RESEARCH REPORT

Changes in the Landscape of Health Care Database Research from 2000 to 2011

Nuria Riera-Guardia, Catherine W. Saltus, Christine L. Bui, David H. Harris, James A. Kaye, Patricia Tennis, Jordi Castellsague, and Susana Perez-Gutthann

August 2013

RTI Press

RTI INTERNATIONAL

About the Authors

Nuria Riera-Guardia, PhD, is a research epidemiologist in the Epidemiology group at RTI Health Solutions, Barcelona, Spain.

Catherine (Kate) W. Saltus, MPH, MA, is a research epidemiologist in the Epidemiology group at RTI Health Solutions, Waltham, Massachusetts.

Christine L. Bui, MPH, is a research epidemiologist in the Epidemiology group at RTI Health Solutions, Research Triangle Park, North Carolina.

David H. Harris, MPH, is a research epidemiologist in the Epidemiology group at RTI Health Solutions, Research Triangle Park, North Carolina.

James A. Kaye, MD, DrPH, is a pharmacoepidemiologist in the Epidemiology group at RTI Health Solutions, Waltham, Massachusetts.

Patricia Tennis, PhD, FISPE, is a senior pharmacoepidemiologist in the Epidemiology group at RTI Health Solutions, Research Triangle Park, North Carolina.

Jordi Castellsague, MD, MPH, is a pharmacoepidemiologist in the Epidemiology group at RTI Health Solutions, Barcelona, Spain.

Susana Perez-Gutthann, MD, PhD, FISPE, is a pharmacoepidemiologist and head of the epidemiology team at RTI Health Solutions, Barcelona, Spain.

RTI Press publication RR-0019-1308

Suggested Citation

Riera-Guardia, N., Saltus, C. W., Bui, C. L., Harris, D. H., Kaye, J. A., Tennis, P., Castellsague, J., & Perez-Gutthann, S. (2013). *Changes in the landscape of health care database research from 2000 to 2011* (RTI Press publication No. RR-0019-1308). Research Triangle Park, NC: RTI Press. Retrieved from http://www.rti.org/rtipress

This publication is part of the RTI Reasearch Report series.

RTI International
3040 East Cornwallis Road
PO Box 12194
Research Triangle Park, NC
27709-2194 USA

Tel: +1.919.541.6000
Fax: +1.919.541.5985
E-mail: rtipress@rti.org
Web site: www.rti.org

doi: 10.3768/rtipress.2013.RR.0019.1308

www.rti.org/rtipress

Changes in the Landscape of Health Care Database Research from 2000 to 2011

Nuria Riera-Guardia, Catherine W. Saltus, Christine L. Bui, David H. Harris, James A. Kaye, Patricia Tennis, Jordi Castellsague, and Susana Perez-Gutthann

Abstract

This study aimed to quantify and characterize the use of automated health care databases for pharmacoepidemiology research over the past decade. To identify studies on drug safety, drug utilization, validation of database data, and disease epidemiology conducted in databases, we reviewed abstracts accepted at the International Conference of Pharmacoepidemiology in 2000 and 2011. The total number of abstracts doubled from 389 in 2000 to 806 in 2011. Abstracts on database studies comprised 35.7 percent (139 of 389) of all abstracts in 2000 and 44.6 percent (360 of 806) in 2011. The most common study objective in both years was drug safety evaluation. Abstracts on validation of database data increased from 2 to 39. Multiple-database abstracts increased from 9 to 43. The number of countries contributing to database abstracts increased from 14 to 22, with the increase primarily in Europe and the Asia-Pacific region. The findings suggest that the use of health care databases in pharmacoepidemiologic research is growing in many countries. The different types of database studies and the number of studies conducted using multiple databases are also increasing. This suggests that larger study populations and greater collaboration among investigators are becoming more common.

Contents

RTI INTERNATIONAL

Introduction

For more than two decades, computerized health databases containing medical care data have been considered important to understanding the real-world use, benefits, and adverse outcomes associated with pharmaceutical and biological therapies.[1] Health databases routinely record information on prescriptions written or dispensed as well as outpatient or hospital diagnoses, procedures, and interventions. These data are collected for administrative or insurance management purposes in claims databases or as part of the electronic medical records in which detailed clinical information is recorded by health care practitioners. These data sources appear to have become the cornerstone of pharmacoepidemiology research, yet almost no quantitative research has been conducted to document this impression. This study aimed to quantify and characterize the use of automated health care databases for pharmacoepidemiology research over the past decade.

