system that mixes demographic, sociological and health data. One of the main safeguards in Taiwan is the Personal Data Protection Act, governing all forms of personal data. The Physicians Act and the Medical Care Act also prohibit the disclosure of health information without good cause [26, 42].

Within the NHIRD, unique and anonymized identifiers allow data linkage across 70 datasets without compromising confidentiality. Only credentialed individuals may access the databases [11]. Despite the lack of known breaches or leaks to date [20], any system is vulnerable. Although the Supreme Administrative Court upheld the use of the NHIRD for research, certain privacy issues, such as the ability to opt out of the NHIRD, remain under discussion [12]. Given the volume of work that must be done to build even basic machine learning tools, global opt-in or opt out rules need to be developed.

Privacy concerns must be addressed with transparency, a lack of mandates and a reliance on credentials that do not evoke the ire of the public, such as the use of biometric data [6, 9]. While there is a good deal of public awareness and sensitivity around facial recognition and fingerprint data, other forms of biometric data (e.g. gait) can go overlooked in privacy legislation.

On the back-end, the use of "-omic" data also presents unique challenges. These data can be used to identify individuals in anonymized datasets. As with facial recognition and fingerprints, legislation to protect genomic data can distract from equally dangerous proteomic or metabolomic data. Once an individual is identified using any biological patterns, it is possible to de-identify the remainder of that individual's data.

Nevertheless, linked databases such as NHIRD in Taiwan or the Government Cloud in Estonia are protected by the highest security standards, and breaches are therefore extremely rare. While a data breach in Taiwan or Estonia could potentially be catastrophic (impacting, for example, voting systems), this potential for a catastrophic breach also means that extreme measures are taken to protect these systems.

Second, while the national EHR exchange system and MediCloud have strengthened data sharing and interoperability, continued improvements will be necessary to permit usage of more health variables [26, 43]. For example, older unstructured data (such as written notes) contain a wealth of information but will inherently be challenging for AI to extract and interpret. Strategies like natural language processing will be necessary to make use of such data effectively. We propose that natural language processing and deeper learning can be

accomplished on MediCloud and that these algorithms can then be applied to localized data systems in other nations.

Third, inherent limitations of the NHIRD, such as a lack of indicators of disease severity and ongoing challenges working with multiple datasets, may create gaps in knowledge or imperfect predictions for both CDSSs and AI systems ("garbage in, garbage out") [11]. Unlike Estonia's Government Cloud, Taiwan does not make all commercial and government services available and is still in the process of building a complete eGovernance infrastructure. As a result, there is less pressure for continuous quality improvement and data integrity.

Fourth, the NHIRD has not been extensively validated. While the existing studies indicate relatively positive findings, a larger-scale effort with larger sample sizes to validate NHIRD diagnoses would be needed to improve confidence in the database [11, 12].

Fifth, health data through wearable technology obtained outside of healthcare settings could not only increase the representativeness of the NHIRD's claims-based data but also increase the false positive error rate, leading to rising costs and increasing medical errors if not implemented with caution [44]. Therefore, AI-based diagnostics must be carefully vetted by human experts who understand the nuances of primary, secondary and tertiary preventive screening tests.

AI can increase the positive predictive value of dangerous screening tests (e.g. screening for prostate cancer using prostate-specific antigen or atrial fibrillation using wearable devices), but the benefits of screening may still outweigh risks. The size, scope and regulatory powers of entities such as the US Preventive Services Task Force must be increased. The United Kingdom's National Institute for Clinical Effectiveness serves as a model for the potential regulatory powers of such entities but remains underfunded.

Finally, usage of existing CDSSs produces alert fatigue, the burden of system maintenance, reliance on computer literacy and disruptions to clinical workflow [8, 14, 45]. Potential solutions include better prioritization of alerts, systematic knowledge reviews, standardized data infrastructure to streamline updates, simplifications and automations in the user interface, CDSS training sessions and user and performance evaluations [8, 14, 45]. Special emphasis should be placed on streamlining systems and automatically integrating them with billing systems to reduce the clinician workload. Importantly, system testing must be rigorously performed to promptly address problems and enhance user confidence early on.

Conclusion

In Taiwan, a patient's NHI IC card is inserted into a card reader, retrieving patient data for the physician to review and update based on the patient's presentation. We propose a system that allows the CDSS to produce predictive algorithms that provide a pretest probability of disease. It would then recommend perhaps a single laboratory test to narrow down the differential diagnosis and encourage the physician to opt for tests with clinical utility while discouraging those unrelated to the underlying diagnosis. The CDSS then returns a list of diagnoses, each with a probability. Eventually, the system would potentially be able to recommend treatments that would be most appropriate for that individual. Ultimately, discretion is left up to the physician who works alongside the system. This is akin to "Level 4" self-driving cars, which keep a driver at the wheel but provide the driver with warnings when dangers are present.

