Research Paper

PAIN®

Prevalence of chronic pain with or without neuropathic characteristics in France using the capture–recapture method: a population-based study

Chouki Chenaf[a,b,c,*], Jessica Delorme[a,b,c], Noémie Delage[a,b,c], Denis Ardid[a,b,c], Alain Eschalier[a,b,c], Nicolas Authier[a,b,c]

Abstract

Capture–recapture methods are increasingly used to determine the prevalence of numerous chronic conditions but have never been used in the context of chronic pain (CP). This study sought to provide up-to-date estimates of the prevalence of people experiencing CP ± neuropathic characteristics in France using the capture–recapture method. In 2013 to 2015, 3 data sources were used: the French prescription drug database (D-list), the national hospital discharge database (H-list), and the French pain center database (P-list). Patients aged 18 years and older treated with analgesic drugs for ≥6 months (D-list) or with a diagnosis of CP ± neuropathic characteristics (H- and P-lists) were included. Two successive capture–recapture analyses were conducted, with log-linear regression for each analysis performed. A total of 63,557 and 9852 distinct cases of CP and chronic neuropathic pain were captured, respectively. The estimated prevalence of CP and chronic neuropathic pain in the adults ranged from 27.2% (95% confidence interval: 26.1-28.4) to 32.7% (26.0-43.3) and from 5.55% (2.89-19.0) to 7.30% (6.40-8.41), respectively. Most patients were female, median ages were 67 (55-80) and 63 (51-76) years for chronic and neuropathic pain, respectively. The analgesic drugs most frequently used in CP patients were paracetamol (62.1%), weak opioids (39.7%), and nonsteroidal anti-inflammatory drugs (32.7%), whereas in neuropathic pain patients, anticonvulsants (45.3%), tricyclic antidepressants (18.1%), and serotonin–norepinephrine reuptake inhibitors (13.3%) were more frequently used. This first electronic health record–based study on CP using the capture–recapture method revealed a high prevalence of CP, with a significant proportion of neuropathic pain patients.

Keywords: Chronic pain, Neuropathic pain, Electronic health record, Prevalence, Capture–recapture, Epidemiology

1. Introduction

Chronic pain (CP) is a major public health issue worldwide in terms not only of its huge impact on patient quality of life[15,23] but also of its significant economic impact on society, with both direct and indirect costs.[8,34,38] Chronic pain, a highly complex condition, is classified into nociceptive, neuropathic, and nociplastic pain and may be affected by socio-environmental or psychological determinants. The epidemiology of CP has been assessed through several studies conducted using community surveys. Published prevalence estimates of CP vary widely, ranging worldwide from 8% to over 60%.[5,8,16] In Europe, estimates derived from different settings have demonstrated that CP affects 10% to 30% of the adult population.[8,40] Discrepancies between pain prevalence rates may be partly due to how CP is defined, what severity of pain is considered, and how patients are selected. In France, only one large epidemiological study is available. Bouhassira et al. conducted a large nationwide postal survey in 2004 that included more than 20,000 patients. The prevalence of CP in the general population was estimated at 31.7%.[6]

Generalizing data from surveys represents an important issue, and representativeness and response rates should be appropriately considered because selection bias and nonresponse bias may seriously affect the validity of a survey. Moreover, no exhaustive system of CP reporting or registration is currently available in France. In this context, the capture–recapture method represents a worthy alternative and may provide more reliable estimates. The capture–recapture approach was originally developed in the fields of biology and zoology to estimate the size of a closed wild animal population.[10] Different samples of animals are captured, counted, and tagged. By calculating the proportion of tagged animals in other samples, the size of the total population can be estimated. By comparing data from several independent overlapping sources, it is possible to adjust for missing cases and to generate estimates of the prevalence of a given condition.

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

[a] *Université Clermont Auvergne, CHU Clermont-Ferrand, Inserm, Neuro-Dol, Service de Pharmacologie médicale, Centres Addictovigilance et Pharmacovigilance, Centre Evaluation et Traitement de la Douleur, Clermont-Ferrand, France,* [b] *Observatoire Français des Médicaments Antalgiques (OFMA)/French Monitoring Centre for Analgesic Drugs, Université Clermont Auvergne, Faculté de Médecine, Clermont-Ferrand, France,* [c] *Institut Analgesia, Faculté de Médecine, Clermont-Ferrand, France*

**Corresponding author. Address: Centre Hospitalier Universitaire de Clermont-Ferrand, Service de Pharmacologie Médicale, BP69, 63003 Clermont-Ferrand, France. Tel.: +33 (0)4 73 751 822; fax: +33 (0)4 73 751 823. E-mail address: chouki.chenaf@uca.fr (C. Chenaf).*

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.painjournalonline.com).

PAIN 00 (2018) 1–9

http://dx.doi.org/10.1097/j.pain.0000000000001347

This method has also been used in epidemiological studies to estimate the prevalence of a disease,[11] and it has been applied to a number of medical conditions.[9,30,37,43,47] However, to our knowledge, no epidemiological studies focusing on the prevalence of CP have used the capture–recapture approach to estimate the size of the CP population.

An updated prevalence estimate for CP in France is needed, since the most recent French data were published in 2008 on patients surveyed in 2004.[6] As CP predominantly affects older adults, and considering global population aging,[25] more up-to-date national estimates would give decision makers valuable insights to develop strategies to meet health care needs, design essential health policies, and allocate health care resources adequately. The aim of this study was to provide up-to-date national estimates of the prevalence of CP with or without neuropathic characteristics in France using the capture–recapture method.

2. Methods

The capture–recapture method, used in this work, estimates the total number of cases of a specific disease after matching cases reported in at least 2 sources. In this study, 3 data sources were used to identify people experiencing CP.

2.1. Settings

This study used electronic health care record data between January 1, 2013, and December 31, 2015, from the Echantillon Généraliste des Bénéficiaires (EGB) database, a representative 1/97th random sample of the population covered by the French national health insurance system (approximately 80% of the French population).[4] The EGB merges 3 distinct databases: (1) the drug reimbursement database, which collects all claims for prescribed and reimbursed drugs, (2) the national hospital discharge summaries database (Programme de Médicalisation des Systèmes d'Information, PMSI), which includes hospital admissions in medical, surgical, and obstetrical wards, and (3) the specialized pain centers database, which includes medical information on full or partial hospitalizations in these centers. The use of this anonymized data for medical research has been approved by the Commission Nationale de l'Informatique et des Libertés, the French data protection authority.

Capture–recapture data consisted of overlapping lists of patients in the target CP population taken from 3 administrative data sources. Each source represented the "capture" stage, ie, each source is regarded as a trapping sample. Overlaps between sources were considered analogous to overlapping "captures" of animals, and thus corresponded to the "recapture" stage, using identification number as tags or marks. The recapture information (ie, source-overlap information or source intersection) can be used to estimate the size of the unobserved population and then the total population under proper assumptions.