Methods

We reviewed and compared the content of published abstracts for the International Conference of Pharmacoepidemiology (ICPE) in 2000 (Barcelona, Spain) and 2011 (Chicago, Illinois [USA]). Abstracts were published in supplements to *Pharmacoepidemiology and Drug Safety* in August 2000 and August 2011.

We defined a database as any longitudinal electronic collection of medical and/or administrative information for individual patients. Using a predefined abstraction process, the four of us who are epidemiologists abstracted data from studies conducted in databases. We abstracted the following information: ICPE year; abstract number; author name; study objective; number of databases involved; and name, country, and world region of the database. Study objectives were classified into five categories: safety endpoint, drug utilization or risk minimization evaluation, disease epidemiology, validation of database variables (e.g., diagnostic codes or algorithms to identify study endpoints, covariates, and exposures), and other (e.g., methodological issues, effectiveness, surveillance studies not involving a prespecified endpoint). We defined a multiple-database study as a study that used two or more databases to select the study subjects. Studies using database linkage to obtain additional information on study subjects already identified in a single database were not considered multiple-database studies. Finally, studies were not considered to be database studies if they were field studies (defined as a study conducted by recruiting patients to participate in the study), health economic studies, or studies of spontaneous reports that were conducted in databases.

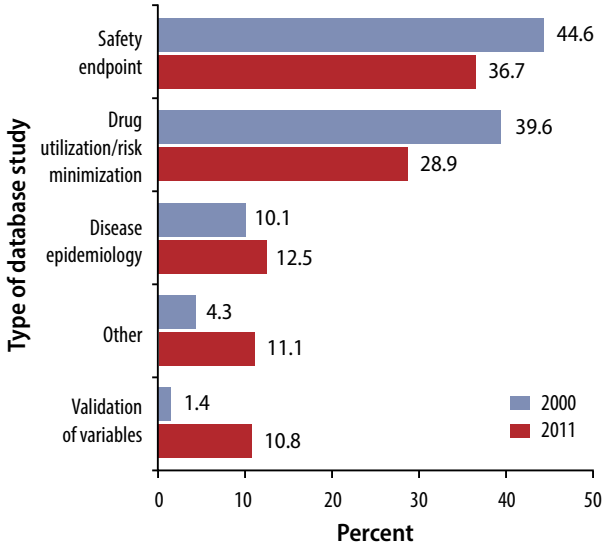
Each research epidemiologist reviewed and abstracted a set of abstracts, and each set was additionally reviewed by a different epidemiologist. The study team met regularly to discuss the consistency of abstractions. We contacted the abstract authors and reviewed published studies to collect essential information not included in the published abstract.

Additionally, we conducted a descriptive analysis and reported the number and percentage of database abstracts by year, study objective, country, and world region. We also enumerated and characterized the multiple-database studies.

Results

The total number of published ICPE abstracts doubled from 389 in 2000 to 806 in 2011. Abstracts on studies conducted in health care databases comprised 35.7 percent (139 of 389) of total ICPE abstracts in 2000 and 44.6 percent (360 of 806) in 2011. Figure 1 shows the distribution of study objectives of database studies by conference year. The most frequent study objective in both 2000 and 2011 was safety endpoint (44.6 percent in 2000 and 36.7 percent in 2011), followed by drug utilization/risk minimization, disease epidemiology, "other" study goal, and validation of variables. Drug utilization studies constituted 36.7 percent of database abstracts in 2000 and 26.4 percent in 2011, whereas risk minimization studies comprised 2.9 percent of studies in 2000 and 2.5 percent in 2011.

Figure 1. Database studies presented at the International Conference of Pharmacoepidemiology in 2000 and 2011, by study objective



The percentage of database studies on safety endpoints and drug utilization/risk minimization was lower in 2011 than in 2000, whereas the percentage of the "other" category of studies increased. The greatest increase occurred for studies on validation of data: from 1.4 percent in 2000 to 10.8 percent in 2011.

The combined categories of studies on safety endpoints and drug utilization/risk minimization accounted for 84.2 percent of all database abstracts in 2000 and 65.6 percent in 2011. Studies categorized as "other" included studies on epidemiological methods, statistical methods, and effectiveness in 2000; in 2011, "other" studies included studies on epidemiological methods, statistical methods, effectiveness, surveillance, and "other"—not falling into any of the previous categories (Table 1).

Table 1. Subcategories of studies categorized as "other"

Other Subcategory	2000	2011
Epidemiological methods	4	14
Statistical methods	1	14
Effectiveness	1	9
Surveillance	0	2
Other	0	1
Total	6	40

Most studies were single-database studies (Table 2), but the number of abstracts describing multiple-database studies increased from 9 (6.5 percent of all database abstracts) in 2000 to 43 (11.9 percent) in 2011. Of the 9 multiple-database abstracts published in 2000, only 1 study (11.1 percent) was conducted in multiple countries (data not shown). However, in 2011, 17 of 43 (39.5 percent) multiple-database abstracts involved more than one country.

Table 2. Number of databases included in each study, by study and year

Number of Databases per Study	2000, n (%)	2011, n (%)	Total
1	**130 (93.5)**	**317 (88.1)**	**447**
>1	**9 (6.5)**	**43 (11.9)**	**52**
2	5 (3.6)	17 (4.7)	22
3	2 (1.4)	12 (3.3)	14
4	1 (0.7)	3 (0.8)	4
5	0	2 (0.6)	2
6	0	1 (0.3)	1
7	0	7 (1.9)	7
8	1 (0.7)	1 (0.3)	2
Total	**139**	**360**	**499**

For the database abstracts, the overall number of database source countries increased from 14 to 22 (Figure 2). In 2000, the United Kingdom (UK), United States of America (USA), and Canada were the database source countries for more than 20 abstracts. In 2011, the USA, UK, Netherlands, Denmark, Taiwan, and Canada were the database source countries for 20 abstracts or more each. Countries not represented in 2000 that were represented in 2011 include Taiwan, Norway, Japan, Finland, Iceland, Israel, Brazil, Lithuania, Portugal, and Switzerland. Only two countries that had database studies in 2000 were not represented in 2011: Ireland and South Africa. Of the countries with fewer than 20 abstracts in 2000, the most notable absolute increase in number of abstracts by 2011 occurred for the Netherlands, Taiwan, Denmark, Sweden, and Norway.

Figure 2. Number and percentage of abstracts presented at the International Conference of Pharmacoepidemiology for research conducted in health care databases, 2000 and 2011, by country of database

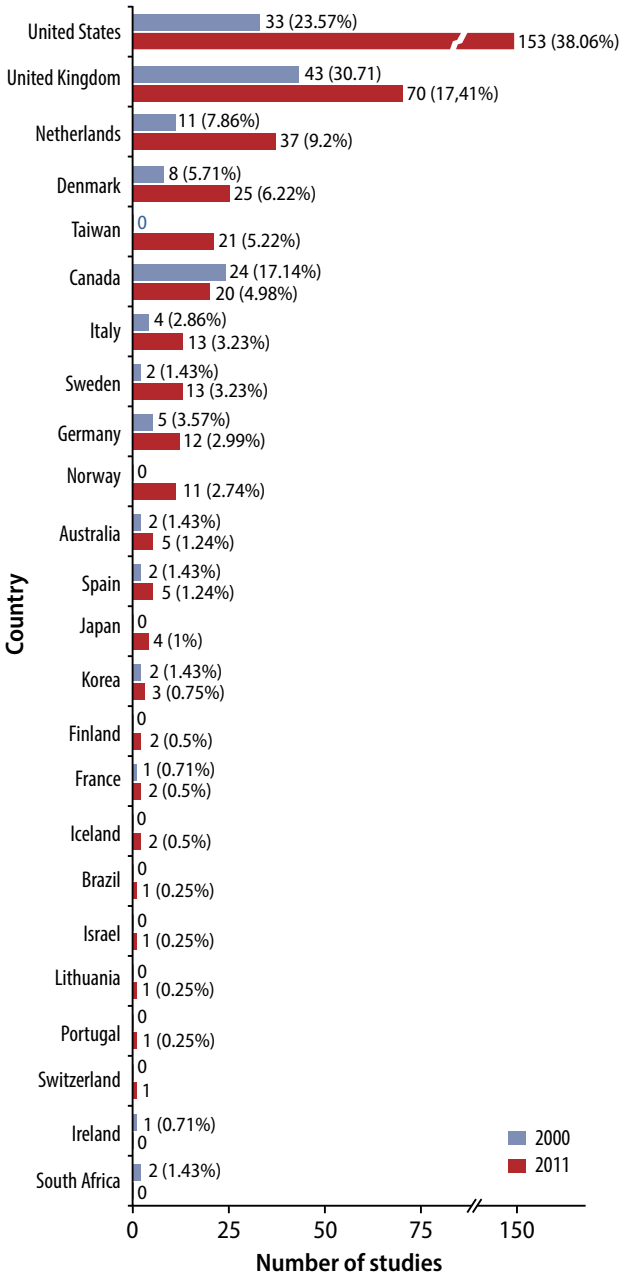
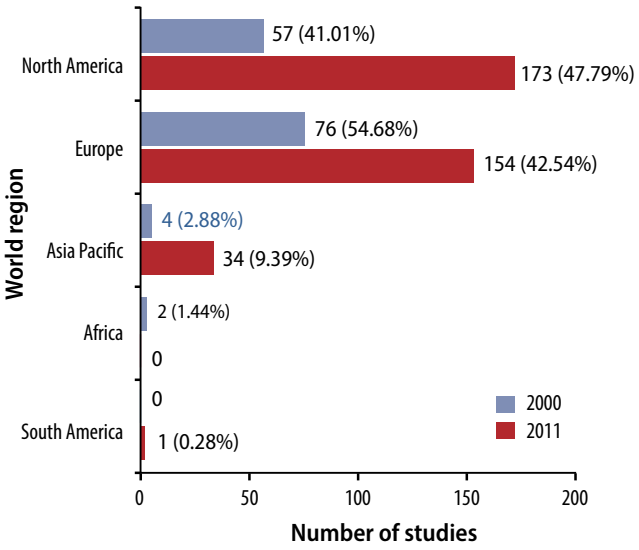