In providing decision support based upon a patient's individual characteristics, we can begin realizing the full potential of precision medicine. With global initiatives to improve data integration and interoperability, Taiwan's existing and robust health data infrastructure offers an exciting first step to build and understanding precision algorithms. In harnessing Taiwan's readily available resources to develop an integrative, big data-driven CDSS, the possibilities of computerized medicine might finally be realized.

References

1. Wennberg JE. Dealing with medical practice variations: a proposal for action. Health Aff. 1984; 3: 6-32.
2. Makary MA, Daniel M. Medical error-the third leading cause of death in the US. BMJ. 2016; 353: 1-5.
3. Institute of Medicine. To err is human: building a safer health system. Washington, DC: National Academy Press; 2000.
4. Engl N, Med J, editors. Case records of the Massachusetts general hospital articles. [cited 2021 Sep 29]. Available from: https://www.nejm.org/medical-articles/case-records-of-the-massachusetts-general-hospital.
5. May TA, Clancy M, Critchfield J, *et al.* Reducing unnecessary inpatient laboratory testing in a teaching hospital. Am J Clin Pathol. 2006; 126: 200-6.
6. Han A, Isaacson A, Muennig P. The promise of big data for precision population health management in the US. Public Health. 2020; 185: 110-16.
7. Castaneda C, Nalley K, Mannion C, Bhattacharyya P, Blake P, Pecora A, Goy A, Suh KS. Clinical decision support systems for improving diagnostic accuracy and achieving precision medicine. J Clin Bioinforma. 2015; 5(4). doi: 10.1186/s13336-015-0019-3.
8. Belard A, Buchman T, Forsberg J, Potter BK, Dente CJ, Kirk A, Elster E. Precision diagnosis: a view of the clinical decision support systems (CDSS) landscape through the lens of critical care. J Clin Monit Comput. 2017; 31: 261-71.
9. El Naqa I, Kosorok MR, Jin J, Mierzwa M, Ten Haken RK. Prospects and challenges for clinical decision support in the era of big data. JCO Clin Cancer Inform. 2018; 2: 1-12.
10. Tenenbaum JD, Avillach P, Benham-Hutchins M, *et al* An informatics research agenda to support precision medicine: seven key areas. J Am Med Inform Assoc. 2016; 23: 791-5.
11. Hsieh CY, Su CC, Shao SC, Sung SF, Lin SJ, Yang YHK, Lai ECC. Taiwan's national health insurance research database: past and future. Clin Epidemiol. 2019; 11: 349-58.
12. Lin LY, Warren-Gash C, Smeeth L, Chen PC. Data resource profile: the national health insurance research database (NHIRD). Epidemiol Health. 2018; 40: e2018062.
13. Lee PC, Kao FY, Liang FW, Lee YC, Li ST, Lu TH. Existing data sources in clinical epidemiology: the Taiwan national health insurance laboratory databases. Clin Epidemiol. 2021; 13: 175-81.
14. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. Npj Digit Med. 2020; 3: 1-10.
15. Ho KF, Chou PH, Chao JCJ, Hsu CY, Chung MH. Design and evaluation of a knowledge-based clinical decision support system for the psychiatric nursing process. Comput Methods Programs Biomed. 2021; 207. doi: 10.1016/j.cmpb.2021.106128.
16. NEJM Catalyst. Healthcare big data and the promise of value-based care. NEJM Catal. 2018.
17. Tsai YW, Chuang YC, Huang WF, See LC, Yang CL, Chen PF. The effect of changing reimbursement policies on quality of in-patient care, from fee-for-service to prospective payment. Int J Qual Heal Care. 2005; 17: 421-6.
18. Cheng T, Zhan X. Pattern recognition for predictive, preventive, and personalized medicine in cancer. EPMA J. 2017; 8: 51-60.
19. Schmidt-Erfurth U, Sadeghipour A, Gerendas BS, Waldstein SM, Bogunović H. Artificial intelligence in retina. Prog Retin Eye Res. 2018; 67: 1-29.
20. Hsing AW, Ioannidis JPA. Nationwide population science: lessons from the Taiwan national health insurance research database. JAMA Intern Med. 2015; 175: 1527-9.
21. Cheng TM. Reflections on the 20th anniversary of Taiwan's single-payer national health insurance system. Health Aff. 2015; 34: 502-10.

