We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

5,800
Open access books available

142,000
International authors and editors

180M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Computer-Aided Pharmacoepidemiology in Drug Use and Safety: Examining the Intersection between Data Science and Medicines Research

Ibrahim Chikowe and Elias Peter Mwakilama

Abstract

Pharmacoepidemiology is a relatively new area of study that focuses on research aimed at producing data about drugs' usage and safety in well-defined populations. Its significant impact on patient safety has translated into improving health care systems worldwide, where it has been widely adopted. This field has developed to an extent that policy and guidelines makers have started using its evidence alongside that produced from randomised controlled clinical trials. Although this significant improvement has been partly attributed to the adoption of statistics and computer-aided models into the way pharmacoepidemiology studies are designed and conducted, certain gaps still exist. This chapter reports some of the significant developments made, along with the gaps observed so far, in the adoption of statistics and computing into pharmacoepidemiology research. The goal is to highlight efforts that have led to the new pharmacoepidemiology developments, while examining the intersection between data science and pharmacology through research narrative reviews of computer-aided pharmacology. The chapter shows the significant number of initiatives that have been applied/adopted to improve pharmacoepidemiology research. Nonetheless, further developments in integrating pharmacoepidemiology with computers and statistics are needed in order to enhance the research agenda.

Keywords: Database, data science, computer-aided, pharmacovigilance, safety, adverse drug reaction

1. Introduction

Pharmacoepidemiology is a research field that applies epidemiological concepts into clinical pharmacology. It is important in the provision of an evidence base for pharmacotherapy, due to the abundance of digital data that is mostly scanty [1, 2]. Pharmacoepidemiology studies aim to quantify patterns of drug use, as well as adverse drug events, and include prescribing, use appropriateness, adherence to treatment regimen and persistence patterns, along with factors that assist in predicting medication use. In addition, pharmacoepidemiology studies involve drug
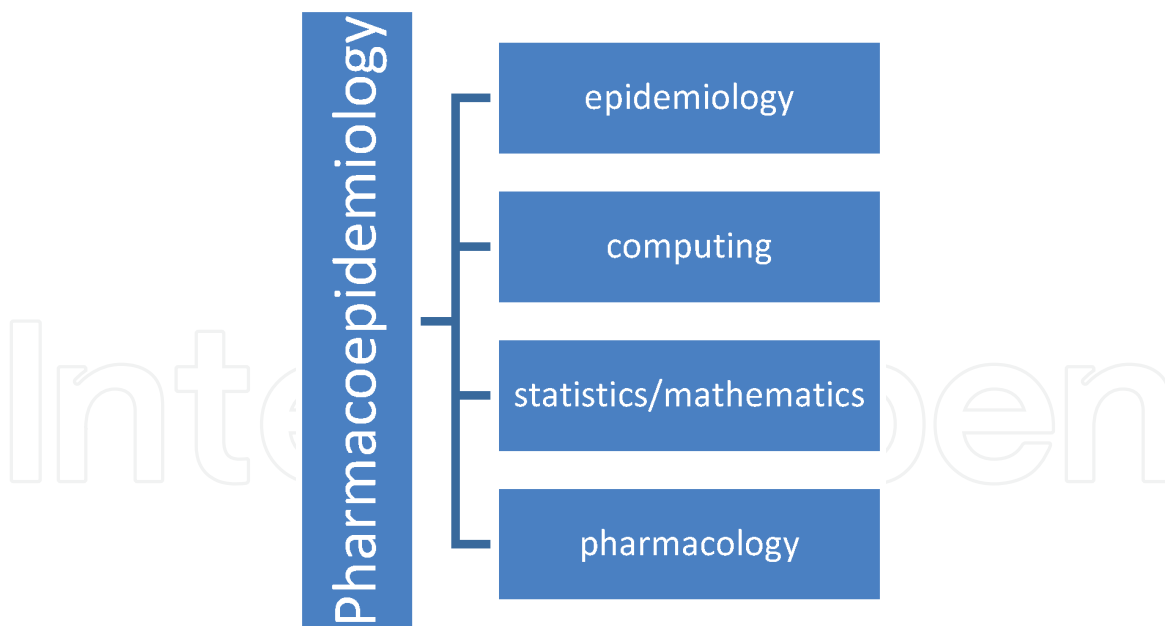
Figure 1.
Main contributors of Pharmacoepidemiology.

safety studies in large populations that focus on common and uncommon, as well as predictable and unpredictable, adverse drug reactions (ADRs) [3]. In this case, all the studies rely on meta-data sources, and include primary data, comprising national data sources and surveys or registries; and secondary data comprising administrative databases, claims databases, as well as primary care electronic health and medical records. **Figure 1** presents the general description of pharmacoepidemiology [4] being a multidisciplinary type of research field which intersects mathematical disciplines with pharmacology.

Recently, it has been established that clinical trial-oriented studies alone are mostly found to be insufficient to provide conclusive data about the drug's safety and occurrence of adverse effects in larger populations, especially the occurrence of idiosyncratic adverse events and other rare events. This is attributed to both the smaller populations and shorter time periods in which the medicines are tested. Additionally, the effectiveness of the medicines is not fully determined by the time the medicines are launched into the market. Post-marketing surveillance, with the help of either statistical or computing models on longitudinal data, becomes a critical tool for solving these challenges. Furthermore, it is important to highlight that adverse drug events and drug's efficacy can vary between clinical trial protocols and health care delivery systems [5–7]. Therefore, pharmacoepidemiology research data has found its way into many aspects of health care systems, such as policy making, drug utilisation and safety decision making, clinical trial design or validation, as well as guidance for the improvement of medical prescription by physicians. Additionally, it is also essential for research and project implementation, methodology development, vaccine and medical devices safety assessment, as well as for minimisation of medication errors and drug-induced toxicities [8].

2. Challenges and opportunities linked to pharmacoepidemiology

Pharmacoepidemiology research provides very important data for the benefit of patients' safety and care since the data generated is more informative and reliable when the study is well designed. Pharmacoepidemiology research offers many advantages, including the use of large patient samples and inclusion of

subpopulations that are under research in uncontrolled conditions [1]. It also describes and estimates the risks and other drug safety or efficacy phenomena in practice [9]. Pharmacoepidemiology approaches make the studies cheaper and faster, when compared to the randomised controlled trials initially performed prior to marketing or after marketing, thus enabling the researchers to assess generic medications, as well as medications after a long period of use. The methods used in pharmacoepidemiology research can also be adapted for their use in pharmacovigilance to assist in unearthing unknown side effects or ADRs, together with the discovery of new drug usages [10].

However, pharmacoepidemiology research also has its own drawbacks, such as contamination of the data with confounding factors and many sources of bias (information bias, selection bias), due to the non-randomised nature of treatment selection, being harder to draw conclusions [1, 11]. In addition, although inclusion of statistical models into pharmacoepidemiology has been already seen, little is known about integrating pharmacology with community behaviour models, such as social networks. Nonetheless, different scholars have suggested several ways of improving pharmacoepidemiology research, including the use of active comparison groups and within-individual designs, as well as propensity scoring [12]. Additionally, pharmacoepidemiology studies have also been improved by triangulation of multiple analytical and data collection approaches, aiming to enhance the confidence in inferred causal relationships [13]. The developments made in the use of databases, computer and statistical models, and big data have led to enormous improvements in the robustness of pharmacoepidemiology studies and the production of reliable data that is being considered as good evidence for inclusion in guidelines, alongside data generated from randomised controlled trials [14].