Figure 3 displays the number and percentage of database abstracts by world region and ICPE year. In 2000, Europe had the highest number of database studies (n=76), followed by North America (n=57). In 2011, North America had the highest number of database studies (n=173). The Asia-Pacific region had the largest proportional growth in database studies, with 4 abstracts in 2000 and 34 abstracts in 2011.

Figure 3. Number of abstracts presented at the International Conference of Pharmacoepidemiology on research conducted in health care databases, 2000 and 2011, by world region of database



Discussion

Over the past decade, the number of pharmacoepidemiology studies conducted using automated health care databases and presented at ICPE has experienced a remarkable expansion, and the number of countries where pharmacoepidemiology research is being conducted using databases has almost doubled. The Asia Pacific region had the largest percentage increase in abstracts on database studies. Studies using multiple databases from within a single country and across multiple countries have become more common. Database validation studies also increased between 2000 and 2011, representing almost 11 percent of the total in 2011. This is an encouraging finding in that the importance of validation to improving the quality of research is generally recognized.

These findings are consistent with the increasing availability of automated health care databases in many countries. Also, the size of the population covered by some databases allows the study of rare events and exposures that are difficult to study in other research settings. Collaborative research has been strongly encouraged by a variety of regulatory agencies and other initiatives. Some examples include the U.S. Food and Drug Administration's (FDA's) Sentinel Initiative[2] and the Observational Medical Outcomes Partnership (OMOP), a collaboration between PhRMA, the FDA, and the Foundation for the National Institutes of Health.[3] Other examples include the European Network of Centers for Pharmacoepidemiology and Pharmacovigilance (ENCePP),[4] which is led by the European Medicines Agency (EMA), and the Innovative Medicines Initiative, which is a public-private partnership between the European Union and the European Federation for Pharmaceutical Industries and Associations.[5]

Because the ICPE was held in Europe in 2000 and in North America in 2011, the different locations could account for part of the results, as it is likely that regional participation, including the number of abstracts, increases when the meeting is conducted in a particular world region.

Overall, the findings suggest that the use of automated health care databases in pharmacoepidemiology research is growing in many countries. The different types of database studies and the number of studies conducted using multiple databases is also increasing, suggesting that larger study populations and greater collaboration among investigators are becoming more common.

References

1. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol. 2005 Apr;58(4):323-7.
2. U.S. Food and Drug Administration [Internet]. FDA's Sentinel Initiative. [updated 2013 May 20; cited 2013 May 31]. Available from: http://www.fda.gov/safety/FDAsSentinelInitiative/ucm2007250.htm
3. Foundation for the National Institutes of Health [Internet]. About OMOP (Observational Medical Outcomes Partnership). [2013; cited 2013 May 31]. Available from: http://omop.org/
4. European Network of Centers for Pharmacoepidemiology and Pharmacovigilance (ENCePP) [Internet]. [updated 2013 July 12; cited 2013 July 12]. Available from: http://www.encepp.eu/
5. Innovative Medicines Initiative [Internet]. Mission. [2010; cited 2013 May 31]. Available from: http://www.imi.europa.eu/content/mission

Acknowledgments

This research was supported by RTI Health Solutions.

www.rti.org/rtipress **RTI Press publication RR-0019-1308**