Creating an open data city for healthcare: A critical review of data management strategy and development in China

Zigeng Fang[1], Qiuchen Lu[2], Long Chen[3]*, Jiayin Meng[4], Zhen Ye[5], Michael Pitt[6]

[1] PhD Candidate, Bartlett School of Sustainable Construction, University College London, London WC1E 6BT, UK. Email: zi.fang.15@ucl.ac.uk

[2] Lecturer, Bartlett School of Sustainable Construction, University College London, London WC1E 6BT, UK. Email: qiuchen.lu@ucl.ac.uk

[3] Lecturer, School of Architecture Building and Civil Engineering, Loughborough University, Loughborough LE11 3TU, UK. Email: l.chen3@lboro.ac.uk (Corresponding Author)

[4] PhD Candidate, Bartlett School of Sustainable Construction, University College London, London WC1E 6BT, UK. Email: jiayin.meng.20@ucl.ac.uk

[5] Associate Professor, Bartlett School of Sustainable Construction, University College London, UK, Email: p.ye@ucl.ac.uk

[6] Professor, Bartlett School of Sustainable Construction, University College London, UK, Email: michael.pitt@ucl.ac.uk

Abstract

The unexpected increased cases of COVID-19 pressured the healthcare system to be exposed to unprecedented challenges, where healthcare management is a complex but essential process to manage and coordinate all information and resources. Critical problems are emerging in this new routine and demanding a higher efficiency of medical data sharing. It is becoming emergent to improve the performances of city systems and create an open data city for reliable information and data sharing (e.g., medical and testing) for better public awareness and healthcare services. This research aims to critically review past efforts in open data city for healthcare from the perspective of data management strategy and development. China was selected as a representative due to its fast development in healthcare infrastructure and medical big data, where 79 out of 4,611 articles have been selected, reviewed, and analyzed. A mixed-method approach has been implemented to review and assess the existing efforts of open data city for healthcare through latitudinal and longitudinal analyses from five aspects based on socio-technical systems: technology, people, infrastructure, processes, and culture. The gaps, missions, and challenges of developing an open data city for healthcare have also been summarised, and an open data city healthcare reinforcement

framework has been proposed accordingly. This research contributes a new multi-dimensional way to rethink the development of open data city in healthcare and helps establish the state-of-the-art of open data city.

Keywords: Open data city, healthcare open data, socio-technical system, data strategies

1. Introduction

Compared with the relatively mild mortality rate of roughly 0.1%, COVID-19 is considerably deadlier with 15% of the infected patients having severe illness and 5% having a critical illness (Flavio Villanustre et al., 2021). For the vulnerable population, the fatality rates are even higher, as persons over 90 years old have a fatality rate of over 14% and patients with coexisting conditions with 10% and 7% fatality rate for the case of cardiovascular disease and diabetes. It is suggested that there is an urgent need for accessible epidemiologic data, as timely, accurate and accessible epidemiologic data is critical for informing public health response efforts (Berry et al., 2020). Accordingly, more and more publicly accessible open data (e.g., Ireland's open data portal (data.goc.ie) and Glasgow's open data website (data.glasgow.gov.uk)) have been made available for tackling the worldwide pandemic (Hunter et al., 2018). With these available datasets, more and more countries have realized the importance of healthcare open data for tackling the COVID-19 pandemic, especially due to the fact that the model built on the open data is reproducible and can be revalidated over its correctness. Although the outbreak of COVID-19 has enlightened the development of this research, the centre of this research is to identify, evaluate, and mitigate the management issues related to the healthcare open data. This is because the implementation of the COVID-19 information governance and strategy needs to rely on the management of modern healthcare open data (Alamo et al., 2020). Derived from the original definition of open data, "the data can be freely modified, used, and shared by anyone for any purpose" (Vetrò et al., 2016), this research defines a new concept, the "open data city" as the city with adequate infrastructure, culture, technology development that supports the accessibility, exploitability, useability, and shareability of the data for both the public and private purposes.

The research experience suggests that the health information technology (HIT)-enabled healthcare system should be treated as a complex adaptive social-technical system (Begun et al., 2003). It is even harder for a country like China to provide high-quality healthcare services to a large population covering rural and urban areas during the pandemic. The Chinese healthcare

information systems are mainly constructed within a hospital, namely the hospital information system (HIS), where Figure 1 shows the generalized Chinese hospital information flow for outpatient networks (Cao, 2006). This hospital-level information sharing process contains different information systems like clinical information system (CIS), laboratory information system (LIS), picture archiving communication system (PACS), radiology information system (RIS), Financial information system (FIS), electronic medical record system (EMRS) and pharmacy information system (PIS). With the popularity of big data technology in healthcare, the massive amount of data scattered throughout these mentioned hospital information sub-systems can provide a basis for the decision-making activities within the hospitals and simplify the process of patient medical treatment. However, the outbreak of the Covid-19 epidemic made people realize that data sharing within the hospital system is far from enough. To better realize the prevention and control of epidemics, information sharing in the context of healthcare should be established on a higher level, for example, at the city level, integrating information from hospitals, other external institutions, and even smart wearable devices. Although it has been mentioned previously that there are more and more available datasets available for the public (Berry et al., 2020; Hunter et al., 2018). Only a few studies (Chang et al., 2003; Yan et al., 2021) have attempted to evaluate the role of these HIS in the healthcare open data at the city level. Therefore, it is suggested that the clinical sub-systems (CIS, LIS, PACS, and RIS) and non-clinical sub-systems (FIS, EMRS, PIS and electronic prescription system) shown in Figure 1 should be re-evaluated against how they can be connected or utilize the healthcare open data in further studies.

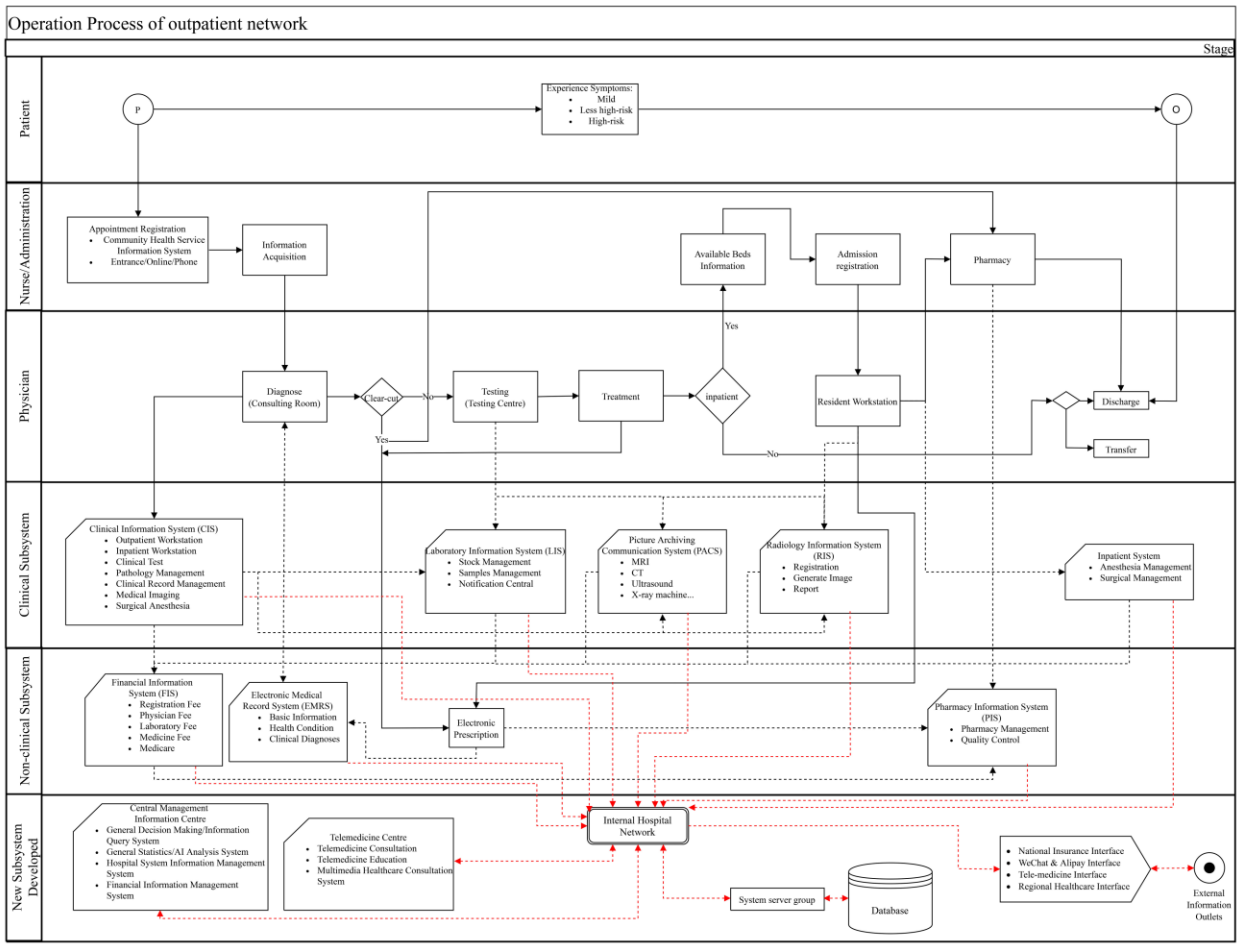


Figure 1 HIS related information flow for operation process of the outpatient network (adapted from Cao, 2006)

There are some conceptual models that were previously applied by HIT studies, including the user acceptance of the information technology model (Venkatesh et al., 2003), the diffusion of innovations theory model (Rogers, 2010), the distributed cognition model (Hutchins, 1995), the seven-step human-computer interaction model (Norman, 1988), and the swiss cheese model (Reason, 2000). Even though they all target one or more crucial facets when a new technology is introduced, the scopes of most of these models are limited in evaluating the adoption of technology itself and cannot address the complex relationships between the different stakeholders, information content, and hardware and software infrastructure. This derives the first research gap of this study: the lack of evaluation of the healthcare open data development from multi perspectives. This research gap exists because it is still unclear whether the data sets current held by the hospital, other government institutions, and technological infrastructure are complete or fit for their tasks. Debates are continuing these days for the accessibility and utilisability of the public to the

healthcare data with the proper consideration of personal information protection. Besides this, the focus of many existing related to healthcare big or open data studies have been focused on the development of the technologies and systems themselves, which have not properly covered the finding of identifying problems and providing solutions for arising issues (e.g., user data privacy protection (Zhou et al., 2022)) related to the introducing of these new healthcare systems and information technologies. This derives the second major research gap of this study: the lack of existing studies that systematically review the existing problems in open data healthcare development, especially targeting the healthcare information engineering and management and sharing from a city level. A systematic and critical literature review is urgently needed to review the current open data city healthcare development to find the potential solutions to the arising challenges from the transformation from big data to open data city healthcare.

This research intends to close this gap through a critical literature review of healthcare open data based on socio-technical system theory and using China as a representative due to its fast development in healthcare infrastructure and medical big data. Although the research context is focused on the healthcare systems in China, the findings from the literature review can be applied to other countries and regions with necessary amendments, where the gaps in current research can be modified accordingly, and the mission and challenges of open data city for healthcare can be shared and discussed. To close the above-identified research gap, three research questions were set up accordingly: (1) How is healthcare open data developed based on the perspectives of the socio-technical systems? (2) What are the gaps and challenges for healthcare open data development in different social and technical dimensions? (3) What are the future directions of healthcare open data development? Through answering these research questions, this research can provide insights into the understanding of cities' healthcare open data development and engineering related management challenges, applying recommendations for city-side healthcare open data development, and using an open data city healthcare reinforcement framework. Although the focus of this study's open data analysis is on the city level, it is important to consider open data practice for both hospital and national levels. This is because the policies and technologies used in these two levels will affect the healthcare open data at the city level. It is also crucial to distinguish the analysis of existing studies and projects into three different levels, as characteristics of stakeholders and data security and culture concerns are different across these levels. For instance, the main national-level target stakeholders are likely the policy maker or

government. However, patients and physicians are more favourited by the hospital level. While for data security, the main target for the hospital level is usually related to patients' information. It is probably replaced with HIS information and national databases for the city and national level. Thus, the levels of information sharing in this research have been separated into three different levels: (1) Hospital level (e.g., a rural clinic, a single hospital, a single medical institution, a single project initiative, etc.); City level (e.g., a group of citywide hospitals, a regional group of rural and city hospital networks, a group of hospital and medical institutions from a province, etc.); National level (e.g., a country-wide genome database network, a country-wide healthcare information sharing platform, etc.). However, this research will focus on the gaps, recommendations, and challenges analysis of complex city-level open data healthcare systems from perspectives of the social-technical system beyond the traditional focus of new technologies. There are three main reasons behind this. Firstly, choosing city level healthcare information sharing target can provide comprehensive coverage from the micro-level open data issue (e.g., a specific HIS system's data interoperability problem) to the macro-level open data issues (e.g., an incentive policy for data sharing). Secondly, it provides better coverage of different socio-technical system stakeholders and perspectives. Lastly, it allows the recommendations, gaps and challenges analysis to be targeted against a complex city system rather than a single technology, hospital, or government policy. The predictive capability of the social-technical system theory combined with the literature review will help to foresee the potential challenges and potentials of healthcare open data. This is because the social-technical system has the potential to improve the add-value of open data through enhanced awareness and evaluation (Davis et al., 2014).

2. Preliminary

2.1 Healthcare information technology and socio-technical systems

Healthcare systems as one kind of organization are usually complex and need to be researched in their entirety, where two sides should be considered conjointly, namely the social and technical dimensions (Sony and Naik, 2020). It is thus necessary to consider the servitization of healthcare information technologies as a socio-technical system to understand the healthcare systems from the combined perspectives of social and technical dimensions that consider behavioural aspects in the development and implementation of technologies (Münch et al., 2022).

The core ideology of the social-technical system theory reflected by the early work of (Clegg, 2000) is the concept that "design is systemic". It was then developed by the UK Tavistock Institute

for identifying the interrelated nature of technological and social facets of the introduction of the coal mining machinery technology (Trist & Bamforth, 1951). Within the socio-technical systems, the core is that it can be improved and can only work satisfactorily when the social and technical dimensions are brought together and treated as interdependent dimensions of a wider system (Chen et al., 2020).

There have been three popular socio-technical system models for classifying interacting and coordinated dimensions. Davis et al. (2014)'s hexagon-shaped social-technical system model considers multiple factors, in addition, to solely the technology itself, including people, infrastructure, technology, culture, process, and goal. Sittig and Singh (2015) further developed a new socio-technical model for adopting health information technology in complex adaptive healthcare systems, which is an eight-dimensional interconnected structure, including hardware and software, clinical content, human-computer interface, people, workflow, communication, and two additional dimensions, internal organizational features, and external rules and regulations. This model was later modified by Wesley et al. (2019) to a six-dimensional model but with two sides from both patients and healthcare providers, which was used for investigating the use of HIT for patient-reported outcomes.

Although originated from social science, the capacity of social-technical systems for optimizing organisations has enlarged the range outside the territory of social science. Many efforts have been made to investigate the servitization of information technologies in healthcare through the social-technical systems approach. For example, Lawler et al. (2011) have used the social-technical system approach to examine a wide range of cognitive ergonomics issues related to HIT (e.g., medication errors and electronic health records). Bonzo et al. (2016) used a large multidomain matrix representing the interactions across social-technical system domains for modelling processes and analysing communication in the hospital operating room. Fayoumi and Williams (2021) proposed an integrated conceptual model that incorporates the concept of a social-technical system to build an enterprise modelling framework that proves the power of providing a holistic IT modernisation for the healthcare system.

Built majorly based on Davis et al. (2014)'s model, this research will conduct a systematic literature review from the perspective of the social-technical system to understand the simultaneous effects of healthcare open data and related HIT across the five selected dimensions.