Having shown that pharmacoepidemiology research is now producing data that is important for health care guidelines and policy development, it is essential that researchers can collaborate with guideline writers to ensure that they frame their questions to get useful answers. On the other hand, pharmacoepidemiology researchers should design their studies in such a way that guideline writers are provided with concrete answers, thus reducing the uncertainty in the evidence base. Additionally, since pharmacoepidemiology depends on statistical and data sciences, there is a need for further development of techniques in these fields to improve the application of pharmacoepidemiology. It is also important to enhance public engagement and capacity building (data resources and researcher base) to take full advantage of future opportunities [1].

3. Computational and statistical models in pharmacoepidemiology

The advent and development of computers has led to the development of databases that have become essential in pharmacoepidemiology. Several Electronic Health Records (EHRs) systems have been developed to keep longitudinal digital records of patient health information that are generated after a series of visits in a hospital setting [15]. EHRs contain patient data related to diseases, medicines and laboratory results, if any, and enable the provision of patient centred treatment by the health care providers [16, 17]. When these databases are linked or nationalised, it prevents patients repeatedly describing their medical histories, in case of treatment transfers. In addition, such data can be accessed by policy makers or researchers [18]. The use of computerised databases has led to a significant reduction in adverse events and prescription errors [19, 20], shorter hospital stays and lower mortality [21], along with better patient tracking, information exchange, efficient handling of information, and real-time data provision [16, 22]. Large

pharmacoepidemiology data bases facilitate research, but they require well trained personnel to produce and handle big data [17, 23]. The use of electronic data has led to a significant reduction in the manual effort of data collection, easy incorporation of regional data into a study, minimal need for recalls, and removal of interviewer bias [24].

3.1 Progress and limitations

3.1.1 Usage of computational and statistical models

So far, a very close link between pharmacology and computational and statistical models has been established (**Figure 1**). In his work, Bentley [25] provides a well organised chapter describing the key statistical models used in the field of pharmacoepidemiology, both at descriptive and inferential analysis levels. Description uses measures of central tendency (e.g. mean), dispersion (e.g. variance), range (e.g. range, maximum and minimum), expressed in tables (e.g. cross-tabulations) and charts but inference may use regression models (e.g. linear, logistic, and Cox). These statistical techniques and descriptions aid in understanding data on usage and effects of drug administration at community level although it is also important to have a good knowledge of the potential errors involved in the design and analysis of pharmacoepidemiology studies [26].

Statistics play a major role in managing the quantifiable errors present in pharmacoepidemiology data analysis and interpretation [27]. Despite a growing interest in applying epidemiology statistical methods in pharmaceutical studies, a proper usage of the statistical techniques in research studies is often still lacking. For example, Suissa [26] states that pharmacoepidemiology observational research studies are hugely affected by information bias (when selecting variables of interest for the study), selection bias (during inclusion and exclusion of subjects), and confounding bias (due to imbalances in covariates). To circumvent these problems, both randomised controlled trials and cohort and case control studies, also used in epidemiological studies [28], have therefore been recommended by several researchers in pharmacoepidemiology [29].

Accordingly, in order to appraise the significance of epidemiological data and the design of studies on drug risk and safety, we reviewed a couple of research studies that have been conducted in developing countries, including in Malawi. We tried to focus on citing the key statistical and computational methods used in such research studies. To achieve this, we have used a similar approach to the one described by Sequi et al. [30] who presented a review of studies to underscore the processes of analysing and reporting data related to paediatric drug utilisation. Out of the 22 studies, the majority (91%) reported at least one descriptive measure, with the mean being the most common one (82%, 18/22), followed by the standard deviation (23%, 5/22). The chi-square test was observed in 12 studies, while graphical analysis was reported in 14 papers. However, only 16 papers reported the number of drug prescriptions and/or packages, while 10 reported the prevalence of the drug prescription. Consequently, the authors observed that only a few of the studies reviewed applied statistical methods and reported data in a satisfactory manner [27].

In a review paper which has set a position on current usage of statistical models in pharmacoepidemiology, Rosli and others [31] systematically reviewed published studies on drug utilisation in hospitalised neonates in Europe, the United States, India, Brazil, and Iran. The findings were not far from those reported by [30] such that a majority (70%) used descriptive statistics to analyse pharmacoepidemiology

data. Nonetheless, some quite remarkable variations were observed regarding to the study design and methodology, sources of data, and sampling process among the selected studies. Of the included studies, 45% were based on cross-sectional or retrospective designs, 40% were prospective, and the remainder (15%) were point prevalence surveys.

Likewise, a 2020 review of 84 drug utilisation studies among neonates by Al-Turkait et al. [32] has shown that median, ranges and mean are frequently reported statistical parameters used for describing pharmacoepidemiology data, and that the style of reporting is mostly descriptive. However, in general public health, Hayat et al. [33] found a variety of statistical methods that were identified in the 216 papers reviewed, whereby 81.9% used an observational study design. 93.1% substantive analysis, 95% used descriptive statistics (tabular or graphical) while statistical inference (t-test, Chi-square, correlation with confidence intervals and p-values) was used in 76%. Logistic regression models were frequently used (38.4%), followed by linear regression models (19.4%).

Sequi et al. [30] recommended that the methodology of drug utilisation studies needs to be improved and we have also observed that drug use in the community is affected by drug availability, pricing, and affordability [34]. Therefore, the logistical and socio-economic aspects of pharmacoepidemiology studies should not be ignored. These two observations were the two key benchmarks for scoring the papers we have found and reviewed. For each study, we extracted information on the study design/type, data sources, period, assessment of variables used and corresponding statistical estimates (incidence, prevalence, pharmacy sales, prescription data), and diagnostic assessment. **Table 1** provides the overall summary details of the included papers.

By analysing **Table 1**, we have noticed that the status of pharmacoepidemiology research in some developing countries, like Malawi, is still at an infancy stage, compared to other developing countries that have adopted advanced inferential analyses into their pharmacoepidemiology research. Our findings do not differ from those reported by Sequi et al. [30], which the majority of the papers focused on the use of descriptive statistics. In addition, few studies clearly demonstrated the use of social/human behaviour network models in pharmacoepidemiology research [44, 45]. The inclusion of social/human behaviour network models into pharmacoepidemiology research is fundamental in the understanding of community structure and behaviour, for instance before mass drug administration during an outbreak such as COVID-19 [46, 47].

3.1.2 Big data in pharmacoepidemiology

Big data is another translational and frontier scientific discipline at the interface of computer science and statistics [48]. This field has found its way into pharmacoepidemiology research by simplifying the data interpretation and trend analysis of the volumes of data produced from many sources in health records [49]. With big data, pharmacoepidemiology research experts and data scientists detect ADRs, and collaborate in signal detection, verification and validation of medication or vaccine safety signals, as well as in the expansion of analytic methodologies for analysing the large volumes of heterogeneous data [14]. For example, the Exploring and Understanding Adverse Drug Reactions (EU-ADR) European project has incorporated innovative research methods in their pharmacovigilance research through the use of a web platform, aiming to provide advanced medication data exploration and assessment features. This enables data scientists and pharmacoepidemiology experts to mine EHRs for drug-events of their interest [4, 50].

Study type/design	Data source(s)	Year	Statistical methods	Variable(s) of interest	Reference
Cross-sectional	Survey questionnaire data	2018	Descriptive (percentages, frequencies, charts, median, ratios) *Excel*	Drug availability, Drug pricing, Affordability	[34]
Controlled trial	Articles	2017	-	Vaccination times, Dosage amounts	[35]
Cross-sectional	Prospective population census, passive surveillance, serological studies and healthcare utilisation surveys	2017	Descriptive (charts, percentages) *Stata*	Pathogen transmission, exposure and susceptibility	[36]
Randomisation	Basic survey	2019	Descriptive (percentages) *Excel & SPSS*	Drug abuse, Prevalence	[37]
Cohort	Anonymised patient record database	2013–2016	Descriptive (percentages), inferential-negative binomial regression (confidence intervals) *Stata*	Incidence and mortality ratios	[38]
Randomized Clinical Trial	Clinical data	2012	Descriptive (proportions), inferential (chi-square test, Kruskal-Wallis test, confidence intervals, incidence rate ratio, p-values, risk ratios) *Software- Stata*	Antiretroviral (ARV) usage, initiation	[39]
Key Informant Interviews (KII) and Focus Groups (FGs)	Recorded and transcribed qualitative data	2019	Thematic analysis *Software-NVivo*	Vaccination trials	[40]
Matched case–control study	Case–control study data	1993	Descriptive (tables, frequencies, percentages) and inferential (conditional logistic regression, relative risks, odds ratio, likelihood ratios, and confidence intervals) *Software- not mentioned*	BCG vaccine, efficacy, leprosy	[41]

Study type/design	Data source(s)	Year	Statistical methods	Variable(s) of interest	Reference
Cross-sectional	Drug prescription data from hospital electronic database	2020	Descriptive (frequencies and percentages for categorical variables) and (means, medians, standard deviations (SD), and interquartile ranges (IQR) for continuous variables). Mean and SD were used for normal distribution and median and IQR were used for skewed distribution. *SPSS and Excel*	Drug utilization	[42]
Retrospective	Pharmacokinetic data of children > = 2 years and adults	2018	Both descriptive and inferential models (mean absolute error from non-linear statistical models) *Stata*	Drug dosing and clearance	[43]

Table 1.
A review of computer aided research studies and usage of statistical models in Pharmacoepidemiology.

3.2 Databases

3.2.1 Importance of databases

Apart from the statistical innovations that have been incorporated into pharmacoepidemiology research, computer databases, networks and software are also playing a critical role in enhancing the field of pharmacoepidemiology, and notable developments have been reported in North America, Europe, and the Asia-Pacific region [51]. The rapid development of computer-aided technology has led to the improvement of electronic health records, which have further led to the advancement of many databases that may be used locally or internationally. Consequently, this has allowed for the possibility of conducting pharmacoepidemiology studies using multiple databases in one or more countries [5]. Several mechanisms have been developed to ensure maximum benefit from the multinational databases and collaborations, such as the creation of research networks [5].

The use of multinational databases enables researchers and policy makers to compare how medications and medical devices are utilised and prescribed, as well as to compare their safety profiles in different settings [51]. It also allows the identification of the underlying factors for the differences or similarities observed, which may include different patient selection, delivery systems and genetic differences [51]. Moreover, it relates drug effects (beneficial or adverse) with differences in ethnic groups (receptor and cytochrome polymorphism effect) and lifestyle (such as dietary habits), among others [52].

Furthermore, the use of multiple databases has overcome sample size problems for rare exposures, outcomes of medications, or rare diseases [5]. While it is challenging to get sufficient power when studying one area, data from multiple databases increase the sample size, thus providing the required statistical power. Additionally, the general use of meta-data may help to solve problems experienced by some countries or areas that do not have their own policies, medications, or medical devices [53]. Therefore, multiple databases provide reference points for such cases. Multiple databases also provide a platform for collaboration and communication amongst researchers in different and distant nations, which has led to the advancement of research in pharmacoepidemiology [5].

3.2.2 Multi-database networks

According to Sturkenboom and Schink [51], electronic healthcare databases have allowed analyses of drug and vaccine utilisation, including investigations of comparative effectiveness and safety. Consequently, both local and international databases have been developed worldwide for use in pharmacoepidemiology. In North America, administrative databases, such as the Health Services Databases in Saskatchewan [54] and the Ontario Health Insurance Plan [55] in Canada, have been set up to manage health care delivery costs, with the fundamental purpose of allowing fiscal tracking and accounting for the delivery of health care from a payer perspective. In the USA, databases managed by Government payers for claims data, for instance Medicaid and Medicare, data are also used in research [56].

Since some of the databases do not cover the entire population, some research networks have been set-up to facilitate multi-database studies that can cover the whole nation. These include the Canadian Drug Safety and Effectiveness Network (CDSEN), set-up in 2007 by the Canadian government, which connects multiple researchers across Canada with expertise in pharmacoepidemiology research [57, 58] as well as the USA Food and Drug Administration (FDA), whom established a

Sentinel Initiative in 2008 with the purpose of refining safety signals that would enable the development of a scalable and transparent organisational structure to study the safety of medical products [59], mainly through the organisation of multiple databases managed via one research governance structure [5, 60].

Similar initiatives have also been adopted in Europe. The EU-ADR [61] was initiated by the European Commission to develop a drug safety surveillance system reliant on connections amongst databases in European countries. This initiative benefits from reliable clinical data obtained from the electronic healthcare records of over 30 million of patients within all the participating countries, thus ensuring an efficient analysis of drug safety issues. Another initiative adopted along the same lines is the Pharmacoepidemiology Research on Outcomes of Therapeutics by an European ConsorTium (PROTECT), which involves 19 collaborative international working groups, networks and research projects in Europe [62]. Nordic countries have established the Nordic Pharmaco-Epidemiological Network (NorPEN), aiming to promote research collaboration and initiate cross-country population-based comparative research in pharmacoepidemiology, for further promotion of safer medication use [63].

The Asian Pharmacoepidemiology Network (AsPEN) was formed in 2008 by four countries, namely Korea, Japan, Australia, and Taiwan, and has currently expanded to Singapore, China, India, Hong Kong, and Thailand [64]. The AsPEN [65] was created to provide mechanisms for supporting pharmacoepidemiology research in Asia, as well as to facilitate the identification and validation of emerging safety issues among the Asian countries. The diversity of the countries provides multi-cultural and ethnic sources of safety data [63, 64]. Nevertheless, this is still an ongoing process, as some countries are still developing their own databases and infrastructures. Special attention should be given to the challenges of handling such multi-complex meta-data, and may involve collaboration of mathematicians, statisticians, epidemiologists and computer scientists (**Figure 1**).

Research networks specialised in certain subpopulations have also been initiated with the goal of studying populations under-represented in clinical trials, such as children, older people, and pregnant women. The most notable networks established for this purpose comprise the Task-force in Europe for Drug Development for the Young (TEDDY) [66]; the European network of population-based registries for the surveillance of congenital anomalies (EUROCAT) [67], for providing early warnings of new teratogenic exposures on congenital anomalies; the Innovative Medicines Initiatives (IMI) [68], for fostering collaboration between different stakeholders (the European Union and the European pharmaceutical industry) in order to address growing challenges in bringing new medicines to market and the rapidly evolving healthcare landscape; the VACCINE.GRID [69], a global network of leading public health organisations concerned with vaccine benefits and risk assessment; and the International Society for Pharmacoepidemiology (ISPE), an international professional organisation dedicated to the open exchange of scientific information for the benefit of people, drug safety in pregnancy, vaccine safety and/or biologics safety [70].

Last but not least, we have also noticed that computational infrastructures have been developed in places where data participants can transform their data locally, as well as execute standardised analytical programs and combine the results [45]. Data science has also been exploited in pharmacoepidemiology research, where it is used in the evaluation of various analytical methods in the context of a network of databases [45, 47]. Common data models that are capable of accommodating heterogeneous databases and executing large-scale statistical analyses [71–73], whose resources sometimes can be downloaded from a website [74], have also been developed. **Table 2** illustrates a few databases that are currently being used as well

Database name	Host(s)	Design	Data	Location	Target population	Data coverage	Reference(s)
Electronic Patient Registration System	Queen Elizabeth Central Hospital	Multiple	Multiple	Malawi	Various	Vital signs data, treatment, demographic data, diagnostic information	[38]
IADB.nl		Multiple	Multiple	Netherlands	Over 500,000 people	Live and stillbirth pregnancy identification, medicine use data, prescriptions from 54 community pharmacies	[52]
DEFF Research Database	Ministry of science, Technology and innovation; Ministry of Culture; Ministry of Education	Multiple	Multiple	Denmark	Countrywide	Dispensed drugs, with potential for linkage to outcomes	[75]
Odense University Pharmacoepidemiological Database (OPED)	University of Southern Denmark	Multiple	Multiple	County of Funen in Denmark	Countrywide	Reimbursed prescriptions	[76]
Disease Analyser Patient Database		Multiple	Multiple	Germany	German, UK, French, and Austrian population	Diagnoses, prescriptions, risk factors (such as smoking and obesity), and laboratory values for approximately 10 million patients	[77]
German Longitudinal Prescription Database (LRx)		Multiple	Multiple	Germany	Countrywide	Diseases, drug utilisation, treatment costs, 60% of prescriptions reimbursed by statutory health insurance funds in Germany	[78]

Database name	Host(s)	Design	Data	Location	Target population	Data coverage	Reference(s)
Database on Veterinary Clinical Research in Homeopathy.		Multiple	Multiple	Germany	Many	200 entries of randomised clinical trials, non-randomised clinical trials, observational studies, drug proving, case reports and case series	[79]
UK General Practice Research Database (GPRD).	UK Department of Health	Longitudinal	Case reports	UK	5 million patients; Countrywide	Collated information from over 500 general physicians' practices	[80, 81]
Clinical Research Database	Memorial Sloan-Kettering Cancer Center (MSKCC)	Multiple	Multiple		Patients on IRB approved studies who have passed through bone marrow transplant (BMT)	Diseases, pathology, infusion, treatment, among others.	[82]
Cancer Research DataBase (CRDB)	Cancer informatics project	Multiple	Mediation and data warehousing		Small molecule	Small molecule data, computational docking results, functional assays, and protein structure data	[83]
Danish Database for Biological Therapies in Rheumatology (DANBIO)	DANBIO	Multiple	Multiple	Denmark	Patients taking Biological treatments	Patients with rheumatoid arthritis (RA), psoriatic arthritis (PsA) and axial spondyloarthritis (Ax SpA), who are followed longitudinally	[84, 85]

Database name	Host(s)	Design	Data	Location	Target population	Data coverage	Reference(s)
The FoodCast Research Image Database (FRIDa)		Multiple	Multiple	Sweden	Wide range of foodstuff and related materials	877 images from eight different categories: natural-food, natural-non-food items. Artificial food-related objects	[86]
Pharmacy Dispensing Database		Multiple	Multiple	Netherlands, Denmark, Norway, Wales, France and Tuscany-Italy	Countrywide	Medicine use data	[87]
Danish National Patient Registry, Norway Medical Birth Registry		Multiple	Multiple	Norway and Denmark	Countrywide	Pregnancy loss identification	[87]
Influenza Research Database (IRD)	Bioinformatics Resource Center	Multiple	Multiple	US	All species of influenza virus sequence data	Influenza virus data, analytical and visualisation tools for influenza virus, personal workbenches for storing data	[88, 89]
Beth Israel Deaconess Medical Centre	Washington heart Centre, Beth Israel hospital, Boston	Multiple	Multiple	USA	Countrywide	Patient problems, medication, lab results	[90]
USDA's National Nutrient Database for Standard Reference, the Dietary Supplement Ingredient Database, the Food and Nutrient Database for Dietary Studies, and the USDA's Food Patterns Equivalents Database	US Department of Agriculture (USDA)	Multiple	Multiple	USA	Foodstuffs	Food and nutrients	[91]

Database name	Host(s)	Design	Data	Location	Target population	Data coverage	Reference(s)
Camden and Islington NHS Foundation Trust (C&I) Research Database	South London and Maussley NHS Foundation Trust (SLaM)	Multiple	Multiple	UK	Countrywide	108,168 mental health patients; 23,538 were receiving active care	[92]
Population and Housing Census (PHC), Health and Welfare Survey (HWS), Socio-Economic Survey (SES), Reproductive Health Survey (RHS), National Disability Survey (NDS), Multiple Indicator Cluster Survey (MICS)	National Statistics Office (NSO)	Interviews, face to face, self-enumeration, internet	Cross sectional	Thailand	Various	General population, health insurance, illness, health services, payment, equity, injury, co-morbidity, income, expenditure, debt, household distribution, family planning, maternal and child health. AIDS, Cancer, infertility, sex education, adolescent health	[93]
Cancer Registry	National Cancer Institute (NCI)	Longitudinal	Case reports	Thailand	All patients	Cancerous diseases, medicines	[93]
Thai Vigibase	Health Product Vigilance Centre (HPVC)		Case reports	Thailand	All patients	Adverse events	[93]
Adverse Events Database	Pharmaceutical and Medical Devices Agency (PMDA)		Case reports	Japan	Countrywide	Adverse events	[93]
National Community Pharmacy Group		Multiple	Multiple	South Africa	Countrywide	Drug utilisation	[94]

Database name	Host(s)	Design	Data	Location	Target population	Data coverage	Reference(s)
South African Medicine Claims Data	Pharmaceutical Benefit Management Company (PBM)	Multiple	Multiple	South Africa	Countrywide	Medicines claims	[95]
Strategic Typhoid Alliance Across Africa (STRATAA)				Malawi, Nepal, and Bangladesh	Countrywide	Demographic data, typhoid disease data	[96]
VigiBase	Uppsala Monitoring Centre (UMC)	Multiple	Multiple	Sweden	Worldwide	Adverse drug events	[97]
District Health Information System 2 (DHIS-2)	Kenya Medical Research Institute (KEMRI), Kamuzu Central Hospital	Multiple	Multiple	Kenya, Malawi, Uganda, Zambia [98]	Various	General health records and drug supply	[99, 100]
Mitishamba Database of Natural Products	University of Nairobi	Anti-Malaria drugs	Natural products	Kenya	Sub-Saharan Africa	Medicinal plants	[101]
International Databases to Evaluate AIDS (IeDEA-EA)	KEMRI, Mbarara Univ. & Tanzania	HIV-AIDS care	Drugs and Personal Protective Equipment (PPEs)	Kenya, Tanzania, Uganda	East African population	HIV care treatment	[102]

Table 2.
Computer databases currently used in pharmacoepidemiology research.

as those comprising data that may be potentially used to improve pharmacoepidemiology research. Although this is not an exhaustive list, these databases may serve as a supplement to those already reported [51].

Although the majority of pharmacoepidemiology research is found in developed countries, most of these databases are open for re-use of data, thus providing an opportunity for enhanced pharmacoepidemiology research, for instance in Asia and Africa [103].

3.2.3 Challenges with use of databases

Databases have limitations that affect their use in pharmacoepidemiology. Bias is one of the challenges and may be categorised into confounding, selection bias and time-related bias [98]. Confounding is further sub classified into confounding by indication, unmeasured or residual confounding, time-dependent confounding, and health user or adherer effect. Selection bias is reported to be associated with database use, being in the subcategories of protopathic bias, losses to follow up, prevalent user bias, and missing data. Another type of bias widely reported is measurement bias, which comes in the form of miscalculation bias, miscalculation of exposure, as well as miscalculation of outcomes. Time-related bias is classified into immortal bias, immeasurable time bias, time-window bias and time-lag bias [98].

4. Conclusions

Through a cross-examination of the intersection between data science principles and pharmacoepidemiology, this chapter has demonstrated that pharmacoepidemiology has greatly evolved over the years, from being a mere research field to one that is playing a significant role in the enhancement of patient safety, as well as in the development of health care guidelines and policies. Our examination of the intersection between data science techniques and pharmacoepidemiology was limited to the policy and research narratives of computer-aided pharmacoepidemiology studies across the globe. The level of evidence generated from several studies indicates that the field is now as important as randomised clinical trials have been, which can be attributed to the adoption of statistical and computational principles and practices. However, it is important to highlight that, although there has been a significant number of initiatives reported to improve pharmacoepidemiology research, the identified gaps and challenges presented in this chapter show that this field still has some potential to grow, for instance by properly integrating the existing data science techniques with appropriate principles and practices. The inclusion of both logistical and social/human behaviour network models into pharmacoepidemiology is strongly recommended.

Acknowledgements

Author contributions

IC conceived the study, performed the review of pharmacoepidemiology databases and participated in the manuscript writing process. EM reshaped the

argument of the study, reviewed research papers on statistical and computing models, and participated in the manuscript writing process. All authors have read and approved the final manuscript.

Conflict of interest

The authors declare no conflict of interest.

Appendices and nomenclature

ADRs	Adverse Drug Reactions
AIDS	Acquired Immunodeficiency Syndrome
ARV	Antiretroviral drugs
AsPEN	Asian Pharmacoepidemiology Network
BCG	BCG-Bacille Calmette-Guerin
BMT	Bone Marrow Transplant
CDSEN	Canadian Drug Safety and Effectiveness Network
COVID-19	Coronavirus Disease 2019
DANBIO	Danish Database for Biological Therapies in Rheumatology
DHIS-2	District Health Information System (version 2)
EHRs	Electronic Health Records
EU-ADR	Exploring and Understanding Adverse Drug Reactions
EUROCAT	European Network of Population-based Registries for the Surveillance of Congenital Anomalies
FDA	Food and Drug Administration
FGD	Focus Groups Discussion
FRIDa	The FoodCast Research Image Database
GPRD	UK General Practice Research Database
HPVC	Health Product Vigilance Centre
HWS	Health and Welfare Survey
IADB.nl	InterAction Database
IeDEA-EA	East African International Databases to Evaluate AIDS
IMI	Innovative Medicines Initiatives
IQR	Interquartile Range
IRD	Influenza Research Database
ISPE	International Society for Pharmacoepidemiology
KIIs	Key Informant Interviews
MICS	Multiple Indicator Cluster Survey
MSKCC	Memorial Sloan-Kettering Cancer Centre
NCI	National Cancer Institute
NDS	National Disability Survey
NorPEN	Nordic Pharmaco- Epidemiological Network
NSO	National Statistical Office
OPED	Odense University Pharmacoepidemiological Database
PBM	Pharmaceutical Benefit Management Company
PHC	Population and Housing Census
PMDA	Pharmaceutical and Medical Devices Agency
PROTECT	Research on Outcomes of Therapeutics by an European ConsorTium
RHS	Reproductive Health Survey
SD	Standard Deviations

SES	Socio-Economic Survey
SPSS	Statistical Package for Social Scientists
STRATAA	Strategic Typhoid Alliance across Africa
TEDDY	Task-force in Europe for Drug Development for the Young
USDA	U.S. Department of Agriculture

Author details

Ibrahim Chikowe[1]* and Elias Peter Mwakilama[2]

1 Pharmacy Department, College of Medicine, University of Malawi, Blantyre, Malawi

2 Mathematical Sciences Department, Chancellor College, University of Malawi, Zomba, Malawi

*Address all correspondence to: chikoweib@yahoo.co.uk; ichikowe@medcol.mw

IntechOpen

References

[1] Davis KAS, Farooq S, Hayes JF, John A, Lee W, MacCabe JH, et al. Pharmacoepidemiology research: delivering evidence about drug safety and effectiveness in mental health. The Lancet Psychiatry. 2020;7(4):363-370.

[2] Salas M, Stricker B. Research Methods for Pharmacoepidemiology Studies. In: S.P. G, editor. Essentials of Clinical Research [Internet]. Dordrecht: Springer; 2008. p. 201-16. Available from: https://doi.org/10.1007/978-1-4020-8486-7_12

[3] Yang Y, West-Strum D. Introduction to Pharmacoepidemiology. In: Yang Y, West-Strum D, editors. Understanding Pharmacoepidemiology. 1st ed. McGraw Hill; 2011.

[4] Nishtala P. Sources of data used in pharmacoepidemiology and pharmacovigilance. Elsevier [Internet]. 2019;391. Available from: https://doi.org/10.1016/B978-0-12-812735-3.00206-5

[5] Lai EC-C, Stang P, Yang Y-HK, Kubota K, Wong ICK, Setoguchi S. International Multi-database Pharmacoepidemiology: Potentials and Pitfalls. Curr Epidemiol Reports. 2015;2(4):229-38.

[6] Lasser KE, Allen PD, Woolhandler SJ, Himmelstein DU, Wolfe SM, Bor DH. Timing of new black box warnings and withdrawals for prescription medications. J Am Med Assoc. 2002; 287(17):2215-20.

[7] Strom BL, Carson JL. Use of Automated Databases for Pharmacoepidemiology Research. Epidemiol Rev [Internet]. 1990 Jan 1;12(1):87-107. Available from: https://doi.org/10.1093/oxfordjournals.epirev.a036064

[8] Lee D, Majumdar SR, Levens H, B. LS, Sean S, L. HR, et al. Special applications of pharmacoepidemiology. In: Strom BL, Kimmel SE, editors. Textbook of Pharmacoepidemiology. 2006.

[9] Nour S, Plourde G. Pharmaco epidemiology in the prevention of adverse drug reactions. In: Nour S, Plourde G, editors. Pharmaco epidemiology and pharmacovigilance [Internet]. Elsevier Inc.; 2019. p. 25-65. Available from: https://doi.org/10.1016/C2017-0-04746-8

[10] Faillie JL, Montastruc F, Montastruc JL, Pariente A. Pharmaco epidemiology and its input to pharmacovigilance. Therapie [Internet]. 2016;71(2):211-216,. Available from: https://doi.org/10.1016/j.therap.2016.02.016

[11] Wettermark B. The intriguing future of pharmacoepidemiology. Eur J Clin Pharmacol [Internet]. 2013;69:43-51,. Available from: https://doi.org/10.1007/s00228-013-1496-6

[12] Suissa S. Statistical Methods in Pharmacoepidemiology. Drug-Safety [Internet]. 1991;6:381-389. Available from: https://doi.org/10.2165/00002018-199106050-00008

[13] Hopf YM, Francis J, Helms PJ, Haughney J, Bond C. Core requirements for successful data linkage: An example of a triangulation method. BMJ Open. 2016;6(10).

[14] Sánchez-Duque JA, Gaviria-Mendoza A, Moreno-Gutiérrez PA, Machado-Alba JE. Big data, pharmacoepidemiology and pharmacovigilance. Rev Fac Med. 2020;68(1):117-20.

[15] Masnoon N, Shakib S, Kalisch-Ellett L, Al. E. Tools for Assessment of the Appropriateness of Prescribing and Association with Patient-Related

Outcomes: A Systematic Review. Drugs Aging [Internet]. 2018;35:43-60. Available from: https://doi.org/10.1007/s40266-018-0516-8

[16] Deloitte. Independent review of New Zealand's Electronic Health Records Strategy [Internet]. New Zealand: Deloite; 2015. Available from: https://www.health.govt.nz/system/files/documents/publications/independent-review-new-zealand-electronic-health-records-strategy-oct15.pdf

[17] Murray MD. Use of Data from Electronic Health Records for Pharmacoepidemiology. Curr Epidemiol Reports. 2014;1(4):186-93.

[18] MOH. Electronic Health Record [Internet]. 2021 [cited 2021 Mar 20]. Available from: https://www.health.govt.nz/our-work/ehealth/digital-health-2020/electronic-health-record

[19] Charles K, Cannon M, Hall R, Coustasse A. Can utilizing a computerized provider order entry (CPOE) system prevent hospital medical errors and adverse drug events? Perspect Health Inf Manag. 2014;11(March 2012).

[20] Nuckols TK, Smith-Spangler C, Morton SC, Asch SM, Patel VM, Anderson LJ, et al. The effectiveness of computerized order entry at reducing preventable adverse drug events and medication errors in hospital settings: A systematic review and meta-analysis. Syst Rev. 2014;3(1):1-12.

[21] Prgomet M, Li L, Niazkhani Z, Georgiou A, Westbrook JI. Impact of commercial computerized provider order entry (CPOE) and clinical decision support systems (CDSSs) on medication errors, length of stay, and mortality in intensive care units: A systematic review and meta-analysis. J Am Med Informatics Assoc. 2017;24(2):413-22.

[22] Harpe SE. Using secondary data sources for pharmacoepidemiology and outcomes research. Pharmacotherapy. 2009;29(2):138-53.

[23] Torre C, Martins AP. Overview of pharmacoepidemiological databases in the assessment of medicines under real-life conditions. In: Lunet N, editor. Current Perspectives on Research and Practice. IntechOpen; 2012. p. 186-193,.

[24] Sink KM, Thomas J, Xu H, Craig B, Kritchevsky S, Sands LP. Dual Use of Bladder Anticholinergics and Cholinesterase Inhibitors: Long-Term Functional and Cognitive Outcomes Kaycee. J Am Geriatr Soc. 2015;33(4): 395-401.

[25] Bentley JP. Biostatistics and Pharmacoepidemiology. In: Yang Y, West-Strum D, editors. Understanding Pharmacoepidemiology. New York: McGraw-Hill; 2011. p. 79-104.

[26] Suissa S. Statistical methods in pharmacoepidemiology: Advances and challenges. Stat Methods Med Res. 2009;18(1):3-6.

[27] Liu R, Tsong Y. Pharmaceutical Statistics. 1st ed. Vol. 218, MBSW 39, Muncie, Indiana, USA, May 16-18, 2016. Muncie, Indiana, USA: Springer International Publishing; 2019. XI, 337.

[28] Filardo G, Adams J, Keung HNT. Statistical Methods in Epidemiology. In: M. L, editor. International Encyclopedia of Statistical Science. Berlin, Heidelberg: Springer; 2011.

[29] Replicable WD, Non R, Wang S, Rassen J, Pinheiro S, Berlin J, et al. Abstracts of the 35th International Conference on Pharmacoepidemiology & Therapeutic Risk Management, Pennsylvania Convention Center, Philadelphia, PA, USA, August 24-28, 2019. Pharmacoepidemiol Drug Saf. 2019;28:5-586.

[30] Sequi M, Campi R, Clavenna A, Bonati M. Methods in pharmaco epidemiology: A review of statistical analyses and data reporting in pediatric drug utilization studies. Eur J Clin Pharmacol. 2013;69(3):599-604.

[31] Rosli R, Dali AF, Aziz NA, Abdullah AH, Ming LC, Manan MM. Drug utilization on neonatal wards: A systematic review of observational studies. Front Pharmacol. 2017;8(FEB).

[32] Al-Turkait A, Szatkowski L, Choonara I, Ojha S. Review of Drug Utilization Studies in Neonatal Units: A Global Perspective. Int J Environ Res Public Health. 2020;17(16):56-69.

[33] Hayat MJ, Powell A, Johnson T, Cadwell BL. Statistical methods used in the public health literature and implications for training of public health professionals. PLoS One. 2020;12(6).

[34] Khuluza F, Haefele-Abah C. The availability, prices and affordability of essential medicines in Malawi: A cross-sectional study. PLoS One. 2019;14(2):1-22.

[35] Gordon SB, Rylance J, Luck A, Jambo K, Ferreira DM, Manda-Taylor L, et al. A framework for Controlled Human Infection Model (CHIM) studies in Malawi: Report of a Wellcome Trust workshop on CHIM in Low Income Countries held in Blantyre, Malawi. Wellcome Open Res. 2017;2(0):1-11.

[36] Darton TC, Meiring JE, Tonks S, Khan MA, Khanam F, Shakya M, et al. The STRATAA study protocol: A programme to assess the burden of enteric fever in Bangladesh, Malawi and Nepal using prospective population census, passive surveillance, serological studies and healthcare utilisation surveys. BMJ Open. 2017;7(6):1-9.

[37] Kuyokwa J, Chiziwa S, Chinkhata M, Muyila D. Epidemiology of Psychoactive Substance Use and Associated Factors among Adolescents: A Descriptive Study of Selected Secondary Schools in South West Education Division, Blantyre, Malawi. Integr J Glob Heal. 2019;3(1:1).

[38] Lewis JM, Abouyannis M, Katha G, Nyirenda M, Chatsika G, Feasey NA, et al. Population Incidence and Mortality of Sepsis in an Urban African Setting, 2013-2016. Clin Infect Dis. 2020;10(1):2547-2552,.

[39] MacPherson P, Lalloo DG, Webb EL, Maheswaran H, Choko AT, Makombe SD, et al. Effect of optional home initiation of HIV care following HIV self-testing on antiretroviral therapy initiation among adults in Malawi: A randomized clinical trial. JAMA - J Am Med Assoc. 2014;312(4):372-9.

[40] Kapumba BM, Jambo K, Rylance J, Gmeiner M, Sambakunsi R, Parker M, et al. Stakeholder views on the acceptability of human infection studies in Malawi. BMC Med Ethics. 2020;21(1):1-15.

[41] Baker DM, Nguyen-Van-Tam JS, Smith SJ. Protective efficacy of BCG vaccine against leprosy in southern Malaŵi. Epidemiol Infect. 1993;111(1):21-5.

[42] Yue Y, Chen L, Choonara I, Xiong T, Ojha S, Tang J, et al. Cross-sectional study of drug utilisation in a Chinese neonatal unit. J Int Med Res [Internet]. 2020;48(5):1-15. Available from: https://doi.org/10.1177/0300060520914197

[43] Tegenge MA, Mahmood I, Jiang Z, Forshee R. Multistep Unified Models Using Prior Knowledge for the Prediction of Drug Clearance in Neonates and Infants. J Clin Pharmacol. 2018;58(7):877-884.

[44] Kiss IZ, Miller J, Simon PL. Mathematics of Epidemics on Networks:

From Exact to Approximate Models. 1st ed. Switzerland: Springer International Publishing; 2017. 413,.

[45] Broeck W V.D., Gioannini C, Gonçalves B, Quaggiotto M, Colizza V, Vespignani A. The GLEaMviz computational tool, a publicly available software to explore realistic epidemic spreading scenarios at the global scale. BMC Infect Dis. 2011;11.

[46] Li MY. An Introduction to Mathematical Modeling of Infectious Diseases. 1st ed. Vol. 2, Mathematics of Planet Earth. Switzerland: Springer International Publishing; 2018. 156.

[47] Xu B, Kraemer MUG, Gutierrez B, Mekaru S, Sewalk K, Loskill A, et al. Open access epidemiological data from the COVID-19 outbreak. Lancet Infect Dis. 2020;20(5):534.

[48] Mayer-Schönberger V, Cukier K. Big Data: A Revolution That Will Transform How We Live, Work, and Think. Reprint. New York: Eamon Dolan/ Mariner Books. ISBN-10; 2014. 544227751 p.

[49] Piontti AP y, Perra N, Rossi L, Samay N, Vespignani A. Charting the Next Pandemic: Modeling Infectious Disease Spreading in the Data Science Age. 1st ed. Switzerland: Springer International Publishing; 2019. 209,.

[50] Oliveira JL, Lopes P, Nunes T, Campos D, Boyer S, Ahlberg E, et al. The EU-ADR Web Platform: delivering advanced pharmacovigilance tools. Pharmacoepidemiol Drug Saf. 2013;22:459-467,.

[51] Sturkenboom M, Schink T. Databases for Pharmacoepidemiological Research. 1st ed. Switzerland: Springer International Publishing; 2021. 280 p.

[52] Visser ST, Schuiling-Veninga CC, Bos JH, Postma MJ, Berg LT de J den. The population-based prescription database IADB.nl: its development, usefulness in outcomes research and challenges. Expert Rev Pharmacoecon Outcomes Res. 2013;13(3):285-292,.

[53] Hennessy S. Use of Health Care Databases in Pharmacoepidemiology. Basic Clin Pharmacol Toxicol. 2006;98:311-313.

[54] Zuiderwijk A, Jeffery K, Janssen M. The Potential of Metadata for Linked Open Data and its Value for Users and Publishers. J e-Democracy Open Gov. 2012;4(2):222-244 ,.

[55] Fry RB, Ray MN, Cobaugh DJ, Weissman NW, Kiefe CI, Shewchuk RM, et al. Racial/ethnic disparities in patient-reported nonsteroidal antiinflammatory drug (NSAID) risk awareness, patient-doctor NSAID risk communication, and NSAID risk behavior. Arthritis Care Res. 2007;57(8):1539-45.

[56] Scales DC, Guan J, Martin CM, Redelmeier DA. Administrative data accurately identified intensive care unit admissions in Ontario. J Clin Epidemiol. 2006;59(8):802-7,.

[57] Hippisley-Cox J. Predicting risk of upper gastrointestinal bleed and intracranial bleed with anticoagulants: Cohort study to derive and validate the QBleed scores. BMJ [Internet]. 2014;349(July):1-21. Available from: http://dx.doi.org/doi:10.1136/bmj.g4606

[58] Lexchin J, Wiktorowicz M, Moscou K, Eggertson L. Provincial drug plan officials' views of the Canadian drug safety system. J Heal Polit Policy Law. 2013;38(3):545-71,.

[59] De Mello NR, Baracat EC, Tomaz G, Bedone AJ, Camargos A, Barbosa IC, et al. Double-blind study to evaluate efficacy and safety of meloxicam 7.5 mg and 15 mg versus mefenamic acid 1500 mg in the treatment of primary dysmenorrhea. Acta Obstet Gynecol Scand. 2004;83(7):667-73.

[60] Platt R, Carnahan RM, Brown JS, Chrischilles E, Curtis LH, Hennessy S, et al. The U.S. Food and Drug Administration's Mini-Sentinel program: status and direction. Pharmacoepidemiol Drug Saf. 2012;21(Suppl 1):1-8.

[61] EU-ADR. Exploring and Understanding Adverse Drug Reactions by Integrative Mining of Clinical Records and Biomedical Knowledge [Internet]. Exploring and Understanding Adverse Drug Reactions. 2020. Available from: http://www.euadr-project.org

[62] Sabaté M, Pacheco JF, Ballarín E, Ferrer P, Petri H, Hasford J, et al. A compilation of research working groups on drug utilisation across Europe. BMC Res Notes. 2014;7(1):1-5.

[63] Furu K, Wettermark B, Andersen M, Martikainen JE, Almarsdottir AB, Sørensen HT. The Nordic Countries as a Cohort for Pharmacoepidemiological Research. Basic Clin Pharmacol Toxicol. 2010;106(2):86-94.

[64] Andersen M, Bergman U, Choi N-K, Gerhard T, Huang C, Jalbert J, et al. The Asian Pharmacoepidemiology Network (AsPEN): promoting multi-national collaboration for pharmaco epidemiologic research in Asia. Pharmacoepidemiol Drug Saf. 2013;22(7):700-4,.

[65] AsPEN. Asian Pharmaco epidemiology Network [Internet]. Databases in Asia: The Potential for Distributed Network Approach. 2018 [cited 2021 May 16]. Available from: https://aspennet.asia/

[66] Ceci A, Giaquinto C, Aboulker J-P, Baiardi P, Bonifazi F, Pasqua O Della, et al. The Task-force in Europe for Drug Development for the Young (TEDDY) Network of Excellence. Pediatr Drugs Vol [Internet]. 2009;11:18-21. Available

from: https://doi.org/10.2165/0148581-200911010-00008

[67] Dolk H. EUROCAT: 25 Years of European surveillance of congenital anomalies. Arch Dis Child Fetal Neonatal Ed. 2005;90(5):355-8.

[68] Laverty H, Meulien P. The Innovative Medicines Initiative −10 Years of Public-Private Collaboration. Front Med. 2019;6(December):1-13.

[69] Grid. V. No Title [Internet]. [cited 2021 May 16]. Available from: http://www.vaccinegrid.org/public.html.

[70] ISPE. International Society for Pharmacoepidemiology [Internet]. [cited 2021 May 16]. Available from: https://www.pharmacoepi.org/

[71] Paris N, Parrot A. MIMIC in the OMOP Common Data Model. medRxiv [Internet]. 2020;(Cdm):2020.08.14. 20175141. Available from: https://doi.org/10.1101/2020.08.14.20175141

[72] Matcho A, Ryan P, Fife D, Reich C. Fidelity assessment of a clinical practice research datalink conversion to the OMOP Common Data Model. Drug Saf Int J Med Toxicol Drug Exp. 2014;37(11):945-59 ,.

[73] Makadia R, Ryan PB. Transforming the premier perspective hospital database into the Observational Medical Outcomes Partnership (OMOP) Common Data Model. EGEMS (Wash DC). 2010;2(1):1110.

[74] Overhage JM, Ryan PB, Reich CG, Hartzema AG, Stang PE. Validation of a common data model for active safety surveillance research. J Am Med Inf Assoc. 2012;19(1):54-60,.

[75] Wettermark B, Zoëga H, Furu K, Korhonen M, Hallas J, Nørgaard M, et al. The nordic prescription databases as a resurce for pharmacoepidemiological research-a literature review.

Pharmacoepidemiol Drug Saf. 2013;22(7):691-699,.

[76] Hallas J, Hellfritzsch M, Rix M, Olesen M, Reilev M, Pottegård A. Osense pharmacoepidemiological database: A review of use and content. Basic Clin Pharmacol Toxicol. 2017;120(5):419-425,.

[77] Becher H, Kostev K, Schröder-Bernhardi D. Validity and representativeness of the Disease Analyser patient database for use in pharmacoepidemiological and pharmacoeconomic studies. Int J Clin Pharmacol Ther. 2009;47(10): 617-626,.

[78] Richter H, Dombrowski S, Hamer H, Hadji P, Kostev K. Use of a German longitudinal prescription database (LRx) in pharmaco epidemiology. Ger Med Sci. 2015;13 (Doc14).

[79] Clausen J, Albrecht H. Database on veterinary clinical research in homeopathy. Homeopathy. 2010;99(3):189-191,.

[80] Bradley SH, Lawrence NR, Carder P. Using primary care data for health research in England – an overview. Futur Healthc J. 2018;5(3):207-12.

[81] Chen YC, Wu JC, Haschler I, Majeed A, Chen TJ, Wetter T. Academic impact of a public electronic health database: Bibliometric analysis of studies using the general practice research database. PLoS One. 2011;6(6):1-7.

[82] Latif M, Maloy M, Klein E, Fuller K, Truong H, Krishtul R. Redesign of an in house electronic data capture system to streamline hematopoietic stem cell informatics needs: Safeguarding over 40 years of data. Biol blood marrow Transplant [Internet]. 2019;25(3):269. Available from: https://doi.org/10.1016/j.bbmt.2018.12.210

[83] Bichutskiy VY, Colman R, Brachmann RK, Lathrop RH. Heterogeneous biomedical database integration using a hybrid strategy: A p53 cancer research database. Cancer Inform. 2006;2(949):277-87.

[84] Hetland ML. DANBIO-powerful research database and electronic patient record. Rheumatol. 2011;50(1):69-77,.

[85] Ibfelt EH, Jensen DV, Hetland ML. The Danish nationwide clinical register for patients with rheumatoid arthritis: DANBIO. Clin Epidemiol. 2016;8: 737-742,.

[86] Foroni F, Pergola G, Argiris G, Rumiati RI. The foodcast research image database (FRIDa). Front Hum Neurosci. 2013;7(FEB):1-19.

[87] Charlton RA, Neville AJ, Jordan S, Pierini A, Damase-Michel C, Klungsøyr K, et al. Healthcare databases in Europe for studying medicine use and safety during pregnancy. Pharmacoepidemiol Drug Saf. 2014;23:586-594.

[88] Rigden DJ, Fernández XM. The 27th annual Nucleic Acids Research database issue and molecular biology database collection. Nucleic Acids Res. 2020;48(D1):D1-8.

[89] Rigden DJ, Fernández XM. The 2018 Nucleic Acids Research database issue and the online molecular biology database collection. Nucleic Acids Res. 2018;46(D1):D1-7.

[90] Halamka JD. Early experiences with big data at an academic medical center. Health Aff. 2014;33(7):1132-8.

[91] Ahuja JKC, Moshfegh AJ, Holden JM, Harris E. USDA food and nutrient databases provide the infrastructure for food and nutrition research, policy, and practice. J Nutr. 2013;143(2):241-9.

[92] Werbeloff N, Osborn DPJ, Patel R, Taylor M, Stewart R, Broadbent M, et al. The camden & Islington research database: Using electronic mental health records for research. PLoS One. 2018;13(1):1-13.

[93] Saokaew S, Sugimoto T, Kamae I, Pratoomsoot C, Chaiyakunapruk N. Healthcare databases in Thailand and Japan: Potential sources for health technology assessment research. PLoS One [Internet]. 2015;10(11):1-20. Available from: http://dx.doi.org/10.1371/journal.pone.0141993

[94] Truter I. Antimicrobial prescribing in Southafrica using a large pharmacy database: A drug utilisation study IIse Truter. J Infect Dis [Internet]. 2014;30(2):52-56,. Available from: https://doi.org/10.1080/23120053.2015.1054181

[95] Obeng-Kusi M. Research methods for conducting pharmaco epidemiological studies using medicines claims data [Internet]. Masters Degree Dissertation. Northwest University; 2019. Available from: http://hdl.handle.net/10394/32816

[96] Thindwa D, Farooq YG, Shakya M, Saha N, Tonks S, Anokwa Y. Electronic data capture for large scale typhoid surveillance, household contact tracing, and health utilisation survey : strategic typhoid alliance across africa and asia. Wellcome Open Res. 2020;5(66).

[97] Ampadu HH, Hoekman J, de Bruin ML, Pal SN, Olsson S, Sartori D, et al. Adverse Drug Reaction Reporting in Africa and a Comparison of Individual Case Safety Report Characteristics Between Africa and the Rest of the World: Analyses of Spontaneous Reports in VigiBase®. Drug Saf. 2016;39(4):335-45.

[98] Prada-Ramallal G, Takkouche B, Figueiras A. Bias in pharmaco epidemiologic studies using secondary health care databases: a scoping review. BMC Med Res Methodol [Internet]. 2019;19(53). Available from: https://doi.org/10.1186/s12874-019-0695-y

[99] Chikumba P. Management of Health Information in Malawi: Role of Technology. Adv Sci Technol Eng Syst J. 2017;2:157-166 ,.

[100] Aywak D, Jaguga CDP, Nkonge NG, Kinuthia R, Ambale C, Awle IA. Pharmacy Practice in Kenya. Can J Hosp Pharm. 2017;70(6):456-462 ,.

[101] Noor AM, Alegana VA, Gething PW, Snow RW. A spatial national health facility database for public health sector planning in Kenya in 2008. Int J Heal Geogr. 2009;8(13).

[102] ISPM. International Databases to Evaluate AIDS (IeDEA-EA) [Internet]. 2021 [cited 2021 Mar 17]. Available from: https://www.iedea.org/regions/east-africa/

[103] Alarkawi D, Ali MS, Bliuc D, Center JR, Prieto-Alhambra D. The Challenges and Opportunities of Pharmacoepidemiology in Bone Diseases. JBMR Plus. 2018;2(4):187-94.