French Medical-Administrative Database for Epidemiology and Safety in Ophthalmology (EPISAFE): The EPISAFE Collaboration Program in Cataract Surgery

Vincent Daien\textsuperscript{a–c} Jean-Francois Korobelnik\textsuperscript{e–g} Cécile Delcourt\textsuperscript{f, g} Audrey Cougnard-Gregoire\textsuperscript{f, g} Marie Noelle Delyfer\textsuperscript{e–g} Alain M. Bron\textsuperscript{h, i} Isabelle Carrière\textsuperscript{b, c} Max Villain\textsuperscript{a, b} Jean Pierre Daures\textsuperscript{d} Sandy Lacombe\textsuperscript{d} Anne Sophie Mariet\textsuperscript{i–n} Catherine Quantin\textsuperscript{i–n} Catherine Creuzot-Garcher\textsuperscript{h, i}

\textsuperscript{a}Department of Ophthalmology, Montpellier University Hospital, \textsuperscript{b}University of Montpellier, \textsuperscript{c}Inserm U1061, and \textsuperscript{d}Department of Biostatistics, Epidemiology and Clinical Research EA2415, University of Montpellier, Montpellier, \textsuperscript{e}Service d’Ophthalmologie, CHU de Bordeaux, \textsuperscript{f}ISPED, Université de Bordeaux, and \textsuperscript{g}Inserm U1219, Bordeaux Population Health Research Center, Bordeaux, and \textsuperscript{h}Department of Ophthalmology, University Hospital Dijon, \textsuperscript{i}Eye and Nutrition Research Group, CSGA, UMR 1324 INRA, \textsuperscript{j}Biostatistics and Bioinformatics (DIM), \textsuperscript{k}Bourgogne Franche-Comté University, \textsuperscript{l}Inserm CIC 1432, \textsuperscript{m}Clinical Epidemiology/Clinical Trials Unit, Clinical Investigation Center, Dijon University Hospital, and \textsuperscript{n}Inserm UMR 1181, Biostatistics, Biomathematics, Pharmacoepidemiology and Infectious Diseases (B2PHI), Bourgogne Franche-Comté University, Dijon, France

Keywords
Big data · Pharmacoepidemiology · Safety · Ophthalmology · Cataract

Abstract
Medical-administrative databases are an important source of big data to assess the epidemiology of diseases and interventions, compare drugs, and investigate rare adverse events. We describe the French national health insurance system databases and the Epidemiology and Safety (EPISAFE) collaboration program in ophthalmology and illustrate the paper with recent studies that used the databases to investigate cataract surgery. The Système national d’information inter-régime de l’assurance maladie (SNIIR-AM) is one of the largest databases of medical and administrative data for a general population, covering nearly 65 million people. The SNIIR-AM database links data for all outpatient-reimbursed health expenditures, hospitalization in all 1,546 French private or public healthcare facilities, and date of death. The use of this database provides information at the scale of the entire country, allowing measurement of the impact of changes in practices and new guidelines. Between 2009 and 2012, a total of 2,717,203 eyes in 1,817,865 patients underwent cataract surgery in France, and the incidence of cataract surgery increased from 9.86 to 11.08/1,000 person-years. The 4-year cumulative risk of retinal detachment was 0.99%, and patients <60 years of age were at higher risk of retinal detachment. The incidence of pseudophakic cystoid macular edema was 0.95%. From 2005 to 2014, from data including more than 6 million procedures, the incidence of endophthalmitis after cataract surgery decreased from 0.15 to 0.05%. The EPISAFE collaboration program encompasses the key issues facing ophthalmology, including public health and public policy concerning disease incidence and prevalence, safety and adverse event monitoring, and clinical guidelines.

© 2017 S. Karger AG, Basel
**Introduction**

Detailed epidemiological data regarding interventions in ophthalmology are important for determining community needs and assessing the potential impact of intervention strategies. The World Health Organization encourages countries to select indicators that are the most appropriate for their epidemiological situation and intervention strategy [1–3].

An increasing number of interventions with varying risks and benefits are becoming available in ophthalmology in the important fields of age-related macular degeneration, diabetic retinopathy, and glaucoma. The number of therapies for age-related macular degeneration will increase in the future since many new agents with different modes of action are currently being tested in phase II clinical trials. Postmarketing medical-administrative studies may be the best way to evaluate the safety of these new agents since it will not be possible to test them all in randomized controlled trials (RCTs).

Substantial work is in progress to provide pooled epidemiological data from population-based and case-control studies as well as randomized trials in Europe [4]. Nevertheless, the number of subjects and patients remains limited, the studies are costly and time-consuming, and the populations studied are not always representative of the whole population of a country.

Another source of information is medical-administrative databases. Such databases have led to significant advances in epidemiological, pharmacoepidemiological, and healthcare economics studies. Many countries support research through their healthcare administrative data; examples are US Medicare (about 55 million people) [5], the UK Clinical Practice Research Datalink (approx. 5 million people) [6], the administrative medical databases of the Régie d’assurance maladie du Québec (approx. 4 million people) [7], and databases in Sweden [8].

Through public entities, the French health insurance system manages all reimbursements of healthcare to all people affiliated with a health insurance scheme in France. The Système national d’information inter-régime de l’assurance maladie (SNIIR-AM) was created in 2003 with the 1999 healthcare funding law to improve the quality of healthcare and contribute to better management of the health insurance system and public health policies.

The result is one of the largest global databases of medical and administrative data in the general population, covering nearly 65 million people. Accessibility to researchers is regularly improved. Since 2010, multiple studies using this system of databases have been conducted in France to assess the epidemiology of rare diseases [9], compare the effectiveness of drugs [10], investigate rare adverse drug events [11], and examine vaccination coverage [12].

We established the Epidemiology and Safety (EPISAFE) collaborative program aiming at assessing the epidemiology and safety of interventions in ophthalmology. This article describes the scope and strategy of this program, as well as the French national health insurance system databases for medical research upon which it is based. We illustrated the paper with recent studies investigating the epidemiology and safety of cataract surgery in France.

**Architecture of the Databases**

**The SNIIR-AM**

Healthcare data are collected prospectively and then transmitted to the National Health Insurance Fund for Salaried Workers (Caisse nationale de l’assurance maladie des travailleurs salariés – CNAMTS), which holds all data in a large digital warehouse called SNIIR-AM [13, 14]. The SNIIR-AM database links data for all outpatient-reimbursed health expenditures from the Données de consommation inter-régimes (DCIR database), hospitalization in all 1,546 French private or public healthcare facilities from the Programme de médicalisation des systèmes d’information (PMSI database), and the date of death from the Institut national des statistiques et des études économiques (INSEE) database. The architecture of SNIIR-AM and the main data recorded are described in Table 1.

The DCIR

The inter-scheme consumption data set includes all outpatient-reimbursed health expenditures. These data are transmitted in real time by the CNAMTS and the various special insurance schemes. This DCIR database is also linked to data from the health insurance medical service that includes information on long-term diseases, occupational accidents, and occupational diseases.

The PMSI

PMSI data are gathered from all private or public hospitals and provided to the CNAMTS for linkage to the SNIIR-AM. Information includes the hospital identifier, the start and end dates of hospital stays, and the diagnosis (1 primary diagnosis, 1 related diagnosis, and associated diagnoses) coded with the International Classification of Diseases, version ICD-10 [15].

Demographics (age, gender, place of residence [municipality code], universal health coverage, affiliate health
insurance scheme) are available in the DCIR and PMSI databases.

Data Protection
Data for each insured person are linked by a unique identification number, the Numéro d’identification au répertoire (NIR). However, to safeguard the identity of patients, the NIR is not accessible in the SNIIR-AM. It is made anonymous by a 2-level procedure, which is completely irreversible [13]. Working on SNIIR-AM data extractions requires having a secured server with sufficient capacity.

Depth of the Databases
The depth of the SNIIR-AM is 3 years plus the current year. The PMSI data represent a sliding range of 10 years plus the current year. The law allows for conservation of successive SNIIR-AM extractions corresponding to 10 years of follow-up.

Echantillon Généralisé des Bénéficiaires
The Echantillon Généralisé des Bénéficiaires (EGB), established in 2005, is a sample of 1/97th of insured people in France, corresponding to 660,000 people [13, 16]. It is a random selection of NIRs representative of the French population by gender and by 5-year increments [13].

The EGB population includes insured persons whether or not they receive healthcare. In contrast, the SNIIR-AM includes only people receiving healthcare. All health insurance schemes are included, except special insurance schemes whose data are progressively being implemented. The data available come from the SNIIR-AM and are also anonymized. Individual data are stored and are accessible for 20 years. Patients who die, emigrate, or are no longer insured are considered out of the cohort: their data are retained but are censored on the release date. Newly insured people are added quarterly to the EGB to retain the 1/97th proportion of insured people in France. Because of the depth of the database and accessibility, the EGB is the database of choice for establishing long-term cohorts of patients with common diseases.

**EPISAFE Collaborative Program**

The epidemiology of common diseases is essential for planning future healthcare provision. Population-based cohort studies are most often based on a sample of the population, whereas French medical-administrative databases are exhaustive and represent the total French population, thus preventing selection bias.

The EPISAFE collaborative program follows the growing collaborations between Bordeaux, Dijon, and Montpellier ophthalmological researchers who have focused for many years on ophthalmic epidemiology through the 3-City studies [17–20]. Our group includes expertise in ophthalmology, epidemiology, biostatistics, and health economics.

We aim to provide data on the nationwide incidence and prevalence of interventions in ophthalmology, as we did recently for cataract surgery [21]. Moreover, many RCTs are initially conducted to demonstrate the efficacy of drugs and report positive results. However, since many procedures in ophthalmology are used in millions of patients, efforts to evaluate safety have emerged [22]. In addition, the statistical power to assess the safety of drugs is often lacking in RCTs because side

---

**Table 1. Simplified architecture of the French SNIIR-AM**

<table>
<thead>
<tr>
<th>SNIIR-AM</th>
<th>DCIR data mart</th>
<th>Healthcare dispensing</th>
<th>PMSI data mart</th>
<th>Hospitals</th>
<th>INSEE Vital records office</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic data</td>
<td>Medical data</td>
<td>Costly long-term diseases</td>
<td>Age, gender</td>
<td>Place of residency</td>
<td>Universal health coverage</td>
</tr>
<tr>
<td>Demographic data</td>
<td>Medical data</td>
<td>Accidents and diseases</td>
<td>Out-of-hospital reimbursements</td>
<td>Date</td>
<td>Drugs: name, form, quantity dispensed</td>
</tr>
<tr>
<td>Demographic data</td>
<td>Medical data</td>
<td>Sick leave</td>
<td>Out-of-hospital reimbursements</td>
<td>Date</td>
<td>Drugs: name, form, quantity dispensed</td>
</tr>
<tr>
<td>Demographic data</td>
<td>Medical data</td>
<td>Sick leave</td>
<td>Out-of-hospital reimbursements</td>
<td>Date</td>
<td>Drugs: name, form, quantity dispensed</td>
</tr>
</tbody>
</table>

SNIIR-AM, Système national d’information inter-régime de l’assurance maladie (National Health Insurance Information System) hosted by the Caisse nationale de l’assurance maladie des travailleurs salariés (National Health Insurance Fund for Salaried Workers); DCIR, Données de consommation inter-régimes (Inter-Scheme Consumption Data); INSEE, Institut national des statistiques et des études économiques (National Institute of Statistics and Economic Studies); PMSI, Programme de médicalisation des systèmes d’information (Program for the Medicalization of Information Systems).
have presented recent studies in cataract surgery that we
agement of diabetic retinopathy.
costeroid safety, retinal and glaucoma surgery, and man-
ministrative studies are suitable include intravitreal injec-
research projects. Current challenges where medical-ad-
collaborative group has regular meetings to coordinate
retinal detachment [27]
factors of pseudophakic
Incidence and risk
effects of drugs or complications of procedures are, hope-
fully, uncommon [23]. Using French medical-adminis-
tative databases and high statistical power, the EPISAFE
collaborative program aims to further explore the safety
of drugs and interventions, as we did recently for endo-
phthalmitis after cataract surgery [24, 25]. The EPISAFE
collaborative group has regular meetings to coordinate
research projects. Current challenges where medical-ad-
ministrative studies are suitable include intravitreal injec-
tion of antivascular endothelial growth factor and corti-
costeroid safety, retinal and glaucoma surgery, and man-
agement of diabetic retinopathy.

EPISAFE Collaborative Program in Cataract Surgery
To illustrate the EPISAFE collaborative program, we
have presented recent studies in cataract surgery that we
have conducted based on French medical-administrative
databases (Table 2).

Data for all patients who underwent primary cataract
surgery in France between January 2009 and December
2012 were extracted from the database [21]. Annual inci-
dence rates were calculated and adjusted to the corre-
sponding-year national population data from INSEE
[26]. Over the 4 years, 2,717,203 eyes in 1,817,865 patients
(59.1% women; mean age, 73.5 ± 0.015 years) underwent
cataract surgery. Between 2009 and 2012, the incidence of
endophthalmitis (POE) decreased from 0.145 to 0.053% during this 10-year period
intracameral antibiotic injection was associated with a lower risk of acute
POE, whereas intraoperative posterior capsule rupture, combined surgery,
and gender (male) were associated with a higher risk of acute POE

\begin{table}
\centering
\caption{Epidemiology and safety (EPISAFE) of cataract surgery in France using medical-administrative databases}
\begin{tabular}{|l|l|m|}
\hline
\textbf{Topic} & \textbf{Population} & \textbf{Main findings} \\
\hline
Epidemiology of cataract surgery in France [21] & 2,717,203 eyes in 1,817,865 patients who underwent cataract surgery in France between 2009 and 2012 & \\
\textbullet\ the incidence of cataract surgery increased from 9.86 to 11.08/1,000 person-years \\
\textbullet\ the incidence of cataract surgery ranged from 1.06/1,000 person-years for patients 40–49 years old to 65.94/1,000 person-years for those 80–89 years old \\
\textbullet\ between 2009 and 2012, the probability of second-eye surgery 12 months after the first-eye surgery increased from 40.6 to 51.2% \\
\textbullet\ the median interval for surgery between eyes was 29 days (interquartile range, 14–86) and decreased over the study period \\
\hline
Incidence and risk factors of pseudophakic retinal detachment [27] & 2,717,203 eyes in 1,817,865 patients who underwent cataract surgery between 2009 and 2012 & \\
\textbullet\ the estimated 4-year cumulative risk of pseudophakic retinal detachment was 0.99% \\
\textbullet\ the median (interquartile range) number of days between cataract surgery and pseudophakic retinal detachment onset was 237 (range, 75–738) \\
\textbullet\ patients <60 years of age were at higher risk of retinal detachment \\
\hline
Incidence and characteristics of pseudophakic cystoid macular edema [28] & 19,980 eyes in 13,556 patients who underwent cataract surgery in Languedoc-Roussillon region in 2010 & \\
\textbullet\ the incidence of pseudophakic cystoid macular edema was 0.95% \\
\textbullet\ the median duration of pseudophakic cystoid macular edema was 2.8 months (range, 1.9–4.9) \\
\textbullet\ in 9.4% patients, pseudophakic cystoid macular edema persisted for 6–12 months and in 8.6% for >12 months \\
\hline
\textbullet\ the incidence of acute postoperative endophthalmitis (POE) decreased from 0.145 to 0.053% during this 10-year period \\
\textbullet\ intracameral antibiotic injection was associated with a lower risk of acute POE \\
\hline
Effectiveness and safety of an intracameral injection of antibiotics for endophthalmitis after cataract surgery with or without perioperative capsular rupture [25] & 3,351,401 eyes of 2,434,008 patients who underwent cataract surgery in France between 2010 and 2014 & \\
\textbullet\ the retinal safety of cefuroxime, assessed by multi-adjusted odds of retinal cystoid macular edema, was not increased for patients receiving cefuroxime injections \\
\textbullet\ the retinal safety of cefuroxime, assessed by multi-adjusted odds of retinal cystoid macular edema, was not increased for patients receiving cefuroxime injections in patients with perioperative capsular rupture of the lens (the major risk factor for POE) and intracameral antibiotic injection, POE incidence was lower with than without cefuroxime (0.37 vs. 0.51%, respectively, p = 0.001), while an increased risk of cystoid macular edema was not identified (5.6% with vs. 7.3% without, respectively, p = 0.12) \\
\end{tabular}
\end{table}

Daien et al.
surgery, with an estimated risk of 0.99% at 4 years after surgery. The median (interquartile range) number of days between cataract surgery and retinal detachment onset was 237 (range, 75–738), and patients <60 years of age were at higher risk of retinal detachment [27]. Thereafter, the incidence of cystoid macular edema after cataract surgery was 0.95% [28]. More recently, 2 papers assessed endophthalmitis after cataract surgery. The first found a decrease in endophthalmitis after cataract surgery from 2005 to 2014 and, within the same period, an increased use of intracameral antibiotics [24]. The second observed that with perioperative capsular rupture of the lens, the incidence of endophthalmitis was lower for patients with than without an intracameral injection of antibiotics, with no increased risk of cystoid macular edema [25].

Although cataract surgery is highly effective and relatively safe, surgical complications are still a major source of morbidity because of the high case load and the increasing number of cataract surgeries. These data are important because cataract surgery is the most commonly performed procedure in people older than 65 years in the developed world [29]. They may allow practitioners to benchmark their complication rates and, if necessary, refine their techniques or indications. The use of French medical-administrative databases also provides information at the scale of the whole country, allowing us to measure the impact of changes in practices and new guidelines. The recommendations of the Société Française d’Ophthalmologie made in 2012 that recommended prophylactic intracameral antibiotic injection of cefuroxime during cataract surgery is a good example, allowing us to precisely assess the evolution of the endophthalmitis rate while recording intraoperative intracameral injection procedures at the same time.

**Strengths and Limitations of French Medical-Administrative Databases**

French medical-administrative data are automatically collected from all outpatient-reimbursed and hospitalization discharge abstracts. This is different in nature and in a certain way completes data from large registries such as the Intelligent Research in Sight (IRIS®) [30] and the Save Sight Registries [31], the UK National Ophthalmology Database [32], and the Swedish National Cataract Register [33]. The exhaustivity of the French database comes from the need for doctors to record accurate diagnoses to receive payment from reimbursement agencies.

**In Epidemiology**

Considering the grades of evidence for the purported quality of study design, studies from medical-administrative databases are at intermediate levels in the evidence hierarchy for evaluating clinical studies [34]. However, in many aspects, they can complement findings from RCTs, which assess outcomes under ideal conditions in selected patients and with limited follow-up [35]. RCTs have higher internal validity because of randomization, and medical-administrative studies have higher external validity since they reflect current real-world practices.

The major advantage of these databases is the huge number of patients they include and the prospective data collection. Therefore, large cohorts can be built from these databases to clarify the epidemiology of diseases nationwide, taking into account geographical variations. The date of death is valuable information recorded since 2008, and the cause of death was also retrieved in a previous study [36]. The SNIIR-AM is a useful data source to estimate the incidence and prevalence of rare diseases or complications such as postoperative endophthalmitis [24, 25]. Mother-to-child studies are possible since this linkage was set up in 2012. Social disadvantage can also be evaluated as a risk factor for disease [37]. Linkage between clinical cohorts such as the Gazel and Constances cohorts and the SNIIR-AM is also possible [38].

**In Pharmacoepidemiology**

The field of pharmacoepidemiology assesses the use, effectiveness, and safety of drugs in real-world practice (after marketing authorization). Therefore, health insurance databases are frequently used for pharmacoepidemiological studies [39]. These databases represent powerful material to compare the efficacy of 2 drugs with the same indication but that have not been compared in RCTs, as was done recently with the SNIIR-AM [10, 40, 41]. Moreover, they allow for stratification of the analyses by subgroups of patients excluded from clinical trials (older patients with comorbidities, polymedicated patients, etc.) because of the large number of patients they include.

Health insurance databases are the tools of choice to assess compliance with recommendations at the population level [42]. However, they cannot be used to accurately assess the reasons for nonadherence to recommendations.

**Limitations**

The main limitation of these databases is their complexity. An extraction of the SNIIR-AM includes several dozen tables, some used only for accountability, contain-
ing hundreds of variables and millions of lines. An SNIIR-AM dictionary is available for users of the health insurance databases. However, training in the complex architecture of these databases has been mandatory since 2013 [43].

A possible limitation in epidemiology is the validity of diagnostic codes. Also, the indications for treatments and paraclinical examinations are not recorded. They are deduced from long-term disease and hospital diagnostic codes and from procedures or drugs used as proxies. Limitations in coding accuracy are inherent to those of the ICD-10 [15]. Quality control before data transmission to the CNAMTS and audits mainly involves the treatment of aberrant or missing data [13]. For example, in the study assessing the epidemiology of cataract surgery, 0.26% of patients were not included because of presumed coding errors [21].

Only the dispensing of reimbursed drugs is recorded in the databases: self-medication with over-the-counter drugs cannot be measured. Detailed information regarding the dose of pharmaceutical forms and the number of units dispensed are available; however, the prescribed dose is not recorded. This can be modeled with pharmacoepidemiological methods. These approaches most frequently use the World Health Organization’s Defined Daily Dose for each drug. Studies on identical twins are limited because they have the same identifier number. Longitudinal studies should track whether patients lost to follow-up are related to changes in identifier number, which remains possible in some circumstances such as marriage or reaching 18 years of age. A national health identifier will soon be created to solve this problem [44, 45]. Other limitations are the lack of smoking status and visual acuity in the database.

Conclusion

The SNIIR-AM and EGB offer great possibilities for medical research in building cohorts with prospective recording of data and unequalled statistical power.

The EPISAFE collaboration program encompasses the key issues facing ophthalmology, including public health and public policy concerning disease incidence and prevalence, utilization of eye care services, safety and adverse event monitoring, and clinical guidelines.

Disclosure Statement

All authors have completed the ICMJE (International Committee of Medical Journal Editors) uniform disclosure form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author). V.D. reports personal fees and nonfinancial support from Alcon, Bayer, Novartis, and Théa, outside the submitted work; J.-F.K. reports personal fees from Alcon, Novartis, Bayer, Théa, Carl Zeiss Meditec, Allimer, Allergan, and Roche, outside the submitted work; C.D. reports personal fees from Allergan, personal fees from Bausch & Lomb, Novartis, and Roche, and grants and personal fees from Laboratoires Théa, outside the submitted work; A.C.-G. discloses no conflict of interest; M.N.D. reports personal fees from Alcon, grants from Horus, and grants and personal fees from Bausch & Lomb and Théa; J.P.D. discloses no conflict of interest; S.L. discloses no conflict of interest; A.S.M. discloses no conflict of interest; C.Q. discloses no conflict of interest; I.C. discloses no conflict of interest; M.V. reports personal fees from Allergan, outside the submitted work; and C.C.-G. reports grants and personal fees from Bausch & Lomb, Horus, and Théa, and personal fees from Alcon, Allergan, Bayer, and Novartis, outside the submitted work.

All authors are independent of financial support. They declare no relationships or activities that could appear to have influenced the submitted work.

References
