Population-Based Clinical Studies Using Routinely Collected Data in Hong Kong, China: A Systematic Review of Trends and Established Local Practices

Derek Wu1, Ronald Nam1,2, Keith Sai Kit Leung, BSc, MBChB1,2,3, Hamza Waraich1, Athaya Purnomo, MD4,5, Oscar Hou In Chou, MSc1,3, Francesco Perone, MD6, Shubhadarshini Pawar, MBBS, MPH1, Fatima Faraz7, Haipeng Liu, PhD3, Jiandong Zhou, PhD9, Tong Liu, MD, PhD10, Jeffrey Shi Kai Chan, MBChB, MPH1 and Gary Tse, MD, PhD, FFPH, FRCP10,11,12

1Epidemiology Research Unit, Cardiovascular Analytics Group, China-UK Collaboration, Hong Kong, China
2Hull University Teaching Hospitals NHS Trust, Hull, UK
3School of Clinical Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China
4Department of Oncology, University of Oxford, Oxford, UK
5Department of Urology, Faculty of Medicine Universitas Brawijaya - Saiful Anwar General Hospital, Malang, Indonesia
6Cardiac Rehabilitation Unit, Rehabilitation Clinic “Villa delle Magnolie”, 81020 Castel Morrone, Caserta, Italy
7Department of Medicine, Rawalpindi Medical University, Rawalpindi, Pakistan
8Research Centre for Intelligent Healthcare, Coventry University, Coventry, UK
9Nuffield Department of Medicine, University of Oxford, Oxford, UK
10Tianjin Key Laboratory of Ionic-Molecular Function of Cardiovascular Disease, Department of Cardiology, Tianjin Institute of Cardiology, Second Hospital of Tianjin Medical University, Tianjin, China
11Kent and Medway Medical School, University of Kent and Canterbury Christ Church University, Canterbury, UK
12Division of Graduate Studies, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China

Received: 19 August 2023; Revised: 23 September 2023; Accepted: 6 October 2023

Abstract

Background: Routinely collected health data are increasingly used in clinical research. No study has systematically reviewed the temporal trends in the number of publications and analyzed different aspects of local research practices and their variations in Hong Kong, China, with a specific focus on research ethics governance and approval.

Methods: PubMed was systematically searched from its inception to March 28, 2023, for studies using routinely collected healthcare data from Hong Kong.

Results: A total of 454 studies were included. Between 2000 and 2009, 32 studies were identified. The number of publications increased from 5 to 120 between 2010 and 2022. Of the investigator-led studies using the Hospital Authority (HA)’s cross-cluster data (n = 393), 327 (83.2%) reported receiving ethics approval from a single cluster/university-based REC, whereas 50 studies (12.7%) did not report approval from a REC. For use of the HA Data Collaboration Lab, approval by a single hospital-based or University-based REC is accepted. Repeated submission of identical ethics applications to different RECs is estimated to cost HK$4.2 million yearly.

Conclusions: Most studies reported gaining approval from a single cluster REC before retrieval of cross-cluster HA data. Substantial cost savings would result if repeated review of identical ethics applications were not required.

Keywords: population-based; cross-cluster; territory-wide; research governance; ethics
Introduction

With the advancement of information technologies, the use of routinely collected electronic health records (EHRs) for clinical and epidemiological research has markedly increased [1]. EHRs contain individual patient information that is collected longitudinally when patients access any affiliated healthcare facilities [2]. A wide range of data including basic demographics; investigations conducted; diagnoses; medications and procedures; and administrative data, including length of stay and hours waited in the emergency department, can be obtained from the data warehouse. In Hong Kong, the Hospital Authority (HA), which is divided into seven administrative clusters, provides care for a population of 7 million residents via its 43 public hospitals, 49 specialist outpatient clinics and 74 general outpatient clinics. HA has maintained a territory-wide clinical database, the Clinical Data Analysis and Reporting System (CDARS), to capture cross-cluster data at the individual patient level since January 1, 1995 [3]. CDARS data have been extensively used to conduct cross-cluster studies [4]. After a governmental call to increase the accessibility of healthcare data for biomedical research, the HA Data Collaboration Lab (HADCL) was established in December 2019, thus enabling investigators not traditionally affiliated with hospitals or local medical schools to use the data for self-initiated research projects. In addition, the Department of Health manages certain primary care clinics, as well as clinics providing services for tuberculosis, chest conditions, social hygiene (sexually transmitted diseases), and dermatology. Both the Department of Health and the HA are essential components of Hong Kong’s healthcare system, and each is dedicated to delivering healthcare services to the public. These clinics are described separately because they are under the Department of Health’s jurisdiction rather than the HA’s. Although both entities share the common goal of public health, they have distinct roles and responsibilities.