22. National Health Insurance Administration. National health insurance annual report 2020-2021. Taipei: National Health Insurance Administration; 2021.
23. Ministry of Health and Welfare. National health expenditure 2019 abstract. Taipei: National Health Insurance Administration; 2021.
24. OECD. Health at a glance 2019. Paris: OECD Indicators; 2019.
25. Lee PC. Zouxiang yunduan bingyi shuangying: jianbao gaige riji [Toward the cloud, A win for patients and physicians: health reform diary. In: Chinese]. Taipei: National Health Insurance Administration; 2019.
26. Li YCJ, Yen JC, Chiu WT, Jian WS, Syed-Abdul S, Hsu MH. Building a national electronic medical record exchange system - experiences in Taiwan. Comput Methods Programs Biomed. 2015; 121: 14-20.
27. Wang CJ, Ng CY, Brook RH. Response to COVID-19 in Taiwan: big data analytics, new technology, and proactive testing. JAMA - J Am Med Assoc. 2020; 323: 1341-2.
28. Tai YL, Chi H, Chiu NC, Tseng CY, Huang YN, Lin CY. The effect of a name-based mask rationing plan in Taiwan on public anxiety regarding a mask shortage during the COVID-19 pandemic: observational study. JMIR Form Res. 2021; 5: 1-9.
29. Lee PC. New app streamlines testing and vaccination. Taipei: Taipei Times; 2021.
30. Dinh-Le C, Chuang R, Chokshi S, Mann D. Wearable health technology and electronic health record integration: scoping review and future directions. JMIR mHealth uHealth. 2019; 7(9). doi: 10.2196/12861.
31. Mohan S, Thirumalai C, Srivastava G. Effective heart disease prediction using hybrid machine learning techniques. IEEE Access. 2019; 7: 81542-54.
32. U.S. Preventive Services Task Force, Curry SJ, Krist AH, *et al.* Screening for cardiovascular disease risk with electrocardiography us preventive services task force recommendation statement. JAMA - J Am Med Assoc. 2018; 319: 2308-14.
33. Lee PC. Zouxiang yunduan bingyi shuangying: jianbao gaige riji 2.0 [Toward the cloud, A win for patients and physicians: health reform diary 2.0; in Chinese]. Taipei: National Health Insurance Administration; 2021.
34. Sun MC, Chan JA. A clinical decision support tool to screen health records for contraindications to stroke thrombolysis-a pilot study. BMC Med Inform Decis Mak. 2015; 15: 105.
35. Chen YF, Lin CS, Hong CF, Lee DJ, Sun C, Lin HH. Design of a clinical decision support system for predicting erectile dysfunction in men using NHIRD dataset. IEEE J Biomed Heal Inform. 2019; 23: 2127-37.
36. Hung CY, Lin CH, Lan TH, Peng GS, Lee CC. Development of an intelligent decision support system for ischemic stroke risk assessment in a population-based electronic health record database. PLoS One. 2019; 14: 1-16.
37. Chen RC, Jiang HQ, Huang CY, Bau CT. Clinical decision support system for diabetes based on ontology reasoning and TOPSIS analysis. J Healthc Eng. 2017; 2017. doi: 10.1155/2017/4307508.
38. Islam M, Yang HC, Poly TN, Jack Li YC. Development of an artificial intelligence–based automated recommendation system for clinical laboratory tests: retrospective analysis of the national health insurance database. JMIR Med Inform. 2020; 8: 1-10.
39. Islam MM, Poly TN, Yang HC, Li YC. Deep into laboratory: an artificial intelligence approach to recommend laboratory tests. Diagnostics. 2021; 11: 1-13.
40. Hou GZ. Huashuo zhihui yiliao zaixia yicheng, xishou zhangji jianzhi AI linchuang juece zhiyuan xitong [another win for AICS, Co-establishing AI CDSS with CCH; in Chinese]. Taipei: TechNews; 2020.
41. McKinney SM, Sieniek M, Godbole V, *et al.* International evaluation of an AI system for breast cancer screening. Nature. 2020; 577: 89-94.

42. Yang CM, Lin HC, Chang P, Jian WS. Taiwan's perspective on electronic medical records' security and privacy protection: lessons learned from HIPAA. Comput Methods Programs Biomed. 2006; 82: 277-82.

43. Wen HC, Chang WP, Hsu MH, Ho CH, Chu CM. An assessment of the interoperability of electronic health record exchanges among hospitals and clinics in taiwan. JMIR Med Inform. 2019; 7(1): e12630.

44. Wu YT, Lee HY. National Health Insurance database in Taiwan: a resource or obstacle for health research?. Eur J Intern Med. 2016; 31: e9-e10.

45. Melton BL. Systematic review of medical informatics–supported medication decision making. Biomed Inform Insights. 2017; 9: 1-7. doi: 10.1177/1178222617697975.

Corresponding author
Kevin Wang can be contacted at: kw2931@cumc.columbia.edu