There were two reasons for adopting the model of Davis et al. (2014). Firstly, the works from Sittig and Singh (2015) and Wesley et al. (2019) are too focused on the technical dimension, but a more balanced model is needed for initiating this literature review. Secondly, other works are suitable to be used within a single hospital or specific HIT context but are too detailed and trivial when investigating the city-level healthcare system.

2.2 From 'big data' to open data city healthcare

Recent data-driven research over large size sharable datasets has provided many opportunities for city-level health care at both system and individual levels (Kostkova et al., 2016). In the light of 'opening' this 'big data', it is easier than ever to identify the diffusion pattern of disease and evaluate the effectiveness of the government policy. With a significant amount of data generated each year, the healthcare data available in 2020 is projected to be 50 times more than in 2011 (Austin & Kusumoto, 2016). Out of all data generated, data from private repositories and commercial organisations is not shareable. The focus, then, should be put on the openable data. Understanding the relationship between open data and 'big data' is crucial for researchers and industry practitioners. Similarly, it is also important to identify how the existing techniques, platforms, and tools' capability of tackling big data's high volume, high dynamism, and heterogeneous complexity can be utilised to benefit the open data practice (Palanisamy & Thirunavukarasu, 2019).

Because open data is required to be freely used, modified, and shared by anyone for any purpose (Vetrò et al., 2016). These requirements of open data have ruled out a significant amount of healthcare data gathered by commercial organisations and private institutions, especially due to the sensitivity of health information. It is suggested that open data is inseparable from the 'big data' because for the healthcare domain, the target data to be opened is normally the 'big data' collected by the regional or central public health department and regulators. On the other hand, there has been an enthusiasm for using 'large' data resources and new information technology (IT) to tackle the existing medical obstacles (Obermeyer & Emanuel, 2016) and help the better decision-making for countering the challenges brought by COVID-19. Therefore, it is important to evaluate the healthcare open data from multiple perspectives and review the arising challenges to find solutions for current obstacles.

3. Research Methodology

This study adopted the mixed systematic literature review (SLR) approach to reviewing healthcare open data city-related studies recently conducted in China in a mix of HIT and architecture, engineering, construction, and facility management (AEC/FM) domains, with the following of the latest 2020 PRISMA systematic literature review standard (Page et al., 2021). Scopus, PubMed, Web of Science, and China National Knowledge Infrastructure (CNKI) databases were selected as they provide good coverage of both the medical and AECFM literature in English and Chinese. Keywords in "open data" city, healthcare, and socio-technical system's sub-groups using both English and Chinese between 2000 and 2021 (Table 1). In general, the selected keywords were classified into three major categories: open data, healthcare, and socio-technical systems. The socio-technical system is further separated into five dimensions according to the model raised by Davis et al. (2014). Inside each category or between sub-categories, keywords were searched using the 'OR' operator. While, between categories, the search query is connected using the 'AND' operator.

Table 1. Keywords searched for systematic literature review

Category	Sub-category	Keywords
Open data city-related		'Open knowledge' OR 'open source' OR 'open research' OR 'open content' OR 'open access' OR 'data sharing' OR 'information sharing' OR 'crowdsourcing' OR 'crowdsensing' OR 'big data' OR 'citizen data' OR 'smart city' OR 'urban accessibility'
AND 'Healthcare-related'		'Universal health care' OR 'primary care' OR 'secondary care' OR 'hospital' OR 'community hospital' OR 'general practice' OR 'private practice' OR 'clinical' OR 'wellness' OR 'medical' OR 'medicine' OR 'medicare'
AND 'social-technical system-related'	Culture	'Culture' OR 'style' OR 'project' OR 'policies' OR 'governance' OR 'agency' OR 'event' OR 'organization' OR 'government'
	People	'People' OR 'group' OR 'record' OR 'portal' OR 'crowd' OR 'end-user' OR 'user involvement' OR 'leadership' OR 'coordination' OR 'role' OR 'responsibility'
	City & Building & Infrastructure	'City' OR 'city lab' OR 'smart city' OR 'urban' OR 'municipal' OR 'metropolitan' OR 'architecture' OR 'building' OR 'transportation' OR 'infrastructure' OR 'foundation'
	Processes & Procedures	'Analytics' OR 'process' OR 'procedure' OR 'application' OR 'uses' OR 'agenda' OR 'plan' OR 'preparation' OR 'step'
	Technology	'Information system' OR 'IoT system' OR 'system' OR 'technology' OR 'data mining' OR 'data harvesting' OR 'dashboard' OR 'information technology' OR 'information display' OR 'clinical information system' OR 'electronic health record' OR 'electronic medical record' OR 'health information system' OR 'hospital information system' OR 'drug prescription and distribution system' OR 'mobile nursing technology' OR 'clinical decision support system' OR 'picture archiving and communication system' OR 'radio frequency identification' OR 'electronic patient record' OR 'computerized physician order entry' OR 'electronic signature' OR 'mobile nursing information system' OR 'financial

	information system' OR 'radiology information system' OR 'nursing information system' OR 'laboratory information system' OR 'pharmacy information system'

The systematic literature review of this article is constructed in two phases. The first phase is the article selection; the second phase is the content analysis. In the article selection phase's article identification stage, firstly, the articles from the keyword search and the primary manual-based reference review will be concatenated with duplicates removed (Figure 2). Secondly, in the screening stage, articles will be screened based on their titles and applying web-based systematic literature review software – Rayyan (Ouzzani et al., 2016). The web-based semi-automated literature review app allows the expedition of the initial screening of titles using algorithms like the support vector machine (SVM) classifier (Cortes & Vapnik, 1995). After the title screening step, articles that lack abstract or are irrelevant to the main topics of open data, healthcare, and socio-technical systems will be further excluded. Then, the full-text review is conducted over the rest of the articles to further screen out the article with missing full text, loosely connected to the selected main topics, with a limited or narrow focus, or with no supportive information for open data city. In total, 25 studies mentioned in this research were written in Chinese and 86 studies were written in English.

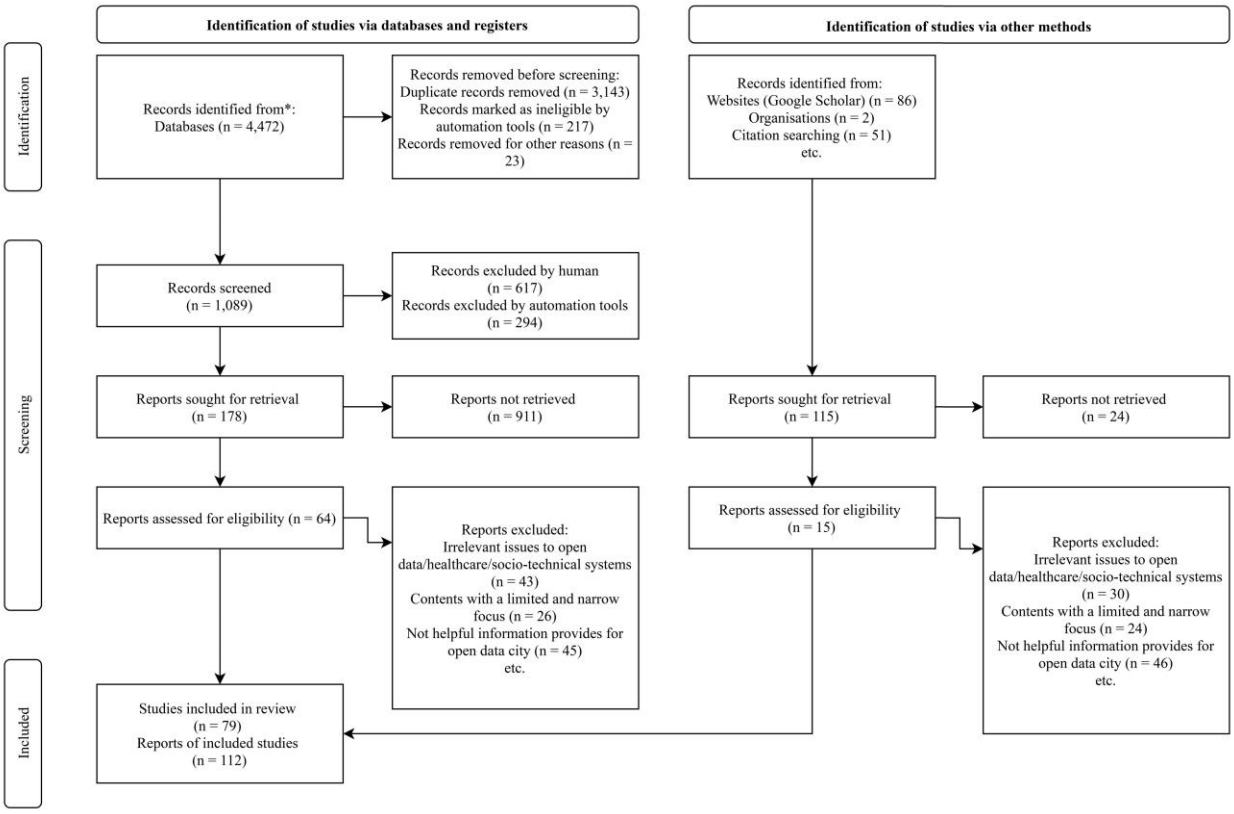


Figure 2 The working flow of the literature screening process

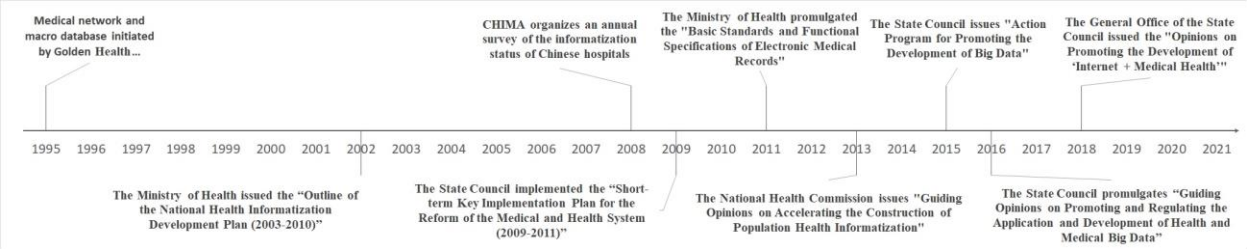
In the article content analysis phase, after identifying and selecting the relevant articles, a mixed analysis method was conducted: (1) the latitudinal analysis is conducted to identify the milestones that marked by the major open data healthcare regulations and guidance published by the Chinese government, and summarise the recent trend of the open data healthcare-related studies; (2) the longitudinal analysis is conducted to view the existing studies of healthcare open data city in the lens of social-technical system's five different dimensions. Accordingly, these studies were discussed over its benefits and limitations against the open data city for healthcare; (3) the gaps, missions, recommendations, and challenges of the open data city for healthcare are summarised from the previous longitudinal analysis, which also inspires the development of a new open data city healthcare reinforcement framework for guiding the research and development of a better integrated and connected healthcare open data powered city.

4. Results and analysis

4.1 Latitudinal analysis

Throughout the analysis of the selected literature, it is shown that the development of China's open data healthcare is mainly driven by the issue of government guidance and regulations, most of

265 which have been published in 2008-2018 (Figure 3). It indicated that the development of healthcare
266 open data had attracted attention in recent years.



267

268 **Figure 3** The milestones of open data healthcare policies in China

269 The result of latitudinal analysis (Figure 4) suggests in 2000-2010, the number of open data-related
270 studies was very limited, and the research topics were mostly constrained to the culture (e.g.,
271 organizational culture and government policy), healthcare information technology infrastructure
272 (e.g., biobanks), and healthcare information technology (e.g., HIS). It was only until the recent
273 decay studies gradually looked into the healthcare open data has grown sufficiently and covered
274 more topics, including the healthcare-related stakeholders (e.g., patients, physicians, etc.) and the
275 medical information management and sharing process (e.g., pharmacy information management
276 process). Overall, out of five social-technical system model dimensions, the 'Culture' and 'City &
277 Building & Infrastructure' dimensions are the most studied categories (Figure 5). While the
278 number of existing studies for 'People' dimensions is limited. Potential explanations for this are
279 the related research materials (e.g., biobank, medical "big data", and government policies) are rich
280 in amount and regarded as the main stream of healthcare open data.

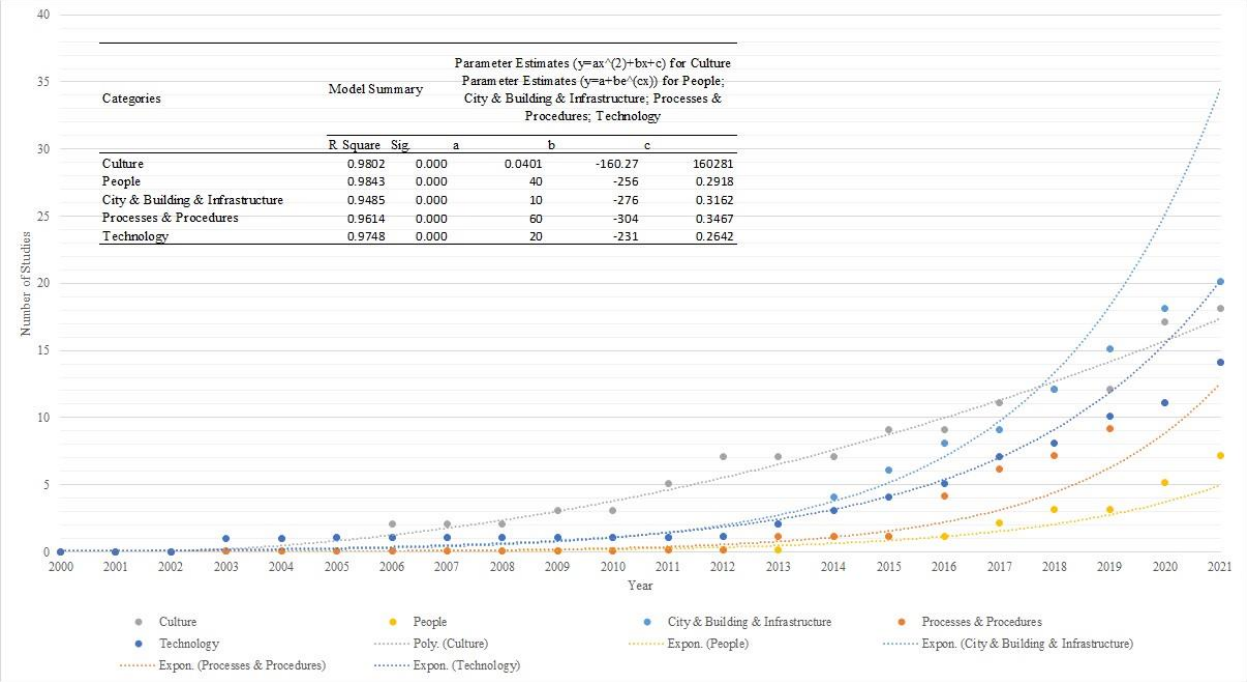


Figure 4 The tendency of China's Healthcare Open data studies in time series

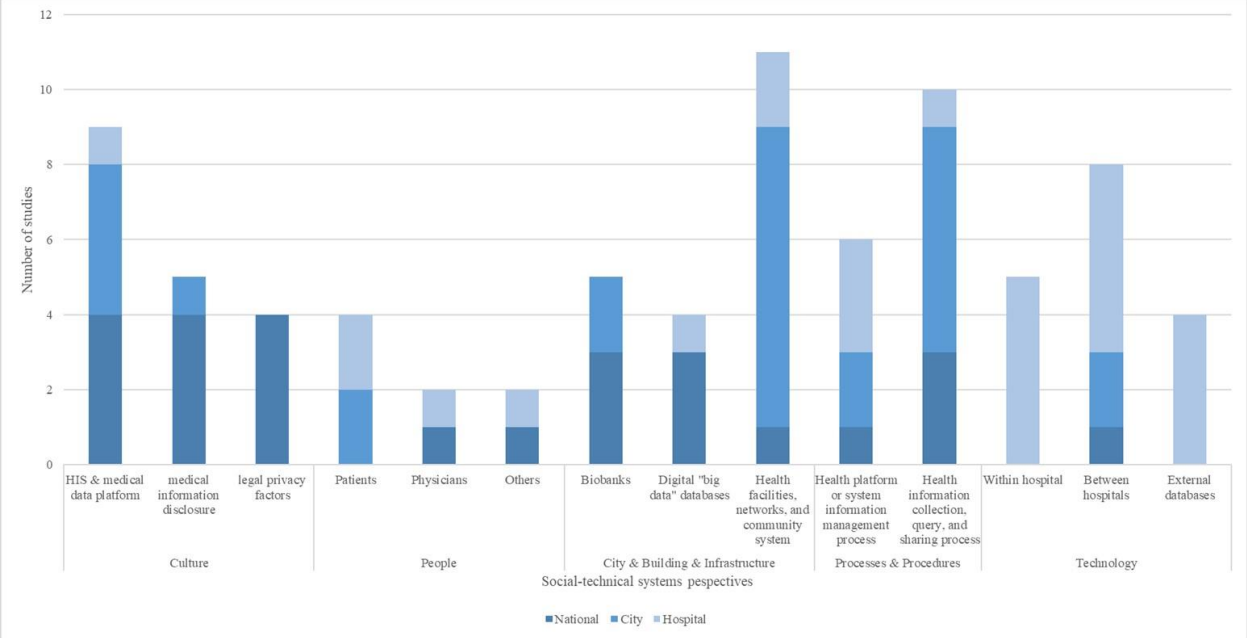


Figure 5 The distribution of the selected studies in different dimensions

4.2 Longitudinal analysis

4.2.1 Culture related to healthcare in China

In the field of culture, as to say the policies and research related to open data and data sharing in the context of healthcare in China, the main analysis can be divided into three parts: construction of hospital information system (HIS) and healthcare big data platform; medical information

disclosure and communication by the government; and legal and regulatory privacy factors that applicable to medical information sharing (Table 2)