To date, review articles on research outputs generated by using EHR databases are limited, and only bibliometric analysis has been performed on the UK’s Hospital Episode Statistics database, Clinical Practice Research Datalink, The Health Improvement Network and QResearch [5, 6]. In the ever-expanding field of scholarly publications, a comprehensive examination conducted by Hemkens et al. has shed light on the quality of studies using routinely collected health data [7]. Unexpectedly, despite the surge in publications, the findings of Hemkens et al. have revealed a disheartening trend: most of these studies lack satisfactory reporting standards, i.e., adherence to established guidelines for reporting research, such as the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines. These guidelines provide standards for reporting observational studies, thus ensuring that readers can accurately interpret and replicate the study’s methods and findings. The insufficient reporting identified in these studies suggests a need for improved adherence to these or similar guidelines, to enhance the quality and transparency of health data research.

However, no study has systematically reviewed publications from Hong Kong and local research practices. The aim of this systematic review is to describe temporal trends in the number of publications and analyze various aspects of local research practices and their variations, with a specific focus on research ethics governance and approval.

Methods

Search Strategy, and Inclusion and Exclusion Criteria

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. PubMed was systematically searched from its inception to March 28, 2023, for studies using routinely collected healthcare data in Hong Kong, on the basis of the following search terms: Hong Kong and (“CDARS” OR “territory-wide” OR...
“population-based”). The inclusion criteria were i) studies using routinely collected health data; ii) studies using patient-level data or aggregate patient-level data from Hong Kong; and iii) study cohorts including patients from more than two centers. The exclusion criteria were any of the following: i) studies using non-routinely collected data (such as surveys, experimental, or prospective registries); ii) non-population-based studies (i.e., those including patients from one or two centers); iii) studies not reporting original findings; or iv) studies using both non-routinely collected data and routinely collected data.

The retrieved studies were screened independently by two investigators. Cases of disagreement were resolved through discussion with a third reviewer to reach a consensus. We specifically excluded prospective registry studies, primarily because the data from these studies would not have been automatically indexed in the health record system but instead stored separately by the researchers. Because the EHR system is updated daily, researchers could theoretically conduct a prospective study by using routinely collected data. Such studies would have been included in this systematic review.

**Data Extraction**

The following data were extracted from each study: name of the first author, year of publication (according to creation date in PubMed), name of the last author, cohort definition, cohort years, cohort size, data source, study type, ethical approval, methods, outcomes and follow-up period. Descriptive statistics was used to summarize the data extracted.

**Counterfactual Analysis of Costs Saved by Not Requiring Repeated Reviews of Identical Ethics Applications**

For a hypothetical healthcare system with \( n \) clusters, assuming that the governance requires approval from all \( n \) RECs for all investigator-initiated studies using cross-cluster data (i.e., data from all \( n \) clusters), the additional person-time required for the review of identical proposals would be \((n-1) \times m \times t \times s \times r\), where \( r \) is the remuneration per hour. Counterfactual analysis was applied to estimate the costs saved by not requiring identical ethics applications to be repeated for the remaining \( n-1 \) clusters.

**Data Availability**

All data arising from this systematic review were extracted from published articles and are detailed in the Supplementary Appendix on the preprint server SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4428255.

**Results**

**Search Results and Study Selection**

Our search strategy returned 1757 entries. After screening of each entry by two independent investigators, 451 studies were selected for inclusion. An additional Google search yielded three additional studies. Therefore, a total of 454 cross-cluster studies were included (Figure 1). Before the year 2000, seven studies, published between 1989 and 1999, were identified. Between 2000 and 2009, 28 studies were identified (2002: \( n = 4 \); 2003: \( n = 5 \); 2004: \( n = 2 \); 2005: \( n = 5 \); 2006: \( n = 2 \); 2007: \( n = 6 \); 2008: \( n = 7 \); 2009: \( n = 1 \)). The number of publications increased from 5 to 120 between 2010 and 2022 (Figure 2A). A summary of the different cross-cluster studies by year of publication is detailed in Table 1. The details of each study, including name of the first author, year of publication, name of the last author, definition of study cohorts, cohort recruitment years, cohort size, data source, study type, REC ethical approval, methods, outcomes and outcome period are shown in Supplementary Table 1.

**Trends in the Number of Publications, Study Designs and Statistical Methods**

The greatest number of patients included was 1487 before 2000; this number increased to 56,167 between 2000 and 2009 and to 7 million in 2018. To define the study cohorts, most studies (\( n = 283 \)) used International Classification of Diseases (Ninth Edition) disease coding, followed by medications...
(n = 64), settings (n = 61, e.g., primary care clinics, accident and emergency departments), exposures (n = 17, e.g., risk factors or vaccines), age (n = 15) and procedures (n = 14). Cohort studies were the most common study design (n = 423) and were followed by case series, case-control or self-controlled case series (n = 22) and cross-sectional studies (n = 1). Four studies reported using two different designs. Cox regression was the most common statistical method (n = 218), and was followed by logistic regression (n = 63), Poisson regression (n = 41), competing risk regression (n = 15), generalized linear/additive models (n = 11), autoregressive integrated moving average (n = 3) and recently machine learning (n = 2). The earliest use of propensity scores to decrease confounding effects was reported in 2014; since then, propensity scores have been calculated for matching, weighting and adjustment. Direct matching has not been reported.

**Local Practices for Ethics Approval for Cross-Cluster Studies**

Seven cluster-based RECs from the HA and the DH accepted applications from all investigators affiliated with staff working in the public hospitals, as well as staff and students affiliated with the local medical schools. Of the 393 cross-studies using HA data, 327 (83.2%) reported receiving ethics approval from a single REC, of which seven studies
did not report the name of the committee, and 50 (12.7%) did not report any information regarding ethics approval (Figure 2B). Moreover, 38 studies (8.4%) obtained approval from the DH REC. For use of HADCL, approval by a single hospital-based or university-based REC is accepted [460]. Three studies reported using HADCL to conduct research studies. All studies associated with COVID-19 vaccines (n = 14) reported approval from both the HA Head Office’s Central REC and DH REC. The data were collected as part of the pandemic measures implemented by the government.

Counterfactual Analysis of Costs Saved by not Requiring Repeated Submission of Identical Applications to Different RECs Within the Same Healthcare System

For a hypothetical healthcare system with \( n = 7 \) clusters, if the governance requires approval from all \( n \) RECs for all investigator-initiated studies using cross-cluster data (i.e., data from all \( n \) clusters), then the additional costs required for reviewing identical ethics applications by \( n-1 \) RECs can be calculated as follows. Assuming that two members (\( m = 2 \)) review each application, and each reviewer takes 0.5 hours to review and comment on the proposal, the person-time required would be \( 6 \times 2 \times 0.5 = 6 \) person-hours. To estimate the number of investigators interested in submitting ethics applications, we assumed two medical schools, each of which has 20 departments comprising ten faculty members each, who will submit ethics applications for themselves and on behalf of their research team members or students; thus, the number of submissions from all projects initiated by the medical schools would be \( 2 \times 20 \times 10 = 400 \).

We also assumed that, within a hypothetical hospital system with 50 hospitals or clinics and two submissions per site, the number of applications from hospital staff would be \( 50 \times 2 = 100 \). Altogether, with a total of 500 faculty members or hospital staff members, \( f \), each submitting \( s \) applications, the total cost would be \( (n-1) \times m \times t \times f \times s \times r \), where \( r \) is the remuneration per hour. \( r \) is estimated from an average salary of a newly qualified specialist of HK$140,000 per month, working 50 hours per week, which is HK$700 per hour. \( 6 \times 2 \times 0.5 \times 500 \times 2 \times 700 = HK$4,200,000 \) (scenario 1). Thus, we can make a counterfactual inference that approximately HK$4.2 million would be saved through the implementation of a single centralized system to review all studies requesting the use of cross-cluster data. The various hypothetical scenarios with varying values of \( m, t, f, s \) and \( r \) are shown in Table 2. Decreasing the number of reviewers from 2 to 1 would halve the costs incurred to HK$2.1 million (scenario 2), whereas increasing the review time from 30 minutes to 36 minutes would increase the costs to HK$5.0 million (scenario 3). If each investigator were to increase the number of applications per year from 2 to 3, the costs would increase to HK$6.3 million (scenario 4). If the invited reviewer were more senior, as reflected by an increase in the hourly salary from HK$700 to HK$750, the costs would increase to HK$4.5 million (scenario 5). If \( f \) were to increase from 500 to 600, the costs for altering \( m, t, s \) and \( r \) would be HK$5.0 (scenario 6), HK$2.5 (scenario 7),
HK$6.0 (scenario 8), HK$7.6 (scenario 9) and HK$5.4 (scenario 10) million, respectively.

**Discussion**

The main findings of this systematic review of 454 cross-cluster studies were as follows. i) The first population-based study was published in 1989, and the number of publications increased from 5 to 120 between 2010 and 2022. ii) Of the 393 investigator-initiated studies using the HA’s cross-cluster data, 327 (83.2%) reported receiving ethics approval from a single cluster/university-based REC, whereas 50 studies (12.7%) did not report obtaining any approval from an REC. For use of HADCL, approval by a single hospital-based or university-based REC is accepted. iii) The costs wasted from requiring repeated submission of identical ethics applications, which is equivalent to the costs saved by not requiring multiple submission of identical ethics applications to different RECs, would be approximately HK$4.2 million per year.

With recent advances in the development of centralized systems for storing health data, clinical or epidemiological studies using such routinely collected data are increasingly being performed. Locally

---

**Table 1** Summary of Cross-Cluster Studies by Year of Publication.

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of publications</th>
<th>Number of investigator-led studies using HA data (HA studies)</th>
<th>Number of HA studies not reporting any approval from a HA cluster REC (%)</th>
<th>Number of HA studies with one approval from one HA cluster REC (%)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-2000</td>
<td>7</td>
<td>6</td>
<td>6 (100%)</td>
<td>0 (0%)</td>
<td>[8–14]</td>
</tr>
<tr>
<td>2000</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>4</td>
<td>1</td>
<td>1 (100%)</td>
<td>0 (0%)</td>
<td>[15–18]</td>
</tr>
<tr>
<td>2003</td>
<td>5</td>
<td>3</td>
<td>3 (100%)</td>
<td>0 (0%)</td>
<td>[19–22]</td>
</tr>
<tr>
<td>2004</td>
<td>2</td>
<td>2</td>
<td>2 (100%)</td>
<td>0 (0%)</td>
<td>[23, 24]</td>
</tr>
<tr>
<td>2005</td>
<td>5</td>
<td>4</td>
<td>3 (75%)</td>
<td>1 (25%)</td>
<td>[25–29]</td>
</tr>
<tr>
<td>2006</td>
<td>2</td>
<td>1</td>
<td>1 (100%)</td>
<td>0 (0%)</td>
<td>[30, 31]</td>
</tr>
<tr>
<td>2007</td>
<td>6</td>
<td>4</td>
<td>4 (100%)</td>
<td>0 (0%)</td>
<td>[32–37]</td>
</tr>
<tr>
<td>2008</td>
<td>7</td>
<td>3</td>
<td>2 (67%)</td>
<td>1 (33%)</td>
<td>[38–44]</td>
</tr>
<tr>
<td>2009</td>
<td>1</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>[45]</td>
</tr>
<tr>
<td>2010</td>
<td>5</td>
<td>2</td>
<td>Not accessible</td>
<td>1 (50%)</td>
<td>[46–50]</td>
</tr>
<tr>
<td>2011</td>
<td>5</td>
<td>4</td>
<td>1 (25%)</td>
<td>3 (75%)</td>
<td>[51–55]</td>
</tr>
<tr>
<td>2012</td>
<td>7</td>
<td>5</td>
<td>3 (60%)</td>
<td>2 (40%)</td>
<td>[56–62]</td>
</tr>
<tr>
<td>2013</td>
<td>10</td>
<td>4</td>
<td>3 (60%)</td>
<td>0 (0%)</td>
<td>[63–72]</td>
</tr>
<tr>
<td>2014</td>
<td>12</td>
<td>8</td>
<td>4 (50%)</td>
<td>4 (50%)</td>
<td>[73–84]</td>
</tr>
<tr>
<td>2015</td>
<td>6</td>
<td>5</td>
<td>2 (40%)</td>
<td>2 (50%)</td>
<td>[85–90]</td>
</tr>
<tr>
<td>2016</td>
<td>13</td>
<td>13</td>
<td>3 (23%)</td>
<td>7 (54%)</td>
<td>[91–103]</td>
</tr>
<tr>
<td>2017</td>
<td>20</td>
<td>19</td>
<td>0 (0%)</td>
<td>19 (100%)</td>
<td>[104–123]</td>
</tr>
<tr>
<td>2018</td>
<td>18</td>
<td>18</td>
<td>3 (17%)</td>
<td>14 (78%)</td>
<td>[124–141]</td>
</tr>
<tr>
<td>2019</td>
<td>27</td>
<td>26</td>
<td>3 (12%)</td>
<td>22 (85%)</td>
<td>[142–168]</td>
</tr>
<tr>
<td>2020</td>
<td>64</td>
<td>64</td>
<td>0 (0%)</td>
<td>62 (97%)</td>
<td>[169–232]</td>
</tr>
<tr>
<td>2021</td>
<td>80</td>
<td>78</td>
<td>1 (1%)</td>
<td>74 (95%)</td>
<td>[233–312]</td>
</tr>
<tr>
<td>2022</td>
<td>120</td>
<td>98</td>
<td>3 (3%)</td>
<td>93 (95%)</td>
<td>[313–431]</td>
</tr>
<tr>
<td>2023 (to 28 March)</td>
<td>28</td>
<td>25</td>
<td>2 (1%)</td>
<td>22 (88%)</td>
<td>[432–459]</td>
</tr>
<tr>
<td>Total</td>
<td>454</td>
<td>393</td>
<td>50 (13%)</td>
<td>327 (83%)</td>
<td>-</td>
</tr>
</tbody>
</table>

1Studies from a commissioned project on COVID-19 vaccines were excluded.
in Hong Kong, a likely reason for the remarkable and rapid increase in the number of big data studies using routinely collected health data between 2015 and 2020 (n = 6 to n = 64, >10 fold increase) was the return of a pharmacy professor from the UK to Hong Kong [73], whose team started using CDARS for epidemiological research. The publication of articles caught the attention of other local researchers and physicians, who followed suit. This story was recently reported in Lancet Psychiatry; interested readers are directed there for an insightful discussion on the profile of the “father of the health-care big data research in Hong Kong” [461].

A survey study of investigators from the Asia Pacific region has identified ten databases, of which four are claims-based, four are EHR systems, and two are registries [3]. However, despite the increasing number of publications, a recent study assessing the quality of studies using routinely collected health data has identified that most studies have insufficient reporting [7]. Another survey study on medical records and EHRs has revealed the gaps in the secondary use of data for research purposes, but has not examined the practices for ethics approval and data governance [462]. Gradual shifts in acceptable research practices have occurred toward a model in which patient consent requirements are waived when retrospective data are used and patients are not contacted [463–465]. Although a comprehensive system for research and data governance is required to safeguard ethical and scientific integrity appropriately [466], the system should not be excessively bureaucratic and should not obstruct researchers from conducting meaningful research. In other jurisdictions, such as Australia, researchers have voiced frustration because of the inconsistencies in the application and decision processes, as well as the long durations required for ethics and governance approvals to be granted across different jurisdictions in Australia [467], thus impeding non-interventional clinical research [468]. The challenges appear to not have been resolved entirely to date [469]. Our study found marked variations in the practices and/or reporting on the ethics approval from the different studies published by Hong Kong investigators. Locally, the hospital network is divided into seven clusters for administrative reasons. Informal discussions among local investigators have occurred regarding the requirements to obtain ethics approval.
from all cluster RECs involving cross-cluster data. The rules for prospective studies are clear, in that when active recruitment or interventions are proposed at various hospital sites, ethics (and administrative) approval should be required at each site, given the resource implications involved. However, in the case of observational studies involving retrospective review of data, without a need to involve colleagues from other sites, the guidelines have been vague. Approval by a single hospital-based or university-based REC is accepted by HA in order to use cross-cluster data from HADCL. However, this was not specifically mentioned for direct data access from CDARS. However, given that HADCL derives all its data from CDARS, the rules for HADCL might logically be inferred to be applicable to studies using CDARS. Most investigators (83.2%) obtained ethics approval from only a single cluster’s REC before using cross-cluster HA data. Thus, most investigators appear to have implicit understanding that approval is required only from a single REC.

The effects of requiring repeated reviews of identical applications can be estimated through the following thought experiments. If a team of investigators conducts a purely retrospective observational study of health records data from different hospitals, and the health system requires them to apply for ethics approval multiple times to RECs of different clusters \((n)\) of a hypothetical healthcare system, this practice would lead to five major problems. First, it would lead to variations in the practices by different RECs within its jurisdiction. Second, confusion could result from different RECs reaching different conclusions. Third, potential breaches by investigators could occur inadvertently if the practices are not aligned, for example, if one REC believes that its jurisdiction extends beyond other clusters (for reviewing cross-cluster data), whereas another does not. Fourth, the process would be highly repetitive, thereby leading to substantial wasting of valuable resources and inappropriate use of public funds. In our hypothetical scenario 1 discussed in the Results section, counterfactual analysis estimated that HK$4.2 million could be saved by not requiring repeated reviews of identical proposals by \(n-1\) clusters. Fifth, inappropriate attribution of authorship might result if RECs require a local principal investigator at each site to satisfy requirements, and researchers are invited as authors for administrative reasons rather than to provide academic input, contrary to the accepted practices described by COPE [470].

The major limitation of our study is that, although we used a systematic approach and adhered to the PRISMA guidelines as much as possible, we fully recognize that some areas of inclusion or exclusion might be subjective. For example, if some data were originally not routinely collected, but policymakers or administrators somehow decide to include such data fields in the system or to code them, then the nature of the data could be changed. If and when these data become routinely collected or routinely coded, then they can be accessed and used for big data research, as exemplified by data from the Multidisciplinary Risk Assessment and Management Programme-Diabetes Mellitus (RAMP-DM), which originally started as a quality improvement project but was subsequently incorporated into the EHR system [313].

**Implications for the Greater Bay Area and Mainland China**

This review of research activities in Hong Kong provides proof of the concept that making health records availability to bona fide researchers can provide opportunities for Chinese researchers to increase visibility internationally by publishing high quality studies. Our team’s original intention was not to specifically examine the REC approval and analyze potential cost savings. Instead, our aim was to examine the evolving trends in the publications over the past two decades to better plan our research strategy in the next 5 years. We inadvertently analyzed aspects of ethics approval because, after examining the publications returned from our searches, we observed substantial variations in practice regarding the use of the data. Given these variations, the research findings from our systematic review have clear implications and may be of interest to local health administrators and policymakers. The implications for the Greater Bay Area are that, given the drive to nurture talents, better linkage across hospitals and accessibility of data within this region may improve patient care and substantially increase the influence of biomedical research. The implications for mainland China have been discussed at length previously, and readers are directed to an excellent scoping review, which provides a helpful detailed summary of the different databases available [471].
Future Directions

The aim of this study was to provide a broad overview of the trends and main details of published studies. Future studies should conduct detailed analyses on the quality of the studies as well as in-depth bibliometric analysis, and assess how both aspects have varied during the past two decades.

Conclusions

A substantial increase in research output using routinely collected health data from Hong Kong was observed. Most studies reported receiving approval from a single cluster REC before retrieval of cross-cluster HA data. According to counterfactual analysis, if a hypothetical healthcare system with $n$ clusters requires its investigators to submit identical ethics application to $n-1$ cluster RECs, then an estimated HK$4.2 million could be saved if such a system is not implemented.

Competing interests

Tong Liu and Gary Tse are board members of CVIA. Neither Tong Liu nor Gary Tse is involved in the peer review or decision-making process of the manuscript. The other authors have no competing interests to disclose.

Author contributions

DW, RL: study conception, screening, data extraction, statistical analysis, preparation of figures, drafting of the manuscript and critical revision of the manuscript. KSKL, HW, AP, OHIC, FP, SP, FF, HL, JZ, TL, JSKC: screening, data extraction, drafting of the manuscript and critical revision of the manuscript.

Funding

No funding was received for this article.

Supplementary Material

Supplementary material for this paper is available from the following link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4428255

REFERENCES


169. Liu APY, Soh SY, Cheng FWC, Ke C, Lau E, Shah BR, Stukel Chung RY, Tsoi KKF, Kyaw Li X, Blais JE, Wong ICK, Tam Sung JJY, Chiu HM, Jung KW, Wong CKH, Tong T, Cheng GHL, 166. 165. 164. 163. 162. 161. 160. 159. 158. 157. 156. 155. 154. 153. 152. 151. 150. 149. 148. 147. 146. 145. 144. 143. 142. 141. 140. 139. 138. 137. 136. 135. 134. 133. 132. 131. 130. 129. 128. 127. 126. 125. 124. 123. 122. 121. 120. 119. 118. 117. 116. 115. 114. 113. 112. 111. 110. 109. 108. 107. 106. 105. 104. 103. 102. 101. 100. 99. 98. 97. 96. 95. 94. 93. 92. 91. 90. 89. 88. 87. 86. 85. 84. 83. 82. 81. 80. 79. 78. 77. 76. 75. 74. 73. 72. 71. 70. 69. 68. 67. 66. 65. 64. 63. 62. 61. 60. 59. 58. 57. 56. 55. 54. 53. 52. 51. 50. 49. 48. 47. 46. 45. 44. 43. 42. 41. 40. 39. 38. 37. 36. 35. 34. 33. 32. 31. 30. 29. 28. 27. 26. 25. 24. 23. 22. 21. 20. 19. 18. 17. 16. 15. 14. 13. 12. 11. 10. 9. 8. 7. 6. 5. 4. 3. 2. 1. 0. D. Wu et al., Population-Based Clinical Studies Using Routinely Collected Data in Hong Kong, China


207. Wong CKH, Wu T, Wong SKH, Law BTT, Grieve E, Ng EKW, et al. Effects of bariatric surgery on kidney diseases, cardiovascular...


235. Ng KS, Vardhanabhuti V. Chest-related imaging investigations during multiple waves of covid-19 infection in Hong Kong. Front Med (Lausanne) 2021;8:704515.


248. Chan KF, Ma TF, Ip MS, Ho PL. Invasive pneumococcal disease, pneumococcal pneumonia and all-cause pneumonia in Hong Kong during the COVID-19 pandemic compared with the preceding 5 years: a retrospective observational study. BMJ Open 2021;11(10):e055575.


study in Hong Kong. BMJ Open 2021;11(6):e045868.


343. Wan EYF, Chui CSL, Mok AHY, Xu W, Yan VKC, Lai FTT, et al. mRNA (BNT162b2) and inactivated (CoronaVac) COVID-19


388. Wan EYF, Chui CSL, Ng VWS, Wang Y, Yan VKC, Lam ICH, et al. Messenger RNA coronavirus


396. Li X, Gao L, Tong X, Chan VKY, Chui CSL, Lai FTT, et al. Autoimmune conditions following mRNA (BNT162b2) and inactivated (CoronaVac) COVID-19 vaccination: a descriptive cohort study among 1.1 million vaccinated people in Hong Kong. J Autoimmun 2022;130:102830.


404. Au PCM, Tan KCB, Cheung BMY, Wong ICK, Li HL, Cheung CL. Association Between SGLT2 Inhibitors vs DPP4 Inhibitors and Renal Outcomes Among Patients With Type 2 Diabetes. J Clin Endocrinol Metab 2022;107(7):e2962–e70.


408. Lai FTW, Chua GT, Chan EWW, Huang L, Kwan MYW, Ma T, et al. Adverse events of special interest following the use of BNT162b2 in adolescents: a


455. Chan GCK, Wong CK, So BYF, Ng JKC, Lui GCY, Szeto CC, et al. Epidemiology and outcomes