2.2. Capture/description of the 3 data sources

The first source was the drug reimbursement database, which comprises all claims for prescribed and reimbursed drugs dispensed in retail pharmacies (including dates dispensed and quantities supplied). Data on analgesic drugs dispensed were extracted according to the Anatomical Therapeutic Chemical classification codes. Strong opioids included morphine (Anatomical Therapeutic Chemical codes N02AA01 and N02AA51), fentanyl (N02AB03), oxycodone (N02AA05 and N02AA55), pethidine (N02AB02), and buprenorphine (N02AE01). Weak opioids included tramadol (N02AX02, N02AX52), codeine (N02AA59), dihydrocodeine (N02AA08 and N02AA58), and opium (N02AA02). Paracetamol (N02BE01), nonsteroidal anti-inflammatory drugs (NSAIDs: M01A), nefopam (N02BG06), and antimigraine drugs (N02C) were also extracted. First-line drugs for neuropathic pain according to current recommendations[19] were also included, namely anticonvulsants (gabapentin N03AX12 and pregabalin N03AX16), tricyclic antidepressants (clomipramine N06AA04 and amitriptyline N06AA09), and serotonin–norepinephrine reuptake inhibitors (SNRIs: duloxetine N06AX21 and venlafaxine N06AX16). Five-percent lidocaine-medicated plaster (N01BB02) used as a specific second-line neuropathic pain drug[19] was also extracted. As antidepressants and antiepileptic drugs are commonly used in indications other than chronic neuropathic pain, patients with a previous history of mental health or epileptic disorders were excluded based on the following *International Classification of Diseases 10th revision (ICD-10)* codes: F00 to F99 (mental and behavioral disorders), and G40 (epilepsy) or G41 (status epilepticus). Chronic pain patients were identified using a 6-month continuous treatment period with an interval of less than 35 days between 2 consecutive dispensations. The 6-month duration was chosen based on the International Association for the Study of Pain expert group proposal,[13] a choice which made sure that the patients were clearly affected by a chronic disease. The 35-day threshold was based on the fact that, in France, prescription drugs are dispensed for a maximum of 4 weeks. To more accurately detect any prescription interruptions, 1 week was added to the maximum prescription duration. The corresponding D-list was composed of patients aged 18 years and older who received 6 months of continuous treatment with analgesic drugs between 2013 and 2015.

The second data source was the national hospital discharge database. Whenever a patient is admitted to a medical, surgical, or obstetrical ward, the principal and associated diagnoses are noted and coded according to *ICD-10*. Chronic pain as a principal or associated diagnosis was identified based on the *ICD-10* codes R5210 (chronic neuropathic pain), R5218 (other chronic intractable pain), and R522 (other CP). The corresponding H-list was composed of patients aged 18 years and older who were diagnosed with CP between 2013 and 2015.

The third source was the French database on specialized pain centers. A pain center is a health care facility that focuses on the diagnosis and management of CP. Full hospitalization and day hospital data were collected, including information on main diagnosis coded according to the *ICD-10*, making it possible to distinguish chronic neuropathic pain from other CP conditions. This P-list included all patients aged 18 years and older admitted between 2013 and 2015.

Two capture–recapture analyses were successively conducted, aiming at first identifying CP patients with or without neuropathic characteristics and then at specifically identifying the subgroup of chronic neuropathic pain patients by restricting the D-list to first- and second-line drugs for neuropathic pain and the P-list and H-list to the R5210 (chronic neuropathic pain) *ICD-10* code.

2.3. Recapture/identification of common patients among sources

After data collection, cases from each list were matched using their unique, and common anonymized identifier, the Numéro d'Inscription au Répertoire (NIR), ie, the French national identification

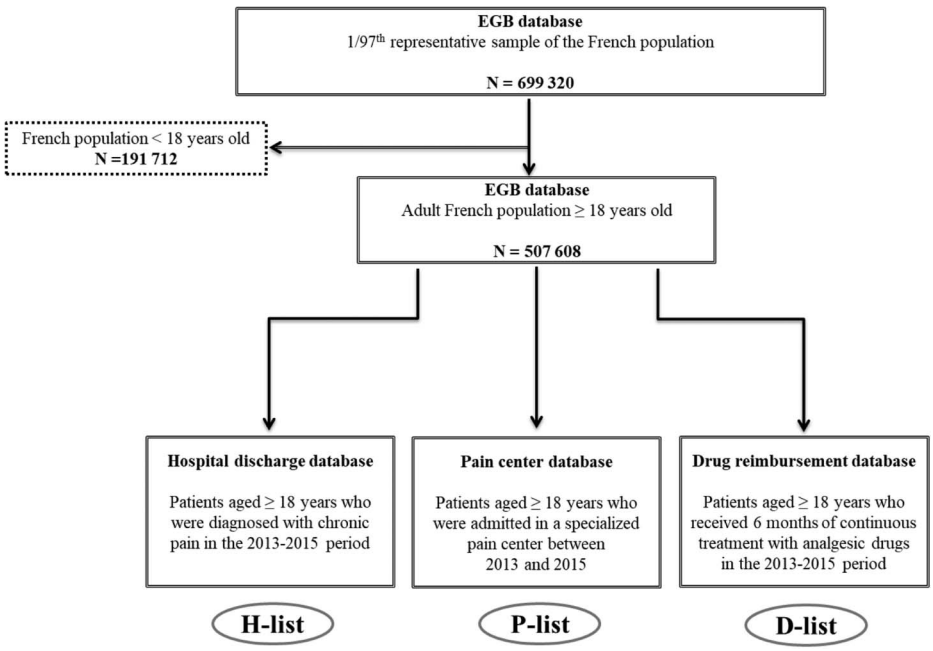

Figure 1. Flow chart of the included patients from the 3 different databases merged in the Echantillon Généraliste des Bénéficiaires (EGB) between 2013 and 2015: the hospital discharge database (H-list), the pain center database (P-list), and the drug reimbursement database (D-list).

number. This process identified cases that appeared on 1, 2, or all 3 of the lists and determined the number of overlaps.

2.4. Capture–recapture estimates

In epidemiology, the validity and the reliability of the estimates hinges on the following underlying assumptions on which the method is based: (1) A closed population, ie, there is no change during the investigation period (no births, no deaths, no immigration, or emigration). (2) Patients can be matched without error between sources, ie, the record-linkage procedure between sources should be reliable (no misclassification of records) because the accurate determination of the number of overlap cases is essential to obtain unbiased estimates. (3) The independence between sources: 2 sources are independent if the probability of a patient being reported in one source does not depend on its probability of being reported in the other source. In the context of 3 sources, the independence assumption is not crucial because it is possible to adjust for potential source dependencies. (4) A homogeneous population, ie, each patient has the same probability of being observed within sources, or, alternatively, the probability of being observed in a source does not depend on the characteristics of the patient (ie, age, sex, severity of disease…).