The need for an effective hospital information system was well recognised as the critical element for the feasibility of healthcare open data at a city level (Li et al., 2020). A series of rules and regulations have been implemented at the national level to promote the establishment and improvement of HIS in China. As early as 1995, the Golden Health Project clearly defined the goal of establishing the internal medical network and the domestic health macro database (liu & Huang, 2009). In the year 2002, the Ministry of Health issued the "Outline of the National Health Informatization Development Plan (2003-2010)", emphasizing the importance of an integrated regional health information system (Luo & Ping, 2011; Zhu, 2011). In 2009, the State Council implemented the "Short-term Key Implementation Plan for the Reform of the Medical and Health System (2009-2011)", requiring the gradual establishment of nationwide residents' health files and a shared medical and health information system (Zhao, 2012). The "Guiding Opinions on Promoting and Regulating the Application and Development of Health and Medical Big Data", promulgated by the State Council in 2016, emphasized the promotion of the sharing and openness of healthcare big data (Wu et al., 2020). In terms of research, many scholars indicated that HIS and big data improve the efficiency of the quality of medical service, along with the reduction of medical expenses, help to optimize the allocation of medical resources, and promote fairness and transparency of medical services (Chen et al., 2018; Li et al., 2019; Luo & Ping, 2011; Wan et al., 2012; Yin et al., 2020; Zhu, 2011). However, due to the insufficient construction and update of relevant data sets, the sharing and opening of high-value clinical data were still lagging (J. Zhang et al., 2020).

With the development of hospital informatization, especially after the spread of the SARS epidemic in 2003, medical information collection and disclosure were regarded as an important element in crisis management and medical service (Lv, 2018; Ma, 2015; Shangguan et al., 2020; Zhang et al., 2006). (Yu, 2006) pointed out that medical information asymmetry is the core cause of conflicts between doctors and patients, a serious social problem that urgently needs to be solved. This was further supported by (Ma, 2015), who argued that medical information disclosure holds a negative relationship with the hospital mortality rate, and better medical information disclosure can effectively eliminate the negative mortality rate by the health institutions. However, it was

worth noting that the foundation of medical and health information disclosure is government information disclosure (Lv, 2018). Unfortunately, China's current medical information opening was still rough and immature (Shangguan et al., 2020). In terms of healthcare research data sharing, platforms such as National Population Health Data Center (NPHDC) and CNGBdb-EBB (E-BioBank) promoted the scientific research and academic development of health and medical care in China. Nonetheless, most of the resources and databases were still low quality and hard to be successfully accessed (Liu & Dong, 2012).

Research on privacy frameworks in China has become popular these days under the circumstance of calling for healthcare data sharing and openness (Chen et al., 2015). Scholars have researched the patients' willingness to share healthcare data, considering privacy protection. Ma et al. (2019) used a sampling method to investigate whether Chinese patients would share their health and healthcare data anonymously. They found that patients are more willing to share data within the hospital and government departments than with other hospitals. Based on evolutionary game theory, Han et al. (2021) analysed the strategies of patients and hospitals in sharing healthcare data in the context of privacy protection. Some scholars in China tried using emerging technologies to build data sharing models to ensure the information security of healthcare data (Xue et al., 2017; Yu et al., 2019).

Overall, there were three major aspects of open data healthcare in the context of culture. These culture-related studies on medical and health data sharing and disclosure came from different aspects, from hospitals' internal sharing (Chen et al., 2018; Li et al., 2019; Luo & Ping, 2011; Wan et al., 2012; Yin et al., 2020; Zhu, 2011) to government medical information disclosure (Lv, 2018; Ma et al., 2015; Shangguan et al., 2020; Zhang et al., 2006) and discussion on patients' privacy (Chen et al., 2015; Han et al., 2021; S. Ma et al., 2019; Xue et al., 2017; Yu et al., 2019). In practice, despite the remarkable achievement of China's hospital informatization in the past 20 years, the quality of data openness and sharing still need to be improved in the areas like city-wide consideration of healthcare information, the medical and non-medical institution information communication, the realization of practice cases, and the consideration of the open healthcare data ownership.

348 **Table 2.** Development of culture for open data healthcare

Category	Level	Objectives	Pros	Cons	References
HIS & healthcare data platform	National	This paper analyses the typical principles of medical information sharing	This study reveals the important components in health and healthcare data sharing to guide Chinese application	This study lacks a detailed pattern for medical sharing	(Zhang et al., 2020)
		This paper summarises the research progress and highlights Chinese medical data sharing	This study analyses the research themes and application practice of healthcare data sharing in China and points out the lack of regional data sharing	This study does not consider the sharing of medical information in the city as a whole	(Yin et al., 2020)
		This paper analyses the content of hospital informatisation in China and discusses its development trend	This study comprehensively introduces the development process of Chinese hospital informatisation	This study lacks the elaboration on medical information sharing	(Liu & Huang, 2009)
		This paper introduces the sharing of medical and health information and relative technologies	This paper shows the development of electronic health records (EHR) and electronic health management (EHM) in China	This study focuses on different components of HIS and does not involve medical information sharing	(Yao, 2006)
	City	This paper concludes four modes for Chinese regional medical information development and analyses the influencing factors of regional medical information construction mode	This study uses quantitative methods and finds that the construction of regional shared medical information platforms is in line with China's development status	This paper is constrained to healthcare information sharing between hospitals, ignoring the data sharing between hospitals and non-medical institutions	(Wan et al., 2012)
		This paper introduces the development status of regional health information	This study illustrates the concept and practice of health informatisation at the city level	This study ignores the data sharing between hospitals and non-medical institutions	(Luo & Ping, 2011)
		This paper introduces the technologies for regional health care big data platform	This study explores the establishment of regional medical big data platforms and data sharing from the perspective of computer and information technology	This study contains no in-depth analysis of the medical information sharing mode in China	(Wu et al., 2020)
		This paper takes Shanghai's regional medical informatisation construction as an example, analysing the related achievements and problems	This study introduces the establishment of the regional healthcare data platform in detail	This study's discussion on health information sharing is not enough; ignoring the data sharing between hospitals and non-medical institutions	(Zhu, 2011)
	Hospital	This paper analyses the current issues in the establishment and management of residents' health records in China	This study takes a Chinese hospital as an example and summarises the achievements of regional medical informatisation from the perspective of health record management	This study selects only one case, which is not enough to summarise the current situation of China's regional medical development	(Zhao, 2012)

Medical information disclosure	National	This paper discussed the connotation of medical and health information disclosure policy based on the analysis of the research background of it	This study defines the study framework of medical and health information disclosure policy	This study only considers the medical information disclosure from the perspective of policy, lack of practical cases	(Lv, 2018)
		This paper analyses the reason for the outbreak of COVID-19 in China from the perspective of crisis management	This study considers the relationship between medical information disclosure and crisis management	The study only focuses on the government control over information, ignoring other forms of medical information sharing	(Shanggu an et al., 2020)
		This paper makes some countermeasures and suggestions to the medical information opening	This study points out that the asymmetry and non-disclosure of medical information are the core of the contradiction between doctors and patients.	This study has insufficient research on medical information disclosure and sharing channels	(Yu, 2006)
		This paper analyses the research theme and hot trends in medical and health information disclosure	This study conducts a systematic literature review on Chinese research focused on medical information disclosure	This study lacks content on current practice in medical and health information disclosure	(Lv, 2017)
	City	This paper analyses the relationship between the information disclosure and the service quality of health institutions	This study uses the method of quantitative analysis and takes Beijing as an example to study the effect of hospital information disclosure in China	This study only considers the information disclosure on the website of medical institutions, ignoring other forms of information sharing	(Ma, 2015)
Legal privacy factors	National	This paper studies the willingness of patients and medical service providers to participate in medical data sharing from the perspective of privacy protection	This study builds a dynamic equation based on evolutionary games and solves stable evolutionary strategies under various conditions	This study fails to consider the influence of government management departments, platform service providers and other stakeholders on the willingness to participate in the process of healthcare data sharing	(Han et al., 2021)
		This paper studies the protection and disclosure of patients' personal medical information from the perspective of law study	This study discusses in detail the factors involved in personal privacy in the disclosure of medical information	This paper only analyses the legal issues and does not establish a complete medical information disclosure system	(Yang, 2017)
		This paper analyses the privacy and security risk factors in medical data under the background of big data	This study considers healthcare data sharing in the context of big data and cloud services	This study only considers the privacy issues of medical big data from the questionnaire survey data and does not provide in-depth solutions.	(Lv & Qiao, 2020)
		This paper studies the privacy and biobanking in China in the context of policy in transition	This study investigates and integrates legal issues related to the sharing of medical big data, especially on biobanks	This study lacks connection with real practice cases	(Chen et al., 2015)

349

4.2.2 People perspective towards healthcare open data

Along with healthcare personnel and patients, who are the main factors within the healthcare open data framework, researchers and policymakers are also important to benefit from medical and healthcare data sharing and openness. Hence, in the context of people related to healthcare open data, three categories are classified for further research: patients, physicians, and others (Table 3).

Patients are the people who are most closely associated with the openness of healthcare data. On the one hand, accurate, complete, comparable and timely data were crucial for disease treatment and surveillance of patients (Tian et al., 2020); on the other hand, the unsound legal system of security privacy protection in China laid a huge risk of healthcare data abuse, which concerns the patients and their relatives. Furthermore, patients can benefit from improved hospital logistics with the integration of medical and non-medical information in the hospital (Xie et al., 2022).

Another important component of people related to healthcare open data is physicians in hospitals and medical institutions. Through the aggregation and real-time analysis of healthcare data from Hospital Information System (HIS), such as Picture Archiving and Communications System (PACS) and Laboratory Information System (LIS), physicians and doctors can adjust their treatment plans in time to make evidence-based and personalized treating decisions (Y. Liu et al., 2018; Yao et al., 2014). What is more, healthcare data sharing from different institutions promoted interoperability in diagnosing, enabling doctors to identify patients with similar diagnose and symptoms and to determine treatment plans in cooperation (Hasan et al., 2021). By recognizing patient population forecasts, doctor scheduling, and hospital structure from daily healthcare-related data, hospital personnel scheduling could be settled to meet the patients' needs (Feng et al., 2021).

Besides, healthcare open data benefits researchers and policymakers in an unprecedented way. Healthcare open data like hospital information disclosure facilitated transparent government transformation and improved government credibility (Lee, 2020). In terms of reasonable policy, available data for scientific research plays an essential role. Healthcare open data creates the opportunity for the researcher to have data-driven insights and data-based analysis to penetrate the policy process more deeply than ever before.

To sum up, in terms of people related to medical and healthcare open data, patients, medical personnel, researchers, and policymakers were the main beneficiaries. There were various benefits

for different groups of these open data "stakeholders". For instance, patients benefitted from better disease treatment and surveillance (Tian et al., 2020) and better integration of medical and non-medical information (Xie et al., 2022); medical practitioners were leveraged from more evidence-based and flexible treatment plans and decision-making (Liu et al., 2018; Yao et al., 2014), as well as, interoperability and cooperation in disease diagnosis (Hasan et al., 2021) and shift scheduling (Feng et al., 2021); while, government benefitted from improved creditability and transparency (Lee, 2020). However, it is also worth noting that there were still many areas of improvement from the perspective of people. They include patient-physician communication's efficiency and legal protection, the channels of medical information sharing, and the lack of healthcare policy supporting information.

390 **Table 3.** Development of people perspective for open data healthcare

Category	Level	Objectives	Pros	Cons	References
Patients	City	This paper connects insurance-system-based claims data in two regions with cancer incidence rates	This study establishes a medical-insurance-system-based cancer surveillance system in China and evaluates its completeness and timeliness	This study fails to discuss the perspective of medical information sharing	(Tian et al., 2020)
		This paper discusses the situations of sharing and application of medical big data	This study mentions and discusses the online health communication for patients and physicians	This study lacks a detailed classification of healthcare data sharing discussion on the benefits to patients	(Liu et al., 2018)
	Hospital	This paper investigates the methods for improving the credibility and interestingness of social media healthcare information	This study takes WeChat public account as a medical information sharing platform and explores the effect on patients	This study is limited to one specific social media, WeChat and no detailed discussion on further medical information sharing	(Fan et al., 2017)
		This paper facilitates the integration of patient-related information with the conventional FM processes	This study demonstrates the potential improvement in hospital logistics and pharmaceutical services due to the integration of medical and non-medical information in the hospital	This study's information integration and sharing between patient-related information and conventional FM processes only stay at the pilot stage	(Xie et al., 2022)
Physicians	National	This paper takes the cloud-based hospital information system as a service for grassroots healthcare institutions	This study analyses the utilisation of different parts of HIS for physicians in the treatment	This study lacks a more detailed discussion of different clients' needs of the HIS	(Yao et al., 2014)
	Hospital	This paper optimises the hospital personnel scheduling through patient prediction based on big data	This study investigates the model for optimising the allocation of various hospital resources using big data, especially on physicians' personnel scheduling	This study is based only on a project level, which lacks a discussion on the city-level	(Feng et al., 2021)
Others	National	This paper investigates the big data strategies for government, society and policy-making	This study emphasises the importance of data for policymakers and governments	This study does not focus on the field of medical and health data sharing	(Lee, 2020)
	Hospital	This paper explores the application of blockchain in medical data sharing	This paper mentions the healthcare data sharing for policymakers and physicians	This study mainly focuses on technology, with only a brief discussion of healthcare stakeholders in the context of healthcare data sharing	(Hasan et al., 2021)

391

4.2.3 infrastructure supported open data healthcare

The "open data" healthcare's city, buildings, and infrastructure dimension in the context of the social-technical system is about examining the physical and digital format of city, building, or even national-wide system infrastructure that supports the "open data" related healthcare projects. The main analysis targets for infrastructure supports can be broadly classified into the following categories given the found literature: biobanks; digital "big data" databases; health facilities, networks, and community systems (Table 4).

Unlike other industries, one of the major sources of "big data" that can be opened - the biobank - is established and operated by hospitals and research institutes. As early as 1992, the Chinese Marrow Donor Program (CMDP) set up its first population-based blood biobank that contains about 1,650,000 specimens (Cheng et al., 2013). Then, in the early 21st century, many different hospitals set up their disease-oriented biobanks to collect different specimen types, including tissue, RNA, DNA, cell, blood, and plasma (Li et al., 2004; Yang et al., 2012). Unfortunately, no unified network infrastructure was applied across the country, and the specimen information could not be communicated and shared among different biobanks efficiently (Cheng et al., 2013). It was also recognized that the lack of an official unified-standard implementation method and procedure was the major obstacle to specimen information sharing, including clinical information, specimen information, and other derived data (Y. Zhang et al., 2015). In order to integrate the existing separated biobanks to establish a virtual biobank, an intra-hospital virtual cancer bank network has been proposed to share the information of bio-samples stored in individual banks using a single hospital pilot study (L. Zhang et al., 2015). More recently, some regional cooperated biobank like the China Health Big Data was built to utilize and store the Low-Dose Chest Computed Tomography (LDCT) from lung cancer screening to improve the Quantitative Computed Tomography (QCT) (Wu et al., 2019). This project gathered 16 clinical centres to use the HIS for data export and sharing.

Besides biobanks that store biological samples and data, digital databases service is another major data source for "open data". Recently, an increasing amount of genome and "big data" databases were established across the entire country. For example, some researchers used the genome sequence data that the 1000 Genomes Project Consortium published to build the population-based virtual Chinese Genome Database (VCCGDB) (Ling et al., 2014). In the year 2018, to meet the demand of development of biomedical and improve the big data management and sharing, the

Beijing Institute of Genomics (BIG) of the Chinese Academy of Sciences founded the BIG Data Center (BIGD) in the year 2016 to provide genetic information sharing database that contains a variety of different resources including the GWH (Genome Warehouse), GSA (Genome Sequence Archive), GEN (Gene Expression Nebulas), GVM (Genome Variation Map), MethBank (Methylation Bank), Science Wikis, and BioCode (Zhang et al., 2018). Other researchers used the more disease-specific data sharing database platform to control, monitor and analyze mumps (Jiang et al., 2019). In the path of "open data", the research team of this project went even one step further to publish the construction code of PIC on GitHub for public utilization and visualization. There were also tidied up and publicly available data like China Physiological Signal Challenge (CPSC) 2018 and 2019 database that offered the ECG data collected from different hospitals and devices (e.g. wearable ECGs), which can be further used for ML-based feature engineering and prediction (Gao et al., 2020). Furthermore, 2019's COVID pandemic also raised the need for an openly accessible coronavirus database for research.

Other than the biobanks and databases, other important infrastructure facilities and networks play an important role in promoting healthcare "Open Data". The 'scattered' patient clinical information and the "data islands" problem were some of the major challenges medical institutions faced. Researchers proposed an encrypted XML based cloud platform to facilitate the information exchange through the web service with the protection of data privacy (Wu & Wang, 2014). To further boost the effectiveness of some existing IT-based community care systems for countering diabetes, some researchers further embedded cloud technology and integrated multi-source data to enable the timely monitoring of patients' health can exchanging the patients' information (Yang & Cao, 2017). Some researchers looked into the data sharing and connection barriers between EMR and the web-based Public Health Information Systems (PHIS) for some regional health information platforms in China to guide public health decision-making and support public health surveillance (Su & Zhao, 2016). One of the major technical challenges that stopped the data sharing between EMR and the PHIS is that there was no standardized data translation and exchange messaging format like Health Level 7 (HL7) existed in China. Also relying on the cloud-based computing and service collaboration model and big data technology, the Big Data Medical Service Platform (BDMSP) was proposed to provide the technical solutions and knowledge sharing capability for the regional medical informatization in China (G. Song et al., 2019). Besides the cloud-based and blockchain protected platforms (Wang et al., 2019b; Al Omar et al., 2019), some

regional healthcare infrastructures provide health system, population health, and patient medical information integration for the rural area (Wang et al., 2016), the Guangdong-Hong Kong-Macao Greater Bay Area (Q. Zhang et al., 2020), the smart city (B. Xu et al., 2018), and multicentre projects (Li et al., 2021a).

To summarize, there were three main exploration areas of the healthcare open data for the infrastructure context. The first group of research investigated the current development progress and characteristics of China's biobank over healthcare information management and communication (Cheng et al., 2013; Li et al., 2004; Wu et al., 2019; Yang et al., 2012; Zhang et al., 2015). The second group of articles looked at the digital "big data" databases, where the spectrum of the articles covers the Genome databases (Ling et al., 2014; Xu et al., 2018b; Zhang et al., 2018), disease-specific database (Jiang et al., 2019), and technology-specific database (Gao et al., 2020). The last research group comes to the health systems, facilities, networks and communities. Articles covered cloud-based community care systems and platforms (Wu & Wang, 2014; Song et al., 2019a; Su & Zhao, 2016; Yang & Cao, 2017), multi-department, regional or citywide health infrastructure (Cai et al., 2020; Li et al., 2021; Wang et al., 2016; Xu et al., 2018a; Zhang et al., 2020b; Wang et al., 2019b; Al Omar et al., 2019). Unfortunately, some limitations restricted a wider sharing of medical information from the healthcare infrastructure's perspective: the scattered storage of genome and patients' clinical information, the lack of unified network infrastructure applied across the country, and the lack of China's standardised data translation and exchange messaging format.

474 **Table 4.** Development of infrastructure for open data healthcare

Category	Level	Objectives	Pros	Cons	References
Biobanks	National	This paper summarises Chinese biobanks thorough investigation and survey of operative, managerial, ethical conditions and challenges of biobanks	This study suggests that more than 20 different types of biobanks exist in China that can potentially be used as the data source for healthcare open data	This study does not provide a unified network to be applied to all the biobanks across China, which stops the sharing of specimen information	(Cheng et al., 2013)
		This paper provides a detailed overview of China's biobanking in (i) characteristics of the current biobank; (ii) development status; (iii) international cooperation; (iv) new prospects for Chinese biobanks	This study identifies the necessity to properly integrate specimens and related comprehensive information and establish sharing platform among biobanks for rational and efficient application of valuable biomedical information	This study lacks sufficient discussion on biobanks' data independence, integration, and communication	(Zhang et al., 2015b)
		This paper introduces a new project, the China Health Big Data, which combines low-dose chest computed tomography (LDCT) scan images and quantitative computed tomography (QCT)	This study showcases the possibility of a nationwide biobank for a large number (16) of HIS's information sharing.	This study is constrained only to the information on the lung cancer LDCT.	(Wu et al., 2019)
	City	This paper explores the conditions of the establishment of the Wuxi city tumour bank applied for molecular biological research	This study recognises the fact that the establishment of a tumour bank is helpful in sharing tumour information resources effectively	This study is constrained to a citywide tumour information resources sharing with have not covered too much information over inter-biobank medical information communication	(Li et al., 2004)
		This paper introduces the details of the establishment of a centralised biobank facility in a common storage environment	The study indicates a growing interest in integrating biomaterial repositories into larger infrastructures to meet research demands	This study's contribution is limited to the challenges of identifying and balancing different stakeholders' interests	(Zhang et al., 2015a)
Digital "big data" databases	National	This paper uses sequencing data published by the 1000 Genomes Project Consortium to construct the Virtual Chinese Genome Database (VCGDB)	The VCGDB developed by this study provides an interactive, user-friendly virtual Chinese genome browser (VCGBrowser) with seamless zooming and a real-time searching function for genome data sharing	This study's Chinese genome data (VCGDB) is still insufficient when considering the data required to illustrate the human genome's complexities fully.	(Ling et al., 2014)
		This paper introduces the National Genomics Data Center (NGDC),	This study indicates the effort to make the NGDC database publicly accessible	The study requires the improvement of the website accessibility,	(Xu et al., 2018b)

		developed by the Chinese Academy of Science	by creating a user-friendly web access portal	especially for downloading of international researchers	
		This paper investigates the epidemiological characteristics of mumps in mainland China from 2004 to 2018 to provide data and analysis for the key population to prevent and control mumps	The study shows how the mump-related healthcare data obtained from the National Population and Health Science Data Sharing Service Platform can be used for the analysis of the patterns and trends for mumps prevention and control	This study's data analysis has not applied advanced statistical analytic techniques in the data analysis, which limits the applicability and effectiveness of the suggestions made	(Jiang et al., 2019)
	Hospital	This paper tidies up and publishes the six opened-accessed standard electrocardiograms (ECG) databases to provide help for feature detection and disease diagnosis algorithms	The study shows the details and processes of making generalised and robust ECG databases that can increase the number of available 'opened' ECG databases	This study's ECG-related databases still need to be further improved with their size and diversity	(Gao et al., 2020)
Health facilities, networks, and community system	National	This paper discusses the current status and problems of public health information systems (PHIS) in China and shares lessons from a pilot project on automatic notifiable infectious disease reporting	The study has re-emphasised the five priority areas that need to be considered to achieve interoperability in health care applications: patient identifiers, semantic interoperability, data interchange standards, core data set and data quality	This study provides no technical support for data translation and exchange in HL7 messaging format in China.	(Su & Zhao, 2016)
	City	This paper presents a novel cloud platform for clinical information exchange between medical institutions based on web services	The study's web-based data exchange cloud platform uses the encrypted and compressed extensible markup language (XML) to ensure clinical information communication's interoperability and privacy	The study only suggests that the clinical information for the patient is still scattered, fragmented and isolated in different medical institutions	(Wu & Wang, 2014)
		This paper learns the lessons from the COVID-19 pandemic and focuses on epidemiological information-sharing strategies for medical institutions in megacities like Shanghai	This study suggests that the government should guide and support the sharing of epidemiological data among medical institutions	This study only suggests that during the patient's visit to the hospital during the pandemic period, the pre-examination staff cannot confirm the personal information provided by the patients	(Cai et al., 2020)
		This paper proposes a multicenter hybrid semi-supervised transfer learning model (MHSTL) based on a unified common data model to ensure multicenter data standardised representation	This study proves the possibility of solving the challenges of limited clinical data and the lack of labels by applying transfer learning to leverage large-scale patient data from multiple hospitals	This study only briefly points out the health infrastructure like the multicentre collaborative research network to share hospital-specific data	(Li et al., 2021)
		This paper analyses health system integration and develops	The study shows the patient's medical information can be communicated	This study only uses a case study to indicate that the county hospitals in	(Wang et al., 2016)

	recommendations for achieving integration	efficiently from upper-level institutions and institutions at the county level	rural China do not share e-health records with other institutions	
	This paper proposes a healthcare data analysis system for a regional medical union to support doctors from different hospitals in assessing patients' health conditions	The study suggests a redesigned clinical process can improve collaboration among general hospitals and community medical centres	This study only suggests some challenges found when patients are referred to different hospitals in the regional medical unions during treatment processes	(Xu et al., 2018a)
	This paper introduces the basic situation and information construction process of five hospital medical institutions	The study suggests that comprehensive interconnection and information sharing can improve healthcare services and innovate medical technology	The study only suggests that the interconnection of the hospital systems is still the biggest problem in developing China's health informatisation.	(Zhang et al., 2020b)
	This paper proposes a blockchain based secure and privacy-preserving electronic health records (EHR) sharing protocol	This study allows the data requester to search the relevant EHR information on the EHR blockchain with encryption ciphertext using the Ethereum platform	This study limits its data sharing limited to the text and numerical based EHR information without taking into account the sharing of graphical information (e.g., X-Ray)	(Wang et al., 2019b)
	This paper presents a patient-centric healthcare data management system with the blockchain storage technology	This study uses cryptographic functions to encrypt patient's data that ensure the pseudonymity	This study does not encounter the interoperability issue between different entities (e.g., diagnostic centre, hospital, doctors, patients) in the healthcare process	(Al Omar et al., 2019)
Hospital	This paper proposes a medical information sharing system to promote medical informatisation development for China's regional medical informatisation	The study suggests it is possible to develop a service pattern that realises the reasonable distribution and sharing of medical resources in a patient-oriented manner	This study has not provided a practical case to prove the designed platform's applicability and validity.	(Song et al., 2019a)
	This paper designs a cloud-based national diabetes community care system that consists of the cloud deployment, the architecture, and the applications model	The study proves that a planned community diabetic care system can be created through sharing information, knowledge, and techniques within a healthcare team	This study only suggests that China's community care services for diabetes mellitus (DM) are still at an early stage, and some existing IT-based community care systems for diabetes are not cost-effective due to the issue of data islands	(Yang & Cao, 2017)

475

4.2.4 Processes of public hospitals in China

The processes and procedures dimension of the socio-technical systems' analysis in the "open data" city healthcare context mainly investigates the current applications and research for China's detailed data related to healthcare information preparation, communication, and storage processes. The priority of this section is to identify the current situation, challenges, and demands for healthcare data information management processes. Therefore, two categories are identified for processes of public hospitals, given the found literature: healthcare service information management process, healthcare information collection and query and sharing processes (Table 5).

One of the wildly discussed topics in the platform or system related healthcare service information management process is tackling low-quality public health services. First, as the unevenly distributed medical resources in China, where resources were concentrated in mega cities' hospitals, there was a general lack of information-sharing platforms and regulations for many small or middle-sized hospitals. Some researchers made some attempts to explore the possibility of combining various latest technologies (e.g., mobile internet technology, cloud computing, IoT.) to create a cloud-based integrated service platform (Zeng & Wu, 2019). Besides the overspent of medical resources over large cities' first-tier hospitals, the shortage of medical personnel was another tough challenge many hospitals face. By integrating the data from some open data sources, some researchers come up with a mixed-integer linear programming (MILP) based optimization system to address the personnel scheduling procedure in the hospital to improve the treatment capabilities (Feng et al., 2021). In comparison, some researchers were more focused on tackling the problem of potential drug safety problems embedded in the drug distribution system and issuing processes due to the drug knowledge-sharing mechanism (Li et al., 2016). Similarly, some researchers looked into solutions for prescription errors due to doctors' career experiences (Bao & Jiang, 2016). By applying the big data-based data mining and recommender technologies utilizing the diagnosis history records, an error-prone way of prescribing medication for the future was designed (Bao & Jiang, 2016). The second group of fiercely discussed topics is about optimizing the supply chain management processes inside the hospital. With the help of the classification and regression tree (CRT) forecasting model, the inventory data was used to help provide a more consistent prediction for market demand based on previous transaction data (Xu & Tan, 2016). Also, scholars suggested that sharing and finding external data outside the healthcare institution is challenging (Li et al., 2020) and will cause service and cost implications to hospital supply chain

management (Xu & Tan, 2016). It was also suggested that the lack of a drug knowledge sharing mechanism and portal (Li et al., 2016) is not the only challenge faced by Chinese health workers. Due to the complexity and uncertainty of many diseases, China's drug procurement planning process has been traditionally bothered by drug overstocking or understocking. Countering that, a big data empowered deep learning (DL) network was used to optimize the drug procurement planning process (Q. Song et al., 2019). This research also used open city data like weather and temperature data and historical morbidities records to optimize drug procurement planning processes.

Through the review of the procedure and process -related healthcare "open data" articles, it is suggested that another group of articles have spent more paragraphs to emphasize the details of health information collection, query, and sharing procedures. Some scholars developed a multi-tiered electrocardiogram (ECG) network system to facilitate clinical data sharing and improve medical information resource integration (Gong et al., 2020). Other scholars tried to solve China's "difficulties and high expense in medical care" problem by effectively using medical resources through information sharing. Some researchers investigate the interactive behaviour between hospitals during the information sharing process based on China's recent constructed medical consortium (Zhang et al., 2017). Research suggested encouraging hospitals' information sharing process, reducing risk-sharing costs, and increasing information using capability are the keys for improving the information sharing processes (Zhang et al., 2017). Furthermore, the uneven distribution of the medical resource occurred not only among the different hospitals in a single city but also among the country's different regions. A Medical-Insurance-System-based Cancer Surveillance System (MIS-CASS) was built accordingly to generate and share cancer-related medical information to address the lack of cancer registries and lengthy cancer statistics collecting time (Tian et al., 2020). In the meantime, a series of rigorous procedures (e.g., data extraction, cleaning, processing, privacy protection, case ascertainment, etc.) were established to ensure the normal function of the whole system (Tian et al., 2020). Although more and more evidence suggests there is an unavoidable trend of "open data", unfortunately, the data sharing process in China was still not uniform with low data utilization and risk of health information security breach (Y. Liu et al., 2018). Countering that, some researchers discussed the potential causes of challenges faced by medical information sharing in China and proposed a regional data sharing strategy (Y. Liu et al., 2018). Except for the problem of inaccessible medical care, some research also focused

on the caring process of chronic disease. Chronic disease care is a long-term, complex process that requires the collaboration and coordination of different clinical professionals. A data mining approach was used to mine the temporal information dependency for type 2 diabetes patients data from three hospitals to understand the caring process's information flow (Sun et al., 2013). It was suggested that the discovered information dependency is vital for guiding the "open data" process and optimizing the governance of chronic disease care (Sun et al., 2013). Recently, some researchers turned their attention to the COVID epidemic. It was suggested that the information leakage and asymmetry of public health emergencies had been argued to distort and alienate the social governance during the COVID pandemic (Wu & Shan, 2021). It was, therefore, recommended by some scholars that artificial intelligence (AI) and machine learning (ML) can be used to enhance traceability and prediction, medical diagnosis and vaccine research with the help of some "big data" (Wu & Shan, 2021). It was also proposed that standardized data collection, sharing, and protection procedures are important in tackling the problem of lacking the proper summary and reflection on the emergency response experience (Wu & Shan, 2021) and the payment mechanisms for the eldercare facilities (Dai et al., 2021). Except for information leakage, the COVID pandemic also raised various challenges regarding epidemic prevention and control. Through exploring the multiple case studies, it was recognized that some latest technology like blockchain is useful not only for the prevention of patient medical information breaches but also the healthcare data sharing, real-time monitoring, identification of non-credible suppliers, and beguiling insurance claims (Hasan et al., 2021). The COVID pandemic locked people inside their shelters. Social media, as a result, played an important role as many citizens' primary information sharing and obtaining platform. Even before the pandemic, there was research focusing on analysing the information sources and authors' authority, as well as the information's length and format's influence during the "open data" processes (Fan et al., 2017).

In summary, there were two main focuses for the healthcare open data process aspects. These articles for healthcare open data process can be separated into two groups; one is platform or system related, and the other one is data collection, query, and sharing related. For the first group, the low-quality public health services caused by the uneven distribution of medical resources was the primary motivation for some research (Feng et al., 2021; Zeng & Wu, 2019). The second motivation for optimizing healthcare open data was to amend the problems or enhance the efficiency of the existing medical information management process (Bao & Jiang, 2016; Feng et

al., 2021; Li et al., 2016; Song et al., 2019b; Xu & Tan, 2016; Li et al., 2020). For the second group, some studies attempted to integrate the medical resource to ensure that the medical services and data are more accessible (Gong et al., 2020; Liu et al., 2018; Tian et al., 2020; Zhang et al., 2017), while other were designed for helping the epidemics (i.e., COVID) governance and information sharing (Fan et al., 2017; Hasan et al., 2021; Wu & Shan, 2021; Dai et al., 2021) and chronic disease care (Sun et al., 2013). Unfortunately, there are still some areas waiting to be improved from the process's perspective, including the unequally distributed medical resources between different ranked hospitals, the medical information leakage and low utilization, and the lack of proper control and environment for the management of medical data quality, preservation, and integration.

579 **Table 5.** Development of processes & procedures for open data healthcare

Category	Level	Objectives	Pros	Cons	References
Health information collection, query, and sharing process	National	This paper proposes a preliminary ingredient drug analysis system that collects all drug names from the website of the China Food and Drug Administration	The study shows the possibility of tackling the drug prescription safety problem within the drug distribution system and issuing processes through a drug knowledge-sharing platform	This study only points out the problem of the existing drug distribution system and issuing processes for traditional Chinese herbal medicine suffer from the ingredients overtaking	(Li et al., 2016)
		This paper documents a framework for answering how some of the major challenges associated with COVID-19 can be alleviated by leveraging blockchain technology	The study suggests that blockchain technology is useful for preventing breaching patient medical information and healthcare data sharing, real-time monitoring, and identification of non-credible suppliers	This study only presents facts about the COVID pandemic and raises various challenges regarding the epidemic prevention and control related to medical information like information leakage	(Hasan et al., 2021)
		This paper studies the application of artificial intelligence to social governance capabilities under public health emergencies	The study proves that artificial intelligence (AI) and machine learning (ML), with the help of medical big data, can enhance traceability and prediction, medical diagnosis and vaccine research	The study only suggests that the information leakage and asymmetry of the public health emergencies are currently distorting and alienating the social governance during the COVID pandemic	(Wu & Shan, 2021)
		This paper investigates the relationships between the characteristics of eldercare facilities and the payment mechanism	This study assists the informed decision-making on the PPP payment mechanisms of the elderly healthcare facilities using shared information from the Ministry of Finance database	This study limits its payment mechanism to viability gap funding (VGF) which is limited to three payment types	(Dai et al., 2021)
	City	This paper establishes an optimisation system with a two-layer mixed-integer linear programming for the hospital personnel scheduling	The study proves the data from the open-source platform can be utilised to analyse the scheduling issue of the healthcare management process	The study briefly suggests that research is still limited to using big or open healthcare data to forecast the future patient population, doctor scheduling and hospital structure	(Feng et al., 2021)
		This research made some attempts to explore the possibility of combining various latest technologies to create a cloud-based integrated service platform	The study suggests that the integrated service platform can be used as a quantitative management platform for government to manage grassroots health effectively	The paper briefly suggests that there is a general lack of information-sharing platforms and regulations for many small or middle-sized hospitals	(Zeng & Wu, 2019)
		This paper develops a multi-tiered electrocardiogram (ECG) network system to solve traditional ECG monitoring problems	The study suggests that a multi-tiered ECG network system can be built to facilitate clinical data sharing and	This study lacks proposing methods for improving effective intersectoral coordination and cooperation for the primary healthcare services' quality	(Gong et al., 2020)

		improvement in medical resource integration efficiency		
	This paper analyses the concept of big data and regional medical information sharing and attempts to promote medical information sharing in China	The study develops the strategy for regional medical information data sharing in the big data era from different aspects	The study only briefly points out that China's current data sharing process is still not uniform with low data utilisation, risk of health information security, and unsatisfied data quality management and integration	(Liu et al., 2018)
	This paper develops a medical-insurance-system-based cancer surveillance system (MIS-CASS) to determine cancer diagnosis	The study proves the applicability of using MIS-CASS to generate and share cancer-related medical information to amend the lack of cancer registries and length of cancer statistics collection time	This study only briefly suggests the fact that the uneven medical resource distribution occurs not only among the single city's hospitals but also between the country's regions	(Tian et al., 2020)
	This paper studies the problem of medical information sharing based on the medical consortium constructed by the Chinese government	The study suggests encouraging hospitals' information sharing behaviour, reducing risk sharing costs, and increasing information using capability are keys to improving the information-sharing processes	This study only points out that hospital information sharing can cost a lot due to a non-universal and imperfect medical environment without providing a practical solution to the problem	(Zhang et al., 2017)
	This paper focuses on the influence of the authority of information sources and authors as well as the format and length of information	The study shows that the information format and authority subscriptions are both important and have a huge impact on the shared information's attractiveness and credibility	This study only briefly suggested that the existing research for analysing open data or medical information sharing on the social media platform is still limited	(Fan et al., 2017)
	This paper proposes an approach to main temporal information dependency in outpatient encounter records	The study suggests that the discovered information dependency is vital for guiding the "open data" process and optimising the governance of chronic disease care	This study suggests that fact that physicians suffer from information overload and have to spend lots of time seeking relevant information in their daily practice	(Sun et al., 2013)
Hospital	This paper designs and implements a universal medicine recommender system framework to the recommendation system to reduce prescription errors	The study suggests it is possible to use an open dataset with data mining and recommender technologies to explore potential knowledge from the records of diagnosis history	This study only focuses on the data cleaning process and the algorithms used for building recommender technology	(Bao & Jiang, 2016)
	This paper proposes a big data-driven approach to predicting the morbidities of acute gastrointestinal infections	The study shows the potential benefit of open city data like weather and temperature data and historical morbidity records to optimise the healthcare management process	This study only points out China's traditional drug procurement planning process's drug overstocking or understocking issues due to inadequate medical information	(Song et al., 2019b)

	This paper demonstrates a case study on classification and regression tree (CRT) forecasting model of inventory data to predict the market demand	This study suggests that inventory data can be used to help provide a more consistent prediction for market demand based on previous transaction data	This study only suggests China's current experience-based supply chain decision-making has suffered the changing policies and license delays problems	(Xu & Tan, 2016)
	This paper proposes a functionality-based assessment approach of hospital resilience to the hazard scenario	This study improves the development of effective engineering measures and informed policies to enhance the seismic resilience of hospital	This study does not provide a hospital seismic resilience information sharing mechanism that can benefit other similar medical institutions	(Li et al., 2020)

580

4.2.5 Technologies (digital tools) developed for sharing and opening healthcare data

The technologies used in the clinical practice in terms of data sharing and opening are about the technology framework and tools utilized based on the hospital information system, regional health information system and multi-level data-sharing platform. Therefore, according to different application scopes of these technologies, literature can be divided into three categories: technologies applied within the hospital; technologies utilized for information exchangeable among various medical institutions; and technologies involved in connecting external non-native healthcare databases (Table 6).

The acquisition and storage of information within the individual hospital are the basis of sharing and openness of healthcare data. The key technical support was guaranteed by the information integration among Hospital Information System (HIS), more precisely, Clinical Information System (CIS), Electronic Medical Record System (EMRS), Pharmacy Information System (PIS), Financial Information System (FIS), Radiology Information System (RIS) and Picture Archiving Communication System (PACS), for the sake of medical information sharing and hospital efficiency enhancement (Chang et al., 2003). This integration procedure involved recognition, translation and transmission of multimedia information flow containing the data of patient profile, registration, clinical images, healthcare reports and medical expenditure etc. (Chang et al., 2003; Yan et al., 2021), and the corresponding health data privacy and security scheme (Xu et al., 2019). For many years, the mainstream of medical informatics development in China has been establishing and improving the HIS through Information and Communication Technology (ICT). Recently, with the development of big data technology, deep neural networks (Song et al., 2019), data processing approach, and machine learning were utilized to make useful explorations in the integration of HIS.

Information transfer among different medical institutions is a milestone for healthcare open data. The realization of such a process mainly depends on the development of cloud computing technology. The cloud-based medical information approach helped to build a comprehensive regional medical system to collect and analyse large-scale medical and biological big data from hospitals, biobanks and other related institutions for clinical diagnosis, laboratory tests and, more advanced, smart healthcare (Zeng & Wu, 2019). Many scholars in China explored relevant technology frameworks in multi-source data integration in the field of healthcare. Data labelling and coding (Cui et al., 2015), data mining (Hu et al., 2017), Bluetooth (Zeng & Wu, 2019),

artificial intelligence and machine learning techniques such as transfer learning were demonstrated in the literature for online registration, diseases diagnosis and disease treatment by building data-driven collaborative models (Li et al., 2021a). Meanwhile, information security for personalized data was emphasized, along with the utilization of medical big data. The key technologies in healthcare privacy protection included data encryption (Wang et al., 2019a), data anonymity, access control, identity authentication and management, authorization, and auditing (Hu et al., 2016).

The ultimate level of healthcare data sharing and openness is beyond the information exchange among medical institutions. For a healthcare open data city, personal data from electronic sensors, smart wear and mobile Internet could be utilized for disease surveillance and prediction. Moreover, multiple external datasets from government and other non-medical institutions were able to be employed for medical service. For example, during COVID-19, location-based databases like Baidu (Chao et al., 2018; F. Villanustre et al., 2021) were used by scholars to predict the spread of disease and build hospital recommendation systems. In addition, big data analytic techniques, geographic information system (GIS) technology (Yan et al., 2013), and mobile Internet technology (Wang et al., 2017) were also implemented appropriately. However, it is also recognised that some digital technologies for open data healthcare have not been established well and are waiting for further investigations, such as electronic signature (e-signature) and the Internet of Things (IoT).

To sum up, recent literature on technologies related to medical data sharing mainly focused on big data analysis, connecting different databases from multiple institutions. For technologies applied within the hospital, the information integration process has involved various HIS information flow (Chang et al.; Yan et al., 2021). Various different modern techniques like Bluetooth (Zeng & Wu, 2019), data mining (Hu et al., 2017), data sharing (Cui et al., 2015), neural network (Li et al., 2021; Song et al., 2019b), location-based databases (Chao et al., 2018; Villanustre et al., 2021), geographic information system (GIS) technology (Yan et al., 2013), blockchain (Xu et al., 2019; Wang et al., 2019a), intelligent shelter hospital system (Zhou et al., 2022), and mobile Internet framework (Wang et al., 2017) have been widely tested and applied in the healthcare big data practice. With the increased awareness in medical crisis management and disease surveillance, GIS techniques and technologies on multi-source data systems are increasingly involved in the

642 healthcare open data framework. However, there were still many areas of improvement from the
643 perspective of technologies. They include the lack of inter-hospital information sharing and the
644 data qualities issues when applying the healthcare technologies.

645 **Table 6.** Development of technologies for open data healthcare

Category	Level	Objectives	Pros	Cons	References
Within hospital	Hospital	This paper discusses the integrating method and working procedure of a digital HIS	This study adopts a unique modularised structure to integrate HIS, RIS and PACS and applies the structure in reality	This study only considers the information exchange within the hospital and does not consider a higher level of information sharing	(Chang et al., 2003)
		This paper utilises the big-data and deep neural network to optimise the drug procurement planning	This study presents a good combination of big data driven approach (deep neural network) and the medical and healthcare field	This study lacks a further discussion on a higher level of healthcare data sharing	(Song et al., 2019b)
		This paper presents a big data analytics platform for the gastroenterology department in a Chinese hospital	This paper introduces the data centre operating system, computing techniques like machine learning, and data processing approaches for medical information	This study is limited in the types of diseases and data range	(Yan et al., 2021)
		This paper proposes a large-scale health data privacy preserving scheme using blockchain technology to prevent privacy leakage and single-point bottleneck	This study allows patients to revoke or add authorised doctors for key management with avoiding medical disputes by disallowing the deleting of IoT data and doctor's diagnosis	This study fails to consider the patient diagnosis information's integration with other existing hospital information systems	(Xu et al., 2019)
		This paper gains experience in the development and utilisation of an intelligent system in Fangcang shelter hospital	This study provides an intelligent platform that is capable of delivering remote consultation and guidance for the implementation of the intelligent shelter hospital system	This study lacks the providing of solutions for ensuring the reliability of the cloud data platform and proper measures for the IoT users' privacy protection	(Zhou et al., 2022)
Between hospitals	National	This paper takes the cloud-based hospital information system as a service for grassroots healthcare institutions	This study emphasises a cloud-based hospital information approach and proposes a cloud-based medical service delivery framework in healthcare data sharing	This study lacks a further discussion on healthcare data sharing from external databases	(Yao et al., 2014)
	City	This paper shows an integrated medical information sharing platform based on the privacy protection system	This study shows key technologies in privacy protection of healthcare data: data encryption, data anonymity, access control, identity authentication authorisation and auditing	This study needs to include more administrative strategies and strategic thinking in a city-level	(Hu et al., 2016)
		This paper targets to improve the data security and the accuracy of the online medical diagnosis service	This study develops a privacy-preserving collaborative model learning scheme that learns a global	This study's privacy-preserving collaborative model learning scheme	(Wang et al., 2019a)

			diagnosis model through local diagnosis models with sensitive medical data protection	requires a trusted authority to issue the key distribution	
	Hospital	This paper presents a cloud-based electronic medical record (EMR) system for medical data sharing between healthcare institutions	This study applies the cloud computing technology and data mining approach in healthcare data sharing between hospitals	This study needs to be improved with a mobile-based end and more comprehensive data	(Hu, et al., 2017)
		This study presents a sample repository and information sharing biobank platform	This study mentions the data labelling and coding techniques for biobank information management	This study's data sharing is limited to medical sharing, which can be improved if added external databases	(Cui et al., 2015)
		This paper presents a multicentre hybrid semi-supervised transfer learning model (MHSTL)	This paper introduces machine learning techniques such as transfer learning in healthcare data sharing	This study lacks a further discussion on a higher level of healthcare data sharing in a city-level	(Li et al., 2021)
		This paper introduces the open HER and OMOP CDM as open source tools for medical distributed data	This study develops a healthcare data platform from distributed sharing data, emphasising an open EHR system	This study lacks a further discussion on a higher level of healthcare data sharing	(Li & Tsui, 2020)
		This paper provides a healthcare service platform using mobile Internet technology, cloud computing, Internet of Things, big data etc	This study introduces smart healthcare and Bluetooth technology in healthcare data sharing and proposes real-time health monitoring with intelligent technology	This study needs to be improved from the perspective of a higher level of data sharing, not constrained within the context of healthcare data	(Zeng & Wu, 2019)
External databases	Hospital	This paper uses Baidu large-scale location-based service (LBS) big data to guide and disperse outpatients	This study presents a long short-term memory (LSTM) based deep learning method in healthcare data sharing and utilises the external platform like Baidu large-scale LBS logs database	The study fails to have a further discussion on a city-level healthcare data sharing	(Chao et al., 2018)
		This paper describes a model of Corona spread to model and track COVID-19 cases by using innovative big data analytics techniques	This study uses mathematical compartmental models on personal relationships, health centre location and spread mechanisms of Coronavirus to predict the behaviour of disease outbreak	The study cannot explain other factors that can affect the trend, such as mobility, weather and others; More comprehensive features need to be added to the model	(Villanustre et al., 2021)
		This paper describes an electronic surveillance system for the early detection of infectious disease epidemics in rural China	This study emphasises the role of the electronic surveillance system, combined with modern communication and GIS technology	This study needs to be improved with longer time series of data in the formal implementation with more surveillance units	(Yan et al., 2013)
		This paper discusses the challenges with the hierarchical medical system in China and how big data and mobile	This study focuses on mobile medical care, pointing out the functions of sensors, smart-watches, and mobile health apps in healthcare data sharing	The study fails to have a further discussion on a city-level healthcare data sharing	(Wang et al., 2017)

	Internet can be used to mitigate these challenges

646

5. Discussion

5.1 Achievements and gaps of current research

After the longitudinal analysis of healthcare open data development from five perspectives, it is possible to re-evaluate how the new technology and healthcare services introduced have changed China's HIS related information flow (Figure 1). There were four major developments (linked with red dash lines) compared with the classic 2006 HIS information flow: (1) the internal hospital network (with the system server group and database); (2) the central management information centre; (3) the telemedicine centre; and (4) the external information outlets. From the culture's perspective, compared with the classic 2006 information flow framework, the hospitals' internal medical information sharing (Chen et al., 2018; Li et al., 2019; Luo & Ping, 2011; Wan et al., 2012; Yin et al., 2020; Zhu, 2011) has been enhanced with the establishment of the hospital system information management system and the hospital database. While the requirement of government medical information disclosure (Lv, 2018; Ma et al., 2015; Shangguan et al., 2020; Zhang et al., 2006) has led to the development of many hospitals' general information query systems. From the People's perspective, the general decision-making/information query system (Ahamed & Ramkumar 2016) has also met the need for more evidence-based and flexible treatment plans and decision-making (Lv, 2018; Ma et al., 2015; Shangguan et al., 2020; Zhang et al., 2006). At the same time, the emergence of social network platforms (e.g., WeChat) (Fan et al., 2017) has contributed to the development of a multimedia healthcare consultation system. From the infrastructure's perspective, the building of system server group and hospital databases and the regional healthcare interface made digital "big data" databases- cloud-based community care systems and platforms (Wu & Wang, 2014; Song et al., 2019a; Su & Zhao, 2016; Yang & Cao, 2017) and regional or city-wide health infrastructure (Cai et al., 2020; Li et al., 2021; Wang et al., 2016; Xu et al., 2018a; Zhang et al., 2020b; Wang et al., 2019b; Al Omar et al., 2019) become possible. From the process's perspective, the increase of hospital information outlets like the national insurance interface has made the integration of the insurance system possible (Tian et al., 2020). Finally, from the technology perspective, due to the demand for information integration of various HIS' information flow (Chang et al.; Yan et al., 2021), hospital system information management system is set up in some hospitals for the coordinating of different HIS. The setting-up of general statistics/AI analysis system has further benefitted prediction, medical diagnosis with the use of "big data" (Wu & Shan, 2021), the data analysis using neural networks (Li et al., 2021;

Song et al., 2019b). However, the development of the healthcare open data related to HIS cannot conceal the fact that many gaps are still waiting to be explored.

Gaps are raised from the different perspectives of socio-technical systems for open data cities. For culture, many studies indicated that there is a lack of consideration of healthcare data information sharing for the city as a whole (Yin et al., 2020; Zhu, 2011), as there is no culture initiative for sharing medical information between hospitals and non-medical institutions (Wan et al., 2012). It is also suggested that, from the culture perspective, the promoting of healthcare data sharing for some projects mainly stays at the policy level. There is a lack of practice cases (Lv, 2018) and a variety of formats and channels of medical information sharing (Ma, 2015; Shangguan et al., 2020; Yu, 2006). From the management of open healthcare data at the city level, there is a lack of enough information and practical cases for both real-world projects and declaring the ownership of open healthcare data (Chen et al., 2015; Han et al., 2021; Lv, 2017; Lv & Qiao, 2020; Yang, 2017). Furthermore, although some researchers have discussed the protection of patient's medical information especially under the background of big data (Han et al., 2021; Yang, 2017; Lv & Qiao, 2020; Chen et al., 2015), there still lacks studies that systematically analyse the critical legal issues involved for the medical information disclosure and the initiatives and barriers for government policy maker to participate in the healthcare open data. From people's perspective, it is demonstrated that the privacy of patients' health-related information lacks proper protection from a legal perspective (Tian et al., 2020). The communication between patients and physicians lacks efficiency. Social media like WeChat is currently the only channel the public and patients rely heavily on to receive medical information sharing (Fan et al., 2017; Y. Liu et al., 2018). The lack of sufficient information is also raised as the biggest challenge for policymakers to make appropriate policies (Lee, 2020). For the city, building, and infrastructure, the "data islands" issue has suffered by many of the open medial data infrastructures like biobanks and hospitals from both inner-city and between cities (Cheng et al., 2013; Li et al., 2004; Y. Zhang et al., 2015). The integration is lacking from different perspectives like: balancing the interests of different stakeholders (L. Zhang et al., 2015), compromising the different requirements of databases (Gao et al., 2020; Ling et al., 2014), biobanks (Wu et al., 2019), general clinical information (Wu & Wang, 2014; Li et al., 2021a; Wang et al., 2016; B. Xu et al., 2018; Q. Zhang et al., 2020), and providing or adopting a better format of sharing (Su & Zhao, 2016; X. Xu et al., 2018) in both city level and national level. While for the processes and procedures perspective, it is found that the

information sharing platform and resources for many small and middle-sized hospitals are still lagged behind (Tian et al., 2020; Zeng & Wu, 2019). There is a lack of universal and quality assured hospital information sharing process and cooperation environment (Gong et al., 2020; Y. Liu et al., 2018; Zhang et al., 2017), which hinders the applicability of the use of open healthcare data in areas like level supply chain management (Xu & Tan, 2016), doctor scheduling (Feng et al., 2021), and drug distribution (Li et al., 2016; Q. Song et al., 2019). From the technology perspective, it is suggested that there is a lack of inter-hospital information sharing for the existing healthcare information technology communication structure (Chang et al., 2003; Cui et al., 2015; Li & Tsui, 2020; Li et al., 2021b; Q. Song et al., 2019; Yao et al., 2014; Zeng & Wu, 2019). The data shared and used in some projects are fragmented, non-standardized, and error-prone (Cui et al., 2015; Hu et al., 2017; Li et al., 2021a; Li et al., 2021b; Zeng & Wu, 2019). Last but not least, the introduction of many new technologies (e.g., smart IoT for telemedicine (Zhou et al., 2022), machine learning (Yan et al., 2021), etc.) and systems (e.g., cloud-based electronic medical record (EMR) system (Hu et al., 2017), cloud-based hospital information system (Yao et al., 2014), etc.) to the healthcare open data environment often requires the integration of various sensitive information (e.g., patients' diagnosis information). However, there is currently a lack of studies that can provide measures for protecting new system users' data privacy, especially during the data integration.

5.2 Mission of open data city for healthcare

To fill the gaps of the existing open data city for healthcare, a city-wide healthcare open data sharing framework based on the socio-technical system's five different dimensions is developed to provide guidance and raise the upcoming missions for a better city level healthcare open data-sharing environment (Figure 6).

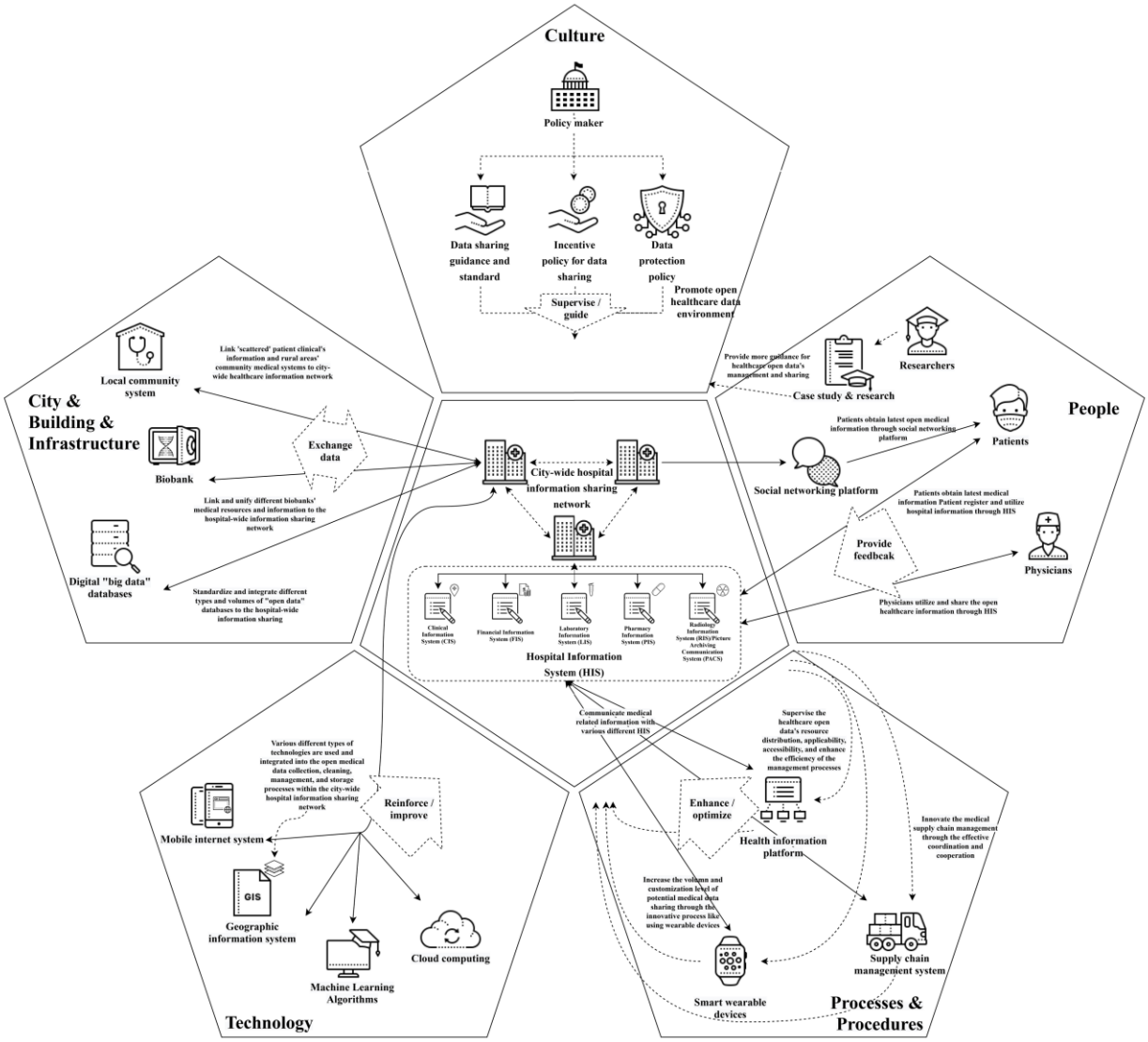


Figure 6 The open data city healthcare reinforcement framework from social-technical system's five dimensions

From the culture aspect, the focus of the open data should be put on how to leverage the government policies to improving the open healthcare data environment. Therefore, the mission for the city-level healthcare open data is to let the government supervise and guide the building of a healthy open healthcare data culture environment. From people's perspective, the concentration is on the different healthcare institution's stakeholders. The mission of the researchers is to provide more case studies and research that can facilitate the development of open healthcare data sharing and management. While for patients and physicians, the mission is to enhance the communication channel of medical information so that the feedback can be provided in an effortless manner. While

for the city, building, and infrastructure perspective, the target is the resources and user points of medical “big data”. The mission is to promote the data exchange between the existing city-wide hospital information sharing network and the established biobanks and digital open healthcare databases and the integration of the local community medical service point’s information system with the city-wide network. For the processes and procedures aspect, the contents discussed is related to many different applications, platforms, and systems that are covered in other social-technical system aspects. However, the concentration is put on the processes themselves rather than the others. The mission of this aspect of the social-technical system is to explore how to improve the utilization efficiency and integration of the healthcare open data to enhance and optimise healthcare data sharing practice. For the final aspect social-technical system, the technology, the focus is on how the currently available technologies can benefit the open healthcare data sharing environment. The mission is to explore the ways of using existing technologies (e.g., mobile internet system, geographical information system, machine learning algorithms, etc.) to improve and reinforce the collection, cleaning, management, and sharing of city-wide healthcare open data.

Recommendations, therefore, were made to help achieve these missions (Figure 6). From the culture perspective, there are three different types of policies recommended to be focused by the city or country’s government and policymakers, include: (1) the guidance and standard for data sharing that can help set up and unify the criterion and technical details for the medical information sharing; (2) the incentive policy for data sharing to initiate the pilot projects and studies for the exploring the applicability and possibility of the open healthcare data; and (3) the data privacy protection policy that outlines the patients' related information storage and destruction requirements and the ownership of the data, which is frequently undermined especially under the current pandemic environment. Due to the importance of healthcare data, government institutions should take the leader role in setting up legal requirements, accountability and the ownership of the healthcare open data.

The recommendations for the people perspective are majorly related to patients. The target stakeholder for this perspective is the patient. It is suggested that, in addition to a more efficient and less error-prone hospital patient information registration and communication channels, more resources should also be put on exploring the possibility of other means of open healthcare data

communication channels like social networking platforms (e.g., WeChat) and mobile apps. How to control the quality and level of medical information that can be obtained through these channels are crucial as more and more patients are relied on these methods for gaining medical information.

The city, building, and infrastructure recommendations are mainly targeted at the healthcare data centres and biobanks' managers. These managers with the management of similar medical information and resource types need to sit together and discuss the way how these resources can be connected and shared in a more efficient and standardised manner.

It is suggested that more research and case studies should be conducted from the processes and procedures perspective to explore how the use of healthcare open data can benefit the key services of medical service management. For instance, the study can be conducted to explore how the use of external healthcare open data can increase the efficiency of hospital supply chain management. Also, it is suggested that technology and financial support should be given especially to the community and rural areas' medical service points to ensure the city-wide medical information can be easily accessed by the medical professionals there. From this perspective, researchers from both the academic and industry should shoulder the responsibility of pilot research and case study development.

Finally, from the technology perspective, many current medical-related technologies and applications (e.g., mobile internet systems, machine learning applications) are still majorly focused on separately within a single hospital, between hospitals, or external databases. It is recommended that those researchers and medical professionals can combine the study covered topic to include both the utilisation of external (e.g., medical information from private institutions or social networks) and different types of potentially useful information (e.g., weather information), and how this information can be connected and integrated to the existing hospital HIS or communicated within the city-level healthcare open data network. Besides, the technology like blockchain should be applied to prevent the leakage of patients' medical information and avoid the potential medical dispute due to the unauthorised changing of a doctor's diagnosis. Both the scientists (e.g., big data scientists), industry technology developers (e.g., blockchain developers), and medical professionals should work together to safeguard patients' medical information and the adoption of new technologies.

5.3 Challenges of open data city for healthcare

The major challenge comes from the integration and the connection of healthcare data in the open data city. This big challenge can be further decomposed into several small confrontations. Firstly, it lacks a national level standard for the transfer of clinical and administrative healthcare data like HL7 (Su & Zhao, 2016). Although policymakers have issued "medical bid data" and "Internet + Medical Health" related policy (Wu et al., 2020), there is no "open healthcare data" specific detailed policies and regulations guidance available. The governance structure and healthcare infrastructure are needed for data management in healthcare (Alhassan et al., 2018; Sharma et al., 2021). Although some large hospitals have started to explore the possibility of healthcare open data, the sharing and using of open medical information in community healthcare service points and rural medical stations are still in infancy (Wang et al., 2016). Moreover, from a technological point of view, there is no proven standardised and effective solution or platform to connect and share medical information at the entire city level. In the process of integrating and sharing healthcare information and resources, it is suspicious to encounter resistance of information sharing between hospitals due to conflicts of interests. For instance, hospitals might be reluctant to share patient-related medical information with their major competitor due to the potential grabbing of the clientele. All the above-mentioned confrontations will challenge the integrating and sharing of the city level healthcare data. It is also suggested the issues like patient information security will potentially let the progress of open city-level healthcare data be overshadowed. Fortunately, the use of blockchain might be a possible solution for safeguarding the information security and the data ownership issue (Xu et al., 2019; Wang et al., 2019b; Al Omar et al., 2019). The discussion over the protection of patient information security at the country level is still at a relatively early stage, with no specific detailed regulations and guidance on the ownership, management, and destruction of patient-related information.

6. Conclusions

This research has conducted a systematic literature review of the previous accomplishments in open data city for healthcare. First, this study provided a critical evaluation of the existing hospitals' HIS's open data development (Figure 1) from a different angle. Then the study further reviewed the development of healthcare open data in five socio-technical systems' dimensions (Figure 2-6). As a result, this study has solved three research questions (section 1) in the field of healthcare open data. To answer the first research question, a mixed-method based review of 79 out of 4,611

articles has been conducted in open data city healthcare in China from the five perspectives of the socio-technical system, including (1) culture, (2) people, (3) infrastructure, (4) processes, and (5) technologies. Based on the latitudinal analysis, it is found that the development of open data city for healthcare is still at its early stage with biases, being policy-driven and focusing on the construction of healthcare infrastructures and adoption of technologies.

Insights have been further provided through the longitudinal analysis of the existing environment in open data city healthcare in China. To answer the second research question, the research gaps and missions of the current studies have been identified in the five perspectives of socio-technical system. These gaps include: (1) from culture's perspective, the lack of culture initiative, practical cases, and consideration for the city as a whole; (2) from people's perspective, the lack of legal protection, information communication efficiency, and sufficient information for policy maker; (3) from infrastructure's perspective, the lack of data connection, databases requirements, and guidance for the format of sharing; (4) from process's perspective, the lack of information sharing quality ensuring, cooperation environment, and platforms for small-sized hospitals; (5) from technology's perspective, the lack of inter-hospital information sharing, data privacy protection, and completed, standardised, and correct data. For the third research question, in order to fill these gaps and achieve the missions, an open data city healthcare reinforcement framework has been proposed for strategies of future development: (1) from culture perspective, three kinds of policies are encouraged for data sharing and protection, including the guidance and standard for data sharing, the incentive policy for data sharing, and the data privacy protection policy; (2) from people perspective, the use of social networking platform should be quality controlled and more specified for healthcare; (3) from city, building, and infrastructure perspective, medical infrastructures' managers should negotiate and explore how their resources can be connected and shared in an efficient and standardised manner; (4) from processes and procedures perspective, it is recommended to focus on the use of healthcare open data for benefits of key medical services; and (5) from technology perspective, research and practice are encouraged to be focused on healthcare open data sharing and integration at the city level.

This study provided three main contributions to the body of knowledge and practice: (1) This study contributes to the understanding of cities' healthcare open data development and management; (2) This study uses Chinese cities as an example to provide potential gaps, missions, and

recommendations for other countries' city-wide healthcare open data development, after taking into account the different countries' healthcare system structures and engineering technologies, regulations and policies, and the interoperability of different HIS and databases; (3) This study provides an open data city healthcare reinforcement framework for the guidance of healthcare city open data development and engineering related management practice.

This study has two major limitations. The first limitation is its focus of analysis and recommendations were targeted at the city level. The second limitation is that when applying the framework and recommendations developed by this study to other countries, although amendments can be made during the application, the types and extent of amendments are unknown. Therefore, future work is needed to evaluate the healthcare open data in different target levels and countries.

This research contributes a new multi-dimensional way to rethink the development of open data city in healthcare and helps establish the state-of-the-art open data city for future research and practice.

Data Availability Statement

All data, models, and code generated or used during the study appear in the submitted article.

Acknowledgement

This research is supported by the grants from the UCL Global Engagement Strategic Partner Funds and UCL Centre for Blockchain Technologies (CBT).

References

Ahamed, B., & Ramkumar, T. (2016). "An intelligent web search framework for performing efficient retrieval of data." *Computers & Electrical Engineering*, 56: 289–299.

Alamo, T., Reina, D., Mammarella, M., & Abella, A. (2020). "Covid-19: Open-data resources for monitoring, modeling, and forecasting the epidemic." *Electronics* (Basel), 9(5): 827. Multidisciplinary Digital Publishing Institute.

Alhassan, I., Sammon, D., & Daly, M. (2018). "Data governance activities: A comparison between scientific and practice-oriented literature." *Journal of Enterprise Information Management*. Emerald Publishing Limited.

Al Omar, A., Bhuiyan, M., Basu, A., Kiyomoto, S., & Rahman, M. (2019). "Privacy-friendly platform for healthcare data in cloud based on blockchain environment." *Future generation computer systems*, 95: 511–521. Elsevier

Austin, C., & Kusumoto, F. (2016). The application of Big Data in medicine: current implications and future directions. *Journal of Interventional Cardiac Electrophysiology*, *47*(1), 51-59.

Bao, Y., & Jiang, X. (2016). An intelligent medicine recommender system framework. *IEEE 11Th conference on industrial electronics and applications (ICIEA)*, 1383–1388. IEEE.

Begun, J., Zimmerman, B., & Dooley, K. (2003). Health care organizations as complex adaptive systems. *Advances in health care organization theory*, *253*, 288.

Berry, I., Soucy, J., Tuite, A., & Fisman, D. (2020). Open access epidemiologic data and an interactive dashboard to monitor the COVID-19 outbreak in Canada. *Canadian Medical Association Journal*, *192*(15), E420-E420.

Bonzo, S., McLain, D., & Avnet, M. (2016). Process modeling in the operating room: a socio-technical systems perspective. *Systems Engineering*, 19(3), 267-277.

Cai, Q., Mi, Y., Chu, Z., Zheng, Y., Chen, F., & Liu, Y. (2020). Demand Analysis and Management Suggestion: Sharing Epidemiological Data Among Medical Institutions in Megacities for Epidemic Prevention and Control. *Journal of Shanghai Jiaotong University (Science)*, *25*(2), 137-139. https://doi.org/10.1007/s12204-020-2166-3

Cao, X. (2006). Research on Hospital Informatization and Hospital Business Process Reorganization. Huazhong University of Science and Technology. Accessed August 08, 2022. https://wap.cnki.net/touch/web/Dissertation/Article/10487-2008026392.nh.html. (Chinese)

Chang, Z., Mei, S., Gu, Z., Gu, J., Xia, L., Liang, S., & Lin, J. (2003). Realization of integration and working procedure on digital hospital infon-nation system. *Computer Standards & Interfaces*, *25*(5), 529-537.

Chao, H., Cao, Y., Zhang, J., Xia, F., Zhou, Y., & Shan, H. (2018). Population Density-based Hospital Recommendation with Mobile LBS Big Data. *2018 IEEE international conference on big data and smart computing (BigComp)*. 37-44. IEEE.

Chen, H. D., Chan, B., & Joly, Y. (2015). Privacy and Biobanking in China: A Case of Policy in Transition. *Journal of Law Medicine & Ethics*, *43*(4), 726-742.

Chen, L., Lu, Q., & Zhao, X. (2020). Rethinking the construction schedule risk of infrastructure projects based on dialectical systems and network theory. *Journal of Management in Engineering*, 36(5), 04020066.

Chen, X., Liu, Z., Wei, L., Yan, J., Hao, T., & Ding, R. (2018). A comparative quantitative study of utilizing artificial intelligence on electronic health records in the USA and China during 2008–2017. *BMC Medical Informatics and Decision Making*, *18*(5), 117.

Cheng, L., Shi, C., Wang, X., Li, Q., Wan, Q., Yan, Z., & Zhang, Y. (2013). Chinese biobanks: present and future. *Genetics Research*, *95*(6), 157-164.

Clegg, C. (2000). Sociotechnical principles for system design. *Applied ergonomics*, *31*(5), 463-477.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*(3), 273-297.

Cui, W., Zheng, P., Yang, J., Zhao, R., Gao, J., & Yu, G. (2015). Integrating Clinical and Biological Information in a Shanghai Biobank: An Introduction to the Sample Repository and Information Sharing Platform Project. *Biopreservation and Biobanking*, *13*(1), 37-42.

Dai, K., Li, S., In Kim, J., & Jae Suh, M. (2021). Identifying Characteristics of PPP Projects for Healthcare Facilities for the Elderly Based on Payment Mechanisms in China. *Journal of Management in Engineering*, 37 (6): 05021009.

Davis, M., Challenger, R., Jayewardene, D., & Clegg, C. (2014). Advancing socio-technical systems thinking: A call for bravery. *Applied ergonomics*, *45*(2), 171-180.

Fan, J., Zeng, Y., & Shao, M. (2017). *How to Improve the Credibility and Interestingness of Social Media Healthcare Information?* Univ Calgary Press. <Go to ISI>://WOS:000412149300053

Fayoumi, A., & Williams, R. (2021). An integrated socio-technical enterprise modelling: A scenario of healthcare system analysis and design. *Journal of Industrial Information Integration*, 23, 100221.

Feng, D., Mo, Y., Tang, Z., Chen, Q., Zhang, H., Akerkar, R., & Song, X. (2021). Data-driven hospital personnel scheduling optimization through patients prediction. *CCF Transactions on Pervasive Computing and Interaction*, *3*(1), 40-56.

Gao, H., Liu, C., Shen, Q., & Li, J. (2020). Representative databases for feature engineering and computational intelligence in ECG processing. In *Feature Engineering and Computational Intelligence in ECG Monitoring* 13-29.

Gong, Z., Zhang, M., Qua, Y., & Xia, Q. (2020). Multi-Tiered Electrocardiogram Network System: Application In a Regional Medical Association In PR China. *Southeast Asian Journal of Tropical Medicine and Public Health*, *51*(5), 785-794.

Han, P., Gu, L., & Zhang, J. (2021). Research on the Willingness of Medical Data Sharing Considering Privacy Protection - From the Perspective of Evolutionary Games. *Journal of Modern Information, 41*(03), 148-158. (Chinese).

Hasan, M., Deng, S., Sultana, N., & Hossain, M. (2021). The applicability of blockchain technology in healthcare contexts to contain COVID-19 challenges. *Library Hi Tech, 39(3)*, 814-833.

Hu, S., Zhou, L., Dong, N., Zhou, Y., Gao, Z., Xu, J., Liang, Z., Tian, T., Jiang, Q., Liu, Y., Burrage, K., Song, J., Wang, Y., Hu, X., Morishita, S., Zhu, Q., & Wang, G. (2016). The Design and Implementation of the Privacy Protection System of a Regional Health Information Platform. *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2016: 964-969.

Hu, S., Lu, L., Jin, X., Jiang, Y., Zheng, H., Xu, Q., Cai, F., Meng, Y., Zhang, C. Hu, X., Shyu, C., Bromberg, Y., Gao, J., Gong, Y., Korkin, D., Yoo, I., & Zheng, J. (2017). The Recommender System for a Cloud-Based Electronic Medical Record System for Regional Clinics and Health Centers in China. *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 1021-1024.

Hunter, E., Mac Namee, B., & Kelleher, J. (2018). An open-data-driven agent-based model to simulate infectious disease outbreaks. *Plos One*, *13*(12).

Hutchins, E. (1995). *Cognition in the Wild*. MIT press.

Jiang, R., Yin, Q., Xu, M., Zhao, Z., Deng, Y., & Che, Y. (2019). Epidemiological characteristics of mumps in mainland China from 2004 to 2018 and key population for prevention and control. *Chinese Journal of Contemporary Pediatrics*, *21*(5), 441-444. (Chinese).

Kostkova, P., Brewer, H., de Lusignan, S., Fottrell, E., Goldacre, B., Hart, G., Koczan, P., Knight, P., Marsolier, C., & McKendry, R. A. (2016). Who owns the data? Open data for healthcare. *Frontiers in public health*, *4*, 7.

Lawler, E., Hedge, A., & Pavlovic-Veselinovic, S. (2011). Cognitive ergonomics, socio-technical systems, and the impact of healthcare information technologies. *International Journal of Industrial Ergonomics*, 41(4), 336-344.

Lee, J. (2020). Big data strategies for government, society and policy-making. *Journal of Asian Finance, Economics and Business*, *7(7)*, 475-487.

Li, B., & Tsui, R. (2020). How to Improve the Reuse of Clinical Data - OpenEHR and OMOP CDM. *Journal of Physics: Conference Series*. IOP Publishing, 1624(3): 032041.

Li, H., Chen, Q., Tang, B., Huang, D., Wang, X., & Liu, Z. (2016). An initial ingredient analysis of drugs approved by China food and drug administration. *China Conference on Knowledge Graph and Semantic Computing*. 650, 104-109.

Li, J., Tian, Y., Li, R., Zhou, T., Li, J., Ding, K., & Li, J. (2021). Improving prediction for medical institution with limited patient data: Leveraging hospital-specific data based on multicenter collaborative research network. *Artificial Intelligence in Medicine, 113*, 102024.

Li, L., Wei, F., Wang, F., Xu, Y., & Gong, P. (2004). Establishment and management of tumor bank applied for molecular biological researches. *Tumor*, *24*(6), 578-580. (Chinese).

Li, Q., Lan, L., Zeng, N., You, L., Yin, J., Zhou, X., & Meng, Q. (2019). A Framework for Big Data Governance to Advance RHINs: A Case Study of China. *IEEE Access*, *7*, 50330-50338.

Li, Y., Lu, C., & Liu, Y. (2020a). Medical Insurance Information Systems in China: Mixed Methods Study. *JMIR Med Inform*, *8*(9), e18780.

Li, Z., Li, N., Cimellaro, G., & Fang, D. (2020b). System dynamics modeling-based approach for assessing seismic resilience of hospitals: Methodology and a case in China. *Journal of Management in Engineering*, 36 (5): 04020050.

Ling, Y., Jin, Z., Su, M., Zhong, J., Zhao, Y., Yu, J., Wu, J., & Xiao, J. (2014). VCGDB: A dynamic genome database of the chinese population. *BMC Genomics*, *15(1)*, 1-13.

Liu, J., & Huang, X. (2009). The emergence and development of hospital informatization in China. *Journal of Intelligence*, *28*(S1), 69-70. (Chinese).
Liu, R., & Dong, C. (2012). Analysis on the current situation of scientific data resource sharing in the field of medicine and health in China. *China Basic Science*, *14*(05), 44-48. (Chinese).
Liu, Y., Cai, H., & Li, G. (2018). Research on Sharing and Application of Medical Big Data. *2018 3rd International Conference on Smart City and Systems Engineering (ICSCSE)*. IEEE, 771-774.
Luo, A., & Ping, J. (2011). Discussion on the construction of regional health information. *Medicine & Philosophy (Humanistic & Social Medicine Edition)*, *32*(07), 48-50. (Chinese).
Lv, H. (2017). Analysis of the Research Themes and Hot Trends of Medical and Health Information Disclosure. *Journal of Modern Information*, *37*(02), 112-118. (Chinese).
Lv, H. (2018). Establishment of Basic Research Framework and Synergistic Dynamic Model of Policy Formulation Level for Medical and Health Information Disclosure Policy. *Journal of Modern Information*, *38*(02), 36-40+177. (Chinese).
Lv, Z., & Qiao, L. (2020). Analysis of healthcare big data. *Future Generation Computer Systems*, *109*, 103-110.
Ma, C. (2015). Research on the Relationship Between Information Disclosure and the Quality of Health Service of Health Institution. *Chinese Hospital Management*, *35*(06), 24-26. (Chinese).
Ma, S., Yu, G., Chen, M., Cui, W., & Wei, M. (2019). Questionnaire survey on the openness and privacy protection of patients' medical and health data. *Chinese Journal of Health Informatics and Management*, *16*(02), 226-231. (Chinese)
Münch, C., Marx, E., Benz, L., Hartmann, E., & Matzner, M. (2022). Capabilities of digital servitization: Evidence from the socio-technical systems theory. *Technological Forecasting and Social Change*, 176, 121361.
Norman, D. (1988). *The psychology of everyday things*. Basic books.
Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the future—big data, machine learning, and clinical medicine. *The New England journal of medicine*, *375*(13), 1216.
Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic reviews*, *5*(1), 1-10.
Page, M., McKenzie, J., Bossuyt, P., Boutron, I., Hoffmann, T., Mulrow, C., Shamseer, L., Tetzlaff, J., Akl, A., & Brennan, S. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Syst Rev, 10 (1): 1–11. BioMed Central.
Palanisamy, V., & Thirunavukarasu, R. (2019). Implications of big data analytics in developing healthcare frameworks–A review. *Journal of King Saud University-Computer and Information Sciences*, 31 (4): 415–425.
Reason, J. (2000). Human error: models and management. *Bmj*, *320*(7237), 768-770.
Rogers, E. M. (2010). *Diffusion of innovations*. An integrated approach to communication theory and research. Routledge. 432-448 .
Shangguan, Z., Wang, M., & Sun, W. (2020). What Caused the Outbreak of COVID-19 in China: From the Perspective of Crisis Management. *International Journal of Environmental Research and Public Health*, *17(9)*, 3279.
Sharma, A., Borah, S., & Moses, A. (2021). Responses to COVID-19: The role of governance, healthcare infrastructure, and learning from past pandemics. *Journal of Business Research*, 122: 597–607.
Sittig, D., & Singh, H. (2015). A new socio-technical model for studying health information technology in complex adaptive healthcare systems. *Cognitive informatics for biomedicine*, 59-80.
Song, G., Wen, Y., Jia, Y., & Liu, H. (2019a). Research on Medical Service System Based on Big Data Technology. *2019 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*. IEEE, 302-304.
Song, Q., Zheng, Y., Huang, Y., Xu, Z., Sheng, W., & Yang, J. (2019b). Emergency Drug Procurement Planning Based on Big-Data Driven Morbidity Prediction. *IEEE Transactions on Industrial Informatics*, *15*(12), 6379-6388.

Sony, M., & Naik, S. (2020). Industry 4.0 integration with socio-technical systems theory: A systematic review and proposed theoretical model. *Technology in society*, 61, 101248.

Su, X., & Zhao, J. (2016). Connecting Information to Promote Public Health. *Frontiers of Engineering Management*, *3*(4), 384-389.

Sun, W., Shen, W., Li, X., Cao, F., Ni, Y., & Liu, H. (2013). Mining information dependency in outpatient encounters for chronic disease care. In *MEDINFO 2013*. IOS Press. 278-282.

Tian, H., Yang, W., Hu, Y., Liu, Z., Chen, L., Lei, L., Zhang, F., Cai, F., Xu, H., Liu, M., Guo, C., Chen, Y., Xiao, P., Chen, J., Ji, P., Fang, Z., Liu, F., Liu, Y., Pan, Y., . . . Ke, Y. (2020). Estimating cancer incidence based on claims data from medical insurance systems in two areas lacking cancer registries in China. *Eclinicalmedicine*, *20*, 100312.

Trist, E., & Bamforth, K. (1951). Some social and psychological consequences of the longwall method of coal-getting: An examination of the psychological situation and defences of a work group in relation to the social structure and technological content of the work system. *Human relations*, *4*(1), 3-38.

Venkatesh, V., Morris, M., Davis, G., & Davis, F. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478.

Vetrò, A., Canova, L., Torchiano, M., Minotas, C. O., Iemma, R., & Morando, F. (2016). Open data quality measurement framework: Definition and application to Open Government Data. *Government Information Quarterly*, *33*(2), 325-337.

Villanustre, F., Chala, A., Dev, R., Xu, L., Furht, B., & Khoshgoftaar, T. (2021). Modeling and tracking Covid-19 cases using Big Data analytics on HPCC system platform. *Journal of Big Data*, *8*(1), 1-24.

Wan, X., Wu, Y., & Shi, Y. (2012). Analysis of Influencing Factors in the Choice of Regional Medical Informatization Construction Mode. *Chongqing Medicine*, *41*(23), 2436-2438. (Chinese)

Wang, F., Zhu, H., Liu, X., Lu, R., Hua, J., Li, H., & Li, H. (2019a). Privacy-preserving collaborative model learning scheme for E-healthcare. *IEEE Access*, 7: 166054–166065.

Wang, X., Birch, S., Ma, H., Zhu, W., & Meng, Q. (2016). The structure and effectiveness of health systems: Exploring the impact of system integration in rural China. *International Journal of Integrated Care*, *16(3)*.

Wang, Y., Sun, L., & Hou, J. (2017). Hierarchical medical system based on big data and mobile internet: A new strategic choice in health care. *JMIR Medical Informatics*, *5*(3).

Wang, Y., Zhang, A., Zhang, P., & Wang, H. (2019b). Cloud-assisted EHR sharing with security and privacy preservation via consortium blockchain. *IEEE Access*, 7, 136704-136719.

Wesley, D., Schubel, L., Hsiao, C., Burn, S., Howe, J., Kellogg, K., Lincoln, A., Kim, B., & Ratwani, R. (2019). A socio-technical systems approach to the use of health IT for patient reported outcomes: patient and healthcare provider perspectives. *Journal of Biomedical Informatics: X*, *4*, 100048.

Wu, F., & Wang, H. (2014). A cloud platform for clinical information exchange based on web services. *Applied Mechanics and Materials*, 556: 6007-6010.

Wu, X., Ding, Y., Jiang, F., Cao, H., Mao, P., & Xie, D. (2020). Research on construction and application on health care big data open laboratory. *Experimental Technology and Management*, *37*(05), 231-233+237. (Chinese)

Wu, Y., Guo, Z., Fu, X., Wu, J., Gao, J., Zeng, Q., Fu, H., Cheng, X., & China Health Big Data project, i. (2019). The study protocol for the China Health Big Data (China Biobank) project. *Quantitative Imaging in Medicine and Surgery*, *9*(6), 1095.

Wu, Y., & Shan, S. (2021). Application of Artificial Intelligence to Social Governance Capabilities under Public Health Emergencies. *Mathematical Problems in Engineering*, *2021*, 6630483.

Xie, X., Fang, Z., Chen, L., Lu, Q., Tan, T., Ye, Z., & Pitt, M. (2022). Facilitating Patient-Centric Thinking in Hospital Facility Management: A Case of Pharmaceutical Inventory. *Buildings*, 12 (7): 888. MDPI.

Xu, B., Li, L., Hu, D., Wu, B., Ye, C., & Cai, H. (2018a). Healthcare data analysis system for regional medical union in smart city. *Journal of Management Analytics*, *5*(4), 334-349.

Xu, J., Xue, K., Li, S., Tian, H., Hong, J., Hong, P., & Yu, N. (2019). Healthchain: A blockchain-based privacy preserving scheme for large-scale health data. *IEEE Internet of Things Journal*, *6(5)*: 8770–8781.

Xu, S., & Tan, K. (2016). Data-driven inventory management in the healthcare supply chain. In *Supply Chain Management in the Big Data Era*. IGI Global, 75-91.

Xu, X., Hao, L., Zhu, J., Tang, B., Zhou, Q., Song, F., Chen, T., Zhang, S., Dong, L., Lan, L., Sang, J., Liang, F., Cao, J., Liu, F., Liu, L., Wang, F., Ma, Y., Zhang, L., Chen, M., . . . Members, B. I. G. D. C. (2018b). Database Resources of the BIG Data Center in 2018. *Nucleic Acids Research*, *46*(D1), D14-D20.

Xue, T., Fu, Q., Wang, C., & Wang, X. (2017). Research on Medical Data Sharing Model Based on Blockchain. *Acta Automatica Sinica*, *43*(09), 1555-1562. (Chinese).

Yan, L., Huang, W., Wang, L., Feng, S., Peng, Y., & Peng, J. (2021). Data-Enabled Digestive Medicine: A New Big Data Analytics Platform. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *18*(3), 922-931.

Yan, W., Palm, L., Lu, X., Nie, S., Xu, B., Zhao, Q., Tao, T., Cheng, L., Tan, L., Dong, H., & Diwan, V. K. (2013). ISS-An Electronic Syndromic Surveillance System for Infectious Disease in Rural China. *Plos One*, *8*(4).

Yang, D. (2017). The protection and disclosure of patients' personal medical information——take the "reference" clauses and personal privacy clauses of the "Regulations on Disclosure of Government Information" as the path. *Northern Legal Science*, *11*(04), 121-131. (Chinese).

Yang, X., & Cao, X. (2017). Design of cloud-based China's community care system for diabetes. *2016 International Conference on Information System and Artificial Intelligence (ISAI)*. IEEE, 228-231.

Yang, Z., Wu, Z., Tan, W., Luo, Y., Xu, Y., & Zhao, C. (2012). Creation and management of a tumor tissue bank. *West China Medical Journal*, *27*, 767-770. (Chinese).

Yao, Q., Han, X., Ma, X. K., Xue, Y. F., Chen, Y. J., & Li, J. S. (2014). Cloud-Based Hospital Information System as a Service for Grassroots Healthcare Institutions. *Journal of Medical Systems*, *38*(9), 1-7.

Yao, Z. (2006). Medical and health informationization based on sharing. *Computer Applications and Software* (10), 4-6. (Chinese).

Yin, H., Zhang, H., Liu, Y., & Wei, M. (2020). Research progress of medical information sharing at home and abroad. *Information studies:Theory & Application*, *43*(01), 177-181+162. (Chinese).

Yu, W., Chen, J., Liu, Q., Hu, Y., Gu, Z., Tian, Z., & Liu, Y. (2019). A new type of blockchain technology for the safe sharing of medical big data. *Journal of Chinese Computer Systems*, *40*(07), 1449-1454. (Chinese).

Yu, Y. (2006). The Suggestions for Medical Information Opening in China. *Chinese Hospital Management*, (06), 7-10. (Chinese).

Zeng, S., & Wu, M. (2019). Based on Public Health Service in Smart Medical Comprehensive Service Platform. 2019 *IEEE International Conference on Computation, Communication and Engineering (ICCCE)*. IEEE, 48-51.

Zhang, J., Li, Y., Gu, Y., Zhu, Y., & He, Q. (2020a). Discussion on the Basic Principles of Health and Medical Data Sharing. *Engineering Science*, *22*(04), 93-100. (Chinese).

Zhang, L., Huang, C., & Gu, H. (2006). Research on the Current Situation and Countermeasures of Information Disclosure in Medical Institutions. *Information Science*(02), 275-278. (Chinese).

Zhang, L., Wu, X., Hu, Y., Wang, X., He, Z., Xie, Y., Pan, K., Wang, N., Dong, Z., Zhang, L., & Ji, J. (2015a). Establishment of a network-based intra-hospital virtual cancer biobank. *Biopreservation and Biobanking*, *13*(1), 43-48.

Zhang, Q., Wang, L., Liu, N., Pei, D., & Jiang, Z. (2017). The research on medical information sharing between hospitals based on evolutionary game theory. *2017 13th IEEE Conference on Automation Science and Engineering (CASE)*. IEEE, 1433-1438.

Zhang, Q., Zhou, X., & Wu, Q. (2020b). Construction and Exploration of Information Interconnection in a Hospital in the Guangdong-Hong Kong-Macao Greater Bay Area. *Proceedings of the 2020 International Symposium on Artificial Intelligence in Medical Sciences*. 203-207.

Zhang, Y., Li, Q., Wang, X., & Zhou, X. (2015b). China biobanking. In *Advances in Experimental Medicine and Biology*, 864, 125-140.

Zhang, Y. S., Xia, L., Sang, J., Li, M., Liu, L., Li, M. W., Niu, G. Y., Cao, J. B., Teng, X. F., Zhou, Q., & Zhang, Z. (2018). The BIG Data Center's database resources. *Yi Chuan, 40(11)*, 1039-1043. (Chinese).

Zhao, L. (2012). Effective ways to improve the management level of residents' health records. *Lantai World*(17), 24-25. (Chinese).

Zhou, Y., Wang, L., Xu, Y., Ding, L., & Tang, Z. (2022). Intelligent Fangcang Shelter Hospital Systems for Major Public Health Emergencies: The Case of the Optics Valley Fangcang Shelter Hospital. *Journal of Management in Engineering*, 38 (1): 05021010.

Zhu, L. (2011). Analysis on the Reform of Regional Medical Information in Shanghai. *Journal of East China University of Science and Technology (Social Science Edition), 26(02)*, 99-104. (Chinese).