From the 3-source capture–recapture data, there are a number of methods to provide estimates of the number of unobserved patients, and particularly, log-linear modelling methods have been used extensively.[27,39] The log-linear method allows for determination of the missing data (total number of cases) from a 2^s contingency table (S being the total number of sources). With 3 sources ($s = 3$), there are 2^3, ie, 8 possible combinations of these sources in which cases do or do not appear (see supplementary figures 1-2, available at http://links.lww.com/PAIN/A633). A 3-source analysis was performed by fitting 8 log-linear models to the data arranged in this 2^3 contingency table. These analyses were performed using STATA's user-written "recap" program for standard 3-source capture–recapture analysis.[1]

By using log-linear methods with 3 lists, estimates are generated by 1 of 8 models, from the simplest, independence of all sources (the "independent" model), to the most complex, the presence of all 2-source interactions (the "saturated" model). In other words, 8 types of log-linear models can be identified: the "independent model" which assumes that all sources are independent (see supplementary table 1, model No. 8 [P H D], available at http://links.lww.com/PAIN/A633), 3 models that includes a 2-source interaction term (see supplementary table 1, model No. 5, 6, or 7, available at http://links.lww.com/PAIN/A633), 3 models that includes 2 terms of 2-source interaction (see supplementary table 1, model No. 2, 3, or 4, available at http://links.lww.com/PAIN/A633), and finally, a "saturated model" that incorporates all possible 2-source interactions (see supplementary table 1, model No. 1 [P H D PH PD HD], available at http://links.lww.com/PAIN/A633). Dependence between sources is incorporated by introducing interaction terms in the models.

This modeling strategy has been validated and fully described elsewhere.[27,39] Briefly, to assess how the various log-linear models fit the data, the log likelihood-ratio test, also known as G^2 or deviance, was used; the lower the value of G^2 the better is the fit of the model. Then, to select the most appropriate model, 2 information criteria were used: the Bayesian information criterion (BIC) and the Akaike information criterion (AIC). The best-fitting model was defined as the one that offered the best balance between the lowest deviance (G^2), the lowest BIC or lowest AIC, and the most parsimonious model (greatest simplicity, ie, the least saturated that includes less interaction terms).[12] Ninety-five–percent goodness-of-fit confidence intervals (95% CIs) were estimated using the likelihood ratio.

Theoretically, the saturated model represents the best model that fits the data perfectly, including all possible interactions. However, although the saturated model provides the least biased estimates, it is also associated with a large variance resulting in poorer precision of the estimates, compared with more parsimonious models. Statisticians view this principle as "a bias vs variance tradeoff." All model selection methods use some notion

Table 1

Characteristics of chronic pain patients included 2013 to 2015.

	P-list*, n = 325	H-list*, n = 2579	D-list*, n = 61,880	P H D†, n = 63,357
Age (y)				
Mean ± SD (min-max)	55.3 ± 15.0 (18-95)	63.7 ± 16.8 (18-105)	66.2 ± 16.8 (18-108)	66.1 ± 16.8 (18-108)
Median (Q1-Q3)	53 (45-66)	64 (52-78)	68 (55-80)	67 (55-80)
Age category, n (%)				
<25	6 (1.5)	27 (1.1)	629 (1.0)	634 (1.0)
25-34	17 (5.3)	109 (4.2)	2300 (3.7)	2408 (3.8)
35-49	105 (32.4)	408 (15.8)	7608 (12.3)	7920 (12.5)
50-64	106 (32.7)	774 (30.0)	15,917 (25.7)	16,346 (25.8)
65-74	50 (15.4)	461 (17.9)	13,276 (21.5)	13,558 (21.4)
≥75	41 (12.7)	800 (31.0)	22,150 (35.8)	22,492 (35.5)
Sex (n, %)				
Male	118 (36.4)	1134 (44.0)	21,848 (35.3)	22,555 (35.6)
Female	207 (63.6)	1445 (56.0)	40,032 (64.7)	40,802 (64.4)
Low-income status (n, %)	36 (11.1)	469 (18.2)	4352 (7.0)	4791 (7.6)
Cancer (n, %)	21 (6.5)	167 (6.5)	3820 (6.2)	3991 (6.3)
Diabetes (n, %)	20 (6.2)	171 (6.6)	4216 (6.8)	4335 (6.8)
Chronic pain *ICD-10* codes (n, %)				
R5210	86 (26.5)	676 (26.2)	n.a	685 (1.1)
R5218	153 (47.1)	1346 (52.2)	n.a	1431 (2.3)
R522	2 (0.6)	557 (21.6)	n.a	558 (0.9)
Analgesic drugs (n, %)‡				
Paracetamol	—	—	38,410 (62.1)	38,410 (60.6)
NSAIDs	—	—	20,251 (32.7)	20,251 (32.0)
Nefopam	—	—	550 (0.9)	550 (0.9)
Triptans	—	—	2404 (3.9)	2404 (3.8)
Weak opioids	—	—	24,553 (39.7)	24,553 (38.8)
Tramadol	—	—	13,582 (21.9)	13,582 (21.9)
Opium	—	—	7402 (12.0)	7402 (12.0)
Codeine	—	—	5294 (8.6)	5294 (8.6)
Strong opioids	—	—	3133 (5.1)	3133 (5.0)
Morphine	—	—	1137 (1.8)	1137 (1.8)
Oxycodone	—	—	1003 (1.6)	1003 (1.6)
Fentanyl	—	—	1296 (2.1)	1296 (2.1)
Antiepileptics	—	—	1303 (2.1)	1303 (2.1)
Gabapentin	—	—	359 (0.6)	359 (0.6)
Pregabalin	—	—	971 (1.6)	971 (1.6)
TCAs	—	—	830 (1.3)	830 (1.3)
SNRIs	—	—	780 (1.2)	780 (1.2)
Lidocaine plaster	—	—	77 (0.1)	77 (0.1)

"—" Represents treatments were not available in the H- and P-lists.

* Occurrences captured in each list.

† Total occurrences captured (multiple occurrences are counted once).

‡ Dispensed at least 3 times during the year of inclusion, except for paracetamol (6 dispensations) because it can be used intermittently to treat a nonspecific fever or flu-like symptoms.

n.a, not available; NSAIDs, nonsteroidal anti-inflammatory drugs; TCAs, tricyclic antidepressants; SNRIs, serotonin–norepinephrine reuptake inhibitors.

of this tradeoff (as the number of parameters in a model increases, bias decreases but variance increases). Typically, the saturated model is the starting point, ie, the default model that is used for testing of goodness-of-fit of the other models, to select the most parsimonious model that will achieve a proper tradeoff between bias and variance. We also conducted a final sensitivity analysis by examining the impact of precision of the estimates of chronic pain ± neuropathic characteristics with increasing model size.

The prevalence of CP with or without neuropathic characteristics was determined by dividing the number of patients obtained from the selected model by the total EGB population aged 18 years and older residing in France (ie, n = 507,608) from which all 3 sources were derived in the 2013 to 2015 period (**Fig. 1**). Data were expressed as frequency and associated percentage for categorical data and as mean ± SD or as median and interquartile range for quantitative data. Statistical analysis was performed using Stata 12.0 software for Windows (StataCorp LP, College Station, TX).

3. Results

3.1. Prevalence of chronic pain in the general population

Between 2013 and 2015, 61,880 CP patients were collected from the D-list, 2579 from the H-list, and 325 from the P-list (**Table 1**). Median ages were 68 (47-64), 64 (52-78), and 53 (45-66) in the D-, H-, and P-lists, respectively. Most of the patients were female, with 64.7%, 56.0%, and 63.6% women in the D-, H-, and P-lists, respectively (**Table 1**). The over-75 age group was the most represented in the H- and D-lists (31.0% and 35.8%, respectively), whereas the 50 to 64 age group was predominant in the P-list. The most common prescription medicines taken by

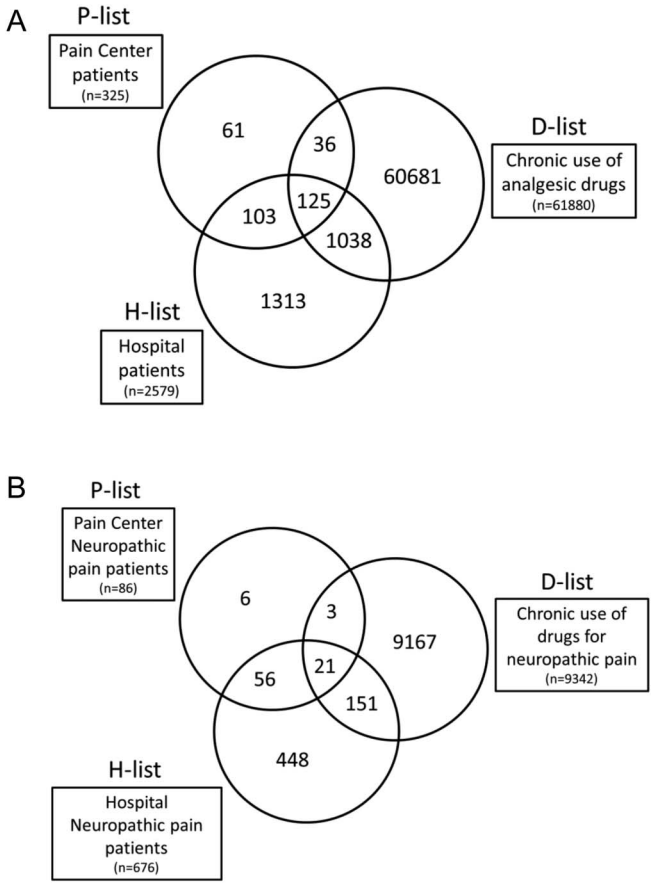

Figure 2. Number of chronic pain patients (A) and chronic neuropathic pain patients (B) in the 2013 to 2015 period matched between 3 data sources: the pain center database (P-list), hospital discharge database (H-list), and reimbursed analgesics database (D-list).

D-list patients were paracetamol (62.1%), weak opioids (39.7%), and NSAIDs (32.7%). Some 5% were taking strong opioid analgesics, and triptans were used by 3.9%. Gabapentinoid anticonvulsants, tricyclic antidepressants, and SNRIs were less frequently prescribed, at 2.1%, 1.3%, and 1.2%, respectively.

Concerning overlaps, there were 228, 1163, and 161 common cases identified between the P- and H-, H- and D-, and P- and D-lists, respectively (**Fig. 2A**). In total, 125 patients overlapped between the 3 sources. Consequently, after cross-referencing, the 3 data sources led to the capture of 63,357 cases. The median age of captured patients was 67 (55-80) years, and the most represented age group was the older than 75 years age group (35.5%). They were predominantly female (64.4%) and low-income patients accounted for 7.6%. Patients affected by cancer and diabetes accounted for 6.3% and 6.8%, respectively.

After log-linear regression, the 3 models that achieved the best balance between the lowest deviance (G^2), lowest AIC, and lowest BIC were (1) the model 2, which incorporated 2 interaction terms (between the P- and H-lists, and between the P- and D-lists) provided an estimate of 140,114 (95% CI: 134,072-146,683) CP patients, ie, a prevalence estimate of 27.6% (140,114/507,608) (95% CI: 26.4-28.9); (2) the model 3, which also included 2 interaction terms (between the P- and H-lists, and between the H- and D-lists) and provided an estimate of 166,177 (95% CI: 131,861-219,972) CP patients, ie, a prevalence estimate of 32.7% (166,177/507,608) (95% CI: 26.0-43.3); and (3) the model 5 that, compared with model 2 and 3, had a lower final estimation of the number of CP patients, 138,107 (95% CI: 132,578-144,092); the corresponding prevalence of CP was estimated at 27.2% (138,107/507,608) (95% CI: 26.1-28.4). Compared with model 2 and 3, the model 5 represented a more parsimonious model, which considered one interaction term between the P- and H-lists (**Table 2**). The sensitivity analysis showed that increasing model size was associated with a parallel decrease in precision (see supplementary table 2, available at http://links.lww.com/PAIN/A633: best precision for model 5 with ±1.1% [one interaction term] vs ±1.2% for model 2 and ±8.7% for model 3 [2 interaction terms] and vs ±16.1% for the most complex model [all possible 2-source interactions], ie, model 1).

3.2. Prevalence of chronic neuropathic pain in the general population

Between 2013 and 2015, 9342 cases were collected from the neuropathic D-list, 676 from the H-list, and 86 cases from the P-list (**Table 2**). Median ages were 63 (47-64), 62 (49-75), and 53 (51-76) years in the D-, H-, and P-lists, respectively. The 50 to 64 age group was the most represented in the P-, H-, and D-lists (38.4%, 30.3%, and 31.1%, respectively). Most of the patients were female, with 64.6%, 56.4%, and 70.9% women in the D-, H-, and P-lists, respectively (**Table 2**). The main specific neuropathic pain treatments identified from the D-list were anticonvulsants (45.3%) such as the gabapentinoids pregabalin (36.5%) and gabapentin (10.8%), tricyclic antidepressants (18.1%), SNRIs (13.3%), and lidocaine plasters (11.4%).

After data cross-referencing, there were 77, 172, and 24 common cases identified between the P- and H-, H- and D-, and P- and D-lists, respectively (**Fig. 2B**). In total, 21 patients overlapped between the 3 sources. Consequently, after cross-referencing, the 3 data sources led to the capture of 9852 cases. The median age of captured patients was 63 (51-76), and the most represented age category was the 50 to 64 age group. Captured patients were mainly female (64.1%) and low-income patients accounted for 9.3%. Patients affected by cancer and diabetes accounted for 6.3% and 6.6%, respectively.

After log-linear regression modelling, the 3 models that achieved the best balance between the lowest deviance (G^2), lowest AIC, and lowest BIC were (1) the model 2 (2 interaction terms) that provided an estimate of 37,049 (95% CI: 32,494-42,698) CP patients, corresponding to a prevalence of 7.30% (37,049/507,608) (95% CI: 6.40-8.41); (2) the model 3, which also included 2 interaction terms, provided an estimate of 28,185 (95% CI: 14,691-96,636) CP patients, ie, a prevalence estimate of 5.55% (28,185/507,608) (95% CI: 2.89-19.0); and (3) the model 5 that, compared with model 2, had lower AIC and BIC and was more parsimonious (only one interaction term, between the P- and H-lists); the estimation of the number of chronic neuropathic pain patients was also slightly lower, 36,567 (95% CI: 32,365-41,700), with a corresponding prevalence of chronic neuropathic pain estimated at 7.20% (36,567/507,608) (95% CI: 6.38-8.22). The sensitivity analysis showed that increasing model size was associated with a parallel decrease in precision (see supplementary table 3, available at http://links.lww.com/PAIN/A633: best precision for model 5 with ±0.92% [one interaction term] vs ±1.0% for the model 2 and ±8.1% for model 3 [2 interaction terms] and vs ±9.8% for the most complex model [all possible 2-source interactions] ie, model 1).

Table 2

Characteristics of chronic neuropathic pain patients included 2013 to 2015.

	P-list*, n = 86	H-list*, n = 676	D-list*, n = 9342	P H D†, n = 9852
Age (y)				
Mean ± SD (min-max)	55.8 ± 14.3 (20-91)	61.5 ± 16.8 (19-100)	63.0 ± 16.2 (18-106)	62.9 ± 16.2 (18-106)
Median (Q1-Q3)	53 (47-64)	62 (49-75)	63 (51-76)	63 (51-76)
Age category, n (%)				
<25	1 (1.2)	5 (0.7)	56 (0.6)	61 (0.6)
25-34	2 (2.3)	36 (5.3)	329 (3.5)	358 (3.6)
35-49	29 (33.7)	133 (19.7)	1628 (17.4)	1728 (17.5)
50-64	33 (38.4)	205 (30.3)	2908 (31.1)	3072 (31.2)
65-74	8 (9.3)	120 (17.8)	1843 (19.7)	1927 (19.6)
≥75	13 (15.1)	177 (26.2)	2578 (27.6)	2706 (27.5)
Sex (n, %)				
Male	25 (29.1)	295 (43.6)	3310 (35.4)	3540 (35.9)
Female	61 (70.9)	381 (56.4)	6032 (64.6)	6312 (64.1)
Low-income status (n, %)	12 (14.0)	85 (12.6)	848 (9.1)	913 (9.3)
Cancer (n, %)	9 (10.5)	39 (5.8)	595 (6.4)	618 (6.3)
Diabetes (n, %)	4 (4.7)	43 (6.4)	622 (6.7)	654 (6.6)
Chronic pain *ICD-10* codes (n, %)				
R5210	86 (100)	676 (100)	n.a	685 (7.0)
R5218	0	0	n.a	0
R522	0	0	n.a	0
Analgesic drugs (n, %)‡				
Antiepileptics	—	—	4465 (47.8)	4465 (45.3)
Gabapentin	—	—	1065 (11.4)	1065 (10.8)
Pregabalin	—	—	3597 (38.5)	3597 (36.5)
TCAs	—	—	1784 (19.1)	1784 (18.1)
SNRIs	—	—	1308 (14.0)	1308 (13.3)
Lidocaine plaster	—	—	1121 (12.0)	1121 (11.4)

"—"Represents treatments were not available in the H- and P-lists.

* Occurrences captured in each list.

† Total occurrences captured (multiple occurrences are counted once).

‡ Dispensed at least 3 times during the year of inclusion.

n.a, not available; TCAs, tricyclic antidepressants; SNRIs, serotonin–norepinephrine reuptake inhibitors.

4. Discussion

To our knowledge, this is the first epidemiological study that provides estimates of the prevalence of people experiencing CP with or without neuropathic characteristics in the general adult (18 years and older) population based on the capture–recapture method using electronic health records. The estimated prevalence of CP was 27.2% (95% CI: 26.1-28.4), ranging from 27.2% to 32.7% (26.0-43.3) and that of CP with neuropathic characteristics was 7.20% (95% CI: 6.38-8.22), ranging from 5.55% (2.89-19.0) to 7.30% (6.40-8.41).

Comparing our results with another population-based survey conducted in France, our prevalence estimate is somewhat lower than that reported by Bouhassira et al.[6] in 2008. Through a population-based survey, they observed that 31.7% (95% CI: 31.1-32.3) of a nationally representative sample of adults reported chronic daily pain. They defined CP as lasting at least 3 months, whereas we used a more restrictive definition of 6 months or more. A European-wide study took a similar pain duration of 6 months and reported a prevalence of 19%.[8] Prevalence estimates of other studies using the 6-month duration of pain have varied from 12% to 49%, irrespective of pain frequency or intensity.[2,7,14,16,17,29,32,33,35,36,42,45] In fact, there has been considerable inconsistency in the choice of pain duration across the studies, as stated by a recent systematic review,[46] which included 86 studies of CP prevalence. Some 68% of the studies reviewed used a pain duration of >3 months or ≥3 months, whereas 18% used a pain duration of ≥6 months, and 11% did not include any duration. Overall, the published prevalence estimates of CP varied greatly in that review, ranging from 8.7% to 64.4%. This was at least partly due to differences in the survey methodology or chosen definition of CP. That systematic review and meta-analysis[46] tried to comprehensively address these methodological issues. Unfortunately, the authors were unable to elucidate the effect of the definition of CP on prevalence estimates and concluded that even "after controlling for many of the potential methodological factors, the studies remained highly heterogeneous." Thus, there has been little consensus on the duration of pain, and an International Association for the Study of Pain expert group has suggested for research purposes that a 6-month duration might be better than 3 months.[13]

The prevalence of chronic neuropathic pain in our study is within the expected ranges previously published in the general population, ranging from 0.9% to 17.9%, depending on the population and methodology used.[6,20,22,24,26,35,48,49,51] Of these studies, although Bouhassira et al.,[6] for instance, found a 6.9% prevalence, very close to ours, 2 others, identifying neuropathic cases from electronic health record databases,[20,22] reported low overall prevalence estimates of around 1%. These low estimates were based only on *ICD-9* diagnoses associated with CP, so the true prevalence was probably underestimated because it was assumed that only patients who presented to a primary care facility could be diagnosed and identified.[26] In a more recent

Table 3

Log-linear models and estimates of the number of chronic pain patients in France 2013-2015.

Chronic pain patients		N	95% CI		df	G^2	AIC	BIC
No.	Model description		Low	High				
1	P H D PH PD HD	221,198	160,100	323,784	0	0	0	0
2	P H D PH PD	140,114	134,072	146,683	1	8.62	6.62	6.67
3	P H D PH HD	166,177	131,861	219,972	1	9.51	7.51	7.56
4	P H D PD HD	64,134	63,912	64,432	1	837.34	835.34	835.39
5	P H D PH	138,107	132,578	144,092	2	11.95	7.95	8.06
6	P H D PD	126,391	121,666	131,505	2	1433.37	1429.37	1429.48
7	P H D HD	77,921	74,283	82,459	2	1281.29	1277.29	1277.39
8	P H D	126,227	121,765	131,029	3	1433.41	1427.41	1427.57

AIC, Akaike information criterion; BIC, Bayesian information criterion; P, H, or D are the 3 lists taken alone: the P-list (pain center patients), H-list (hospital patients), or D-list (chronic use of analgesic drugs >6 months), respectively; PH, PD, or HD are the interactions between the different lists; CI, confidence interval; df, degrees of freedom; G^2, goodness-of-fit test.

electronic health record–based study,[44] both diagnosis and medication were extracted from a primary care electronic health record database, which led to the identification of more neuropathic pain cases compared with the previous 2 studies[20,22]: prevalence estimates ranged from 1.5% (for certain neuropathic pain) to 11.2% (for certain + probable neuropathic pain).

In our study, patients with CP were older and more likely to be female, as in previous findings. According to our patients' demographics, almost two-thirds of CP sufferers were women and one-third of CP patients were aged 75 years or older. Chronic neuropathic pain was also more prevalent in women (64.1% vs 35.9% in men) and increased with age, peaking at 50 to 64 years. These findings are consistent with the known demographic profile of CP patients with or without neuropathic characteristics from population surveys[3,6,8,18,41,50], and so, they provide an external validation of our study. Furthermore, the analgesic drug profile of our CP patients is qualitatively quite similar to the results of the European-wide study by Breivik et al.[8] in which the most commonly prescribed drugs for CP were NSAIDs (44%), weak opioid analgesics (23%), and paracetamol (18%). However, our findings showed that paracetamol was more frequently used than NSAIDs (ratio 1.9: 62.1% vs 32.7%). This is in line with the data on France in the European-wide survey (ratio 1.5: NSAIDs 25% vs paracetamol 38%), whereas NSAIDs were much more widely used in Italy (75%), Austria (62%), and Germany (62%) than paracetamol (6%, 2%, and 4%, respectively).[8] Conversely, concerning strong opioid analgesics, our findings were similar, with 5% of patients taking a strong opioid. Antiepileptics and tricyclic antidepressants were not commonly found in our study (2.1% and 2.5%, respectively), which is also consistent with Breivik et al. (2% and 3%, respectively). However, in neuropathic pain patients, anticonvulsants (45.3%), tricyclic antidepressants (18.1%), and SNRIs (13.3%) were the most frequently used drugs, in line with the updated NeuPSIG recommendations[19]; lidocaine patches, a second-line drug, being used in 11.4% in this study.

4.1. Strengths and limitations

The capture–recapture method avoids the problems of selection bias and generalization to the whole population, which arise from survey methods such as questionnaires and telephone or face-to-face interviews. Interestingly, Hook and Regal[28] have even suggested the use of the capture–recapture technique for supposed exhaustive surveys. Estimation of the prevalence of CP by capture–recapture methods is also particularly appropriate, since the care of patients is frequently split between hospitals, pain centers, and primary care facilities. Although the capture–recapture approach is a simple and attractive statistical approach for estimating the size of hard-to-reach populations, the results must be interpreted with caution because the validity of capture–recapture estimates depends on potential violations of certain underlying conditions:

Table 4

Log-linear models and estimates of the number of chronic neuropathic pain patients in France 2013-2015.

Chronic neuropathic pain patients		N	95% CI		df	G^2	AIC	BIC
No.	Model description		Low	High				
1	P H D PH PD HD	30,250	14,696	114,161	0	0	0	0
2	P H D PH PD	37,049	32,494	42,698	1	0.14	−1.86	−1.80
3	P H D PH HD	28,185	14,691	96,636	1	0.15	−1.85	−1.79
4	P H D PD HD	9900	9868	9958	1	152.42	150.42	150.48
5	P H D PH	36,567	32,365	41,700	2	0.81	−3.57	−3.47
6	P H D PD	27,882	25,214	31,068	2	544.52	540.52	540.63
7	P H D HD	10,584	10,135	11,397	2	395.97	391.97	392.08
8	P H D	28,403	25,790	31,496	3	545.64	539.64	539.80

AIC, Akaike information criterion; BIC, Bayesian information criterion; P, H, or D are the 3 lists taken alone: the P-list (pain center patients), H-list (hospital patients), or D-list (chronic use of neuropathic pain drugs >6 months), respectively; PH, PD, or HD are the interactions between the different lists; CI, confidence interval; df, degrees of freedom; G^2, goodness-of-fit test.

(1) The identification of common cases across sources is essential. The failure of this assumption could introduce a bias, resulting in either overestimation or underestimation of the estimates. This was probably not an issue in our study; it was indeed possible to reliably match members of the population appearing on different lists because they were uniquely identified by their NIR, the French national identification number. Overall, this unique identifier permitted perfect record-linkage, assuming no misclassifications of records.
(2) The closed population assumption for our 3-year duration of sampling is difficult to fully achieve because additions (births or immigration) or deletions (deaths or immigration) cannot be totally ruled out. However, to relax this assumption, the sources had the same geographic coverage and the same time frame, such that the birth, death, and migrations rates were assumed negligible. Moreover, it has been shown that prevalence estimates were robust to violations of the closure assumption, although precision decreased.[31]
(3) The traditional assumption of truly independent data sources may also be difficult to achieve. Nevertheless, log-linear models have been proposed to take account of dependence between sources, and in our study, both of the final models chosen encompassed interaction between the P- and H-lists (model 5, **Tables 3 and 4**). Introducing a second interaction term in the final models (model 2, **Tables 3 and 4**) did not change much the population size estimates.
(4) The homogeneity assumption: indeed, whether a patient is identified by a particular source may depend on several covariates such as age, sex, or severity of disease. Violation of this assumption can be handled by stratification of the population into more homogeneous subgroups, perform capture–recapture analysis for each of the distinct strata, and then add the results for the total estimate. To improve the model by reducing heterogeneity and the bias in the estimates, stratified analyses were performed according to age and sex variables: interestingly, the total population size estimate did not change much (see supplementary figure 3, available at http://links.lww.com/PAIN/A633). Stratifications on other important covariates such as severity of disease or treating physician were not possible because these variables were not available in the 3 sources.

In the context of this study, it can reasonably assumed that the assumptions (1) and (2) hold. However, in epidemiology, it is likely that the assumptions (3) and (4) may be violated, but we mentioned approaches that relaxed these assumptions: the effects of heterogeneity could be partly reduced by stratification based on available covariate information, and dependence between sources was taken into account by introducing interaction terms in the final log-linear model.

Overall, the use of 3 samples is a major strength since when 3 samples are cross-referenced, as in our study, data show that the usual biases associated with the capture–recapture method can be limited[21]: indeed, log-linear modelling allows for controlling and reducing (at least partially) dependence and heterogeneity, making these 3-source capture–recapture models more powerful.

5. Conclusion

This first epidemiological study on CP using the capture–recapture method revealed a high prevalence of CP, affecting more than one-quarter of the French adult population while suggesting a considerable burden of CP. Using electronic health record data represents an essential source of public health surveillance, and here, we show that the capture–recapture method yields accurate estimates of the prevalence of CP, offering a valuable, more straightforward, and cheaper method of achieving up-to-date estimates than conventional surveys.

Conflict of interest statement

None of the authors have any conflicts of interest to disclose.

This study was not funded by any sponsor.

Acknowledgments

Author contributions: C. Chenaf, A. Eschalier, and N. Authier developed the concept and devised the study. C. Chenaf had full access to all of the study data and assumes responsibility for data integrity and the accuracy of the data analysis. The data were analyzed by C. Chenaf. The manuscript was written by C. Chenaf, J. Delorme, A. Eschalier, and N. Authier. The manuscript was revised by N. Delage, D. Ardid, A. Eschalier, and N. Authier.

Appendix A. Supplemental digital content

Supplemental digital content associated with this article can be found online at http://links.lww.com/PAIN/A633.

Article history:
Received 26 February 2018
Received in revised form 25 June 2018
Accepted 11 July 2018
Available online 19 July 2018

References

[1] An Der Heiden M. RECAP: Stata module to perform capture-recapture analysis for three sources with goodness-of-fit based confidence intervals. Available at: https://ideas.repec.org/c/boc/bocode/s456859.html. Accessed April 30, 2018.
[2] Andersson HI, Ejlertsson G, Leden I, Rosenberg C. Chronic pain in a geographically defined general population: studies of differences in age, gender, social class, and pain localization. Clin J Pain 1993;9:174–82.
[3] Bartley EJ, Fillingim RB. Sex differences in pain: a brief review of clinical and experimental findings. Br J Anaesth 2013;111:52–8.
[4] Bezin J, Duong M, Lassalle R, Droz C, Pariente A, Blin P, Moore N. The national healthcare system claims databases in France, SNIIRAM and EGB: powerful tools for pharmacoepidemiology. Pharmacoepidemiol Drug Saf 2017;26:954–62.
[5] Blyth FM, March LM, Brnabic AJ, Jorm LR, Williamson M, Cousins MJ. Chronic pain in Australia: a prevalence study. PAIN 2001;89:127–34.
[6] Bouhassira D, Lantéri-Minet M, Attal N, Laurent B, Touboul C. Prevalence of chronic pain with neuropathic characteristics in the general population. PAIN 2008;136:380–7.
[7] Brattberg G, Thorslund M, Wikman A. The prevalence of pain in a general population. The results of a postal survey in a county of Sweden. PAIN 1989;37:215–22.
[8] Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. Eur J Pain 2006;10:287–333.
[9] Cameron CM, Coppell KJ, Fletcher DJ, Sharples KJ. Capture-recapture using multiple data sources: estimating the prevalence of diabetes. Aust N Z J Public Health 2012;36:223–8.
[10] Capture-recapture and multiple-record systems estimation I: history and theoretical development. International working group for disease monitoring and forecasting. Am J Epidemiol 1995;142:1047–58.
[11] Chao A, Tsay PK, Lin SH, Shau WY, Chao DY. The applications of capture-recapture models to epidemiological data. Stat Med 2001;20:3123–57.
[12] Chen MH, Huang L, Ibrahim JG, Kim S. Bayesian variable selection and computation for generalized linear models with conjugate priors. Bayesian Anal 2008;3:585–614.
[13] Classification of chronic pain. Descriptions of chronic pain syndromes and definitions of pain terms/prepared by the International Association for

the Study of Pain, Task Force on Taxonomy. 1994. Available at: http://www.iasp-pain.org/PublicationsNews/Content.aspx?ItemNumber=1673. Accessed April 30, 2018.
[14] Dominick CH, Blyth FM, Nicholas MK. Unpacking the burden: understanding the relationships between chronic pain and comorbidity in the general population. PAIN 2012;153:293–304.
[15] Duenas M, Ojeda B, Salazar A, Mico JA, Failde I. A review of chronic pain impact on patients, their social environment and the health care system. J Pain Res 2016;9:457–67.
[16] Elliott AM, Smith BH, Penny KI, Smith WC, Chambers WA. The epidemiology of chronic pain in the community. Lancet 1999;354:1248–52.
[17] Eriksen J, Jensen MK, Sjøgren P, Ekholm O, Rasmussen NK. Epidemiology of chronic non-malignant pain in Denmark. PAIN 2003;106:221–8.
[18] Fillingim RB, King CD, Ribeiro-Dasilva MC, Rahim-Williams B, Riley JL. Sex, gender, and pain: a review of recent clinical and experimental findings. J Pain 2009;10:447–85.
[19] Finnerup NB, Attal N, Haroutounian S, McNicol E, Baron R, Dworkin RH, Gilron I, Haanpää M, Hansson P, Jensen TS, Kamerman PR, Lund K, Moore A, Raja SN, Rice ASC, Rowbotham M, Sena E, Siddall P, Smith BH, Wallace M. Pharmacotherapy for neuropathic pain in adults: a systematic review and meta-analysis. Lancet Neurol 2015;14:162–73.
[20] Gajria C, Murray J, Birger R, Banarsee R, Bennett DL, Tan K, Field M, Rice AS, Majeed A. Identification of patients with neuropathic pain using electronic primary care records. Inform Prim Care 2011;19:83–90.
[21] Gallay A, Nardone A, Vaillant V, Desenclos JC. The capture-recapture applied to epidemiology: principles, limits and application [in French]. Rev Epidemiol Sante Publique 2002;50:219–32.
[22] Gore M, Dukes E, Rowbotham DJ, Tai K-S, Leslie D. Clinical characteristics and pain management among patients with painful peripheral neuropathic disorders in general practice settings. Eur J Pain 2007;11:652–64.
[23] Gureje O, Von Korff M, Simon GE, Gater R. Persistent pain and well-being: a world health organization study in primary care. JAMA 1998;280:147–51.
[24] Gustorff B, Dorner T, Likar R, Grisold W, Lawrence K, Schwarz F, Rieder A. Prevalence of self-reported neuropathic pain and impact on quality of life: a prospective representative survey. Acta Anaesthesiol Scand 2008;52:132–6.
[25] He W, Goodkind D, Kowal P. An aging world: 2015. Washington, DC: U.S. Government Publishing Office, 2016. Available at: https://www.census.gov/content/dam/Census/library/publications/2016/demo/p95-16-1.pdf. Accessed April 24, 2018.
[26] van Hecke O, Austin SK, Khan RA, Smith BH, Torrance N. Neuropathic pain in the general population: a systematic review of epidemiological studies. PAIN 2014;155:654–62.
[27] Hook E, Regal R. Capture-recapture methods. Lancet 1992;339:742.
[28] Hook EB, Regal RR. The value of capture-recapture methods even for apparent exhaustive surveys. The need for adjustment for source of ascertainment intersection in attempted complete prevalence studies. Am J Epidemiol 1992;135:1060–7.
[29] Johannes CB, Le TK, Zhou X, Johnston JA, Dworkin RH. The prevalence of chronic pain in United States adults: results of an Internet-based survey. J Pain 2010;11:1230–9.
[30] Jouanjus E, Pourcel L, Saivin S, Molinier L, Lapeyre-Mestre M. Use of multiple sources and capture-recapture method to estimate the frequency of hospitalizations related to drug abuse: capture-recapture estimates of drug abuse disorders. Pharmacoepidemiol Drug Saf 2012;21:733–41.
[31] Kendall WL. Robustness of closed capture–recapture methods to violations of the closure assumption. Ecology 1999;80:2517–25.
[32] Kurita GP, Sjøgren P, Juel K, Højsted J, Ekholm O. The burden of chronic pain: a cross-sectional survey focussing on diseases, immigration, and opioid use. PAIN 2012;153:2332–8.
[33] Landmark T, Romundstad P, Borchgrevink PC, Kaasa S, Dale O. Associations between recreational exercise and chronic pain in the general population: evidence from the HUNT 3 study. PAIN 2011;152:2241–7.
[34] McDermott AM, Toelle TR, Rowbotham DJ, Schaefer CP, Dukes EM. The burden of neuropathic pain: results from a cross-sectional survey. Eur J Pain 2006;10:127.
[35] de Moraes Vieira EB, Garcia JBS, da Silva AAM, Mualem Araújo RLT, Jansen RCS. Prevalence, characteristics, and factors associated with chronic pain with and without neuropathic characteristics in São Luís, Brazil. J Pain Symptom Manage 2012;44:239–51.
[36] Nakamura M, Nishiwaki Y, Ushida T, Toyama Y. Prevalence and characteristics of chronic musculoskeletal pain in Japan. J Orthop Sci 2011;16:424–32.
[37] O'Callaghan FJ, Shiell AW, Osborne JP, Martyn CN. Prevalence of tuberous sclerosis estimated by capture-recapture analysis. Lancet 1998;351:1490.
[38] Phillips CJ. The cost and burden of chronic pain. Br J Pain 2009;3:2–5.
[39] Regal RR, Hook EB. Goodness-of-fit based confidence intervals for estimates of the size of a closed population. Stat Med 1984;3:287–91.
[40] Reid KJ, Harker J, Bala MM, Truyers C, Kellen E, Bekkering GE, Kleijnen J. Epidemiology of chronic non-cancer pain in Europe: narrative review of prevalence, pain treatments and pain impact. Curr Med Res Opin 2011;27:449–62.
[41] Rovner GS, Sunnerhagen KS, Björkdahl A, Gerdle B, Börsbo B, Johansson F, Gillanders D. Chronic pain and sex-differences; women accept and move, while men feel blue. PLoS One 2017;12:e0175737.
[42] Schopflocher D, Taenzer P, Jovey R. The prevalence of chronic pain in Canada. Pain Res Manag 2011;16:445–50.
[43] Scublinsky D, Venarotti H, Citera G, Messina OD, Scheines E, Rillo O, Arturi A, Hofman J, Somma LF, Casado G, Iannantuono RF, Gonzalez CD. The prevalence of rheumatoid arthritis in Argentina: a capture-recapture study in a city of Buenos Aires province. J Clin Rheumatol 2010;16:317–21.
[44] Shadd JD, Ryan BL, Maddocks HL, McKay SD, Moulin DE. Neuropathic pain in a primary care electronic health record database. Eur J Pain 2015;19:715–21.
[45] Sjøgren P, Ekholm O, Peuckmann V, Grønbaek M. Epidemiology of chronic pain in Denmark: an update. Eur J Pain 2009;13:287–92.
[46] Steingrímsdóttir ÓA, Landmark T, Macfarlane GJ, Nielsen CS. Defining chronic pain in epidemiological studies—a systematic review and meta-analysis. PAIN 2017;158:2092–2107.
[47] Törner A, Stokkeland K, Svensson Å, Dickman PW, Hultcrantz R, Montgomery S, Duberg AS. The underreporting of hepatocellular carcinoma to the cancer register and a log-linear model to estimate a more correct incidence. Hepatology 2017;65:885–92.
[48] Torrance N, Smith BH, Bennett MI, Lee AJ. The epidemiology of chronic pain of predominantly neuropathic origin. Results from a general population survey. J Pain 2006;7:281–9.
[49] Toth C, Lander J, Wiebe S. The prevalence and impact of chronic pain with neuropathic pain symptoms in the general population. Pain Med 2009;10:918–29.
[50] Tsang A, Von Korff M, Lee S, Alonso J, Karam E, Angermeyer MC, Borges GLG, Bromet EJ, Demytteneare K, de Girolamo G, de Graaf R, Gureje O, Lepine JP, Haro JM, Levinson D, Oakley Browne MA, Posada-Villa J, Seedat S, Watanabe M. Common chronic pain conditions in developed and developing countries: gender and age differences and comorbidity with depression-anxiety disorders. J Pain 2008;9:883–91.
[51] Yawn BP, Wollan PC, Weingarten TN, Watson JC, Hooten WM, Melton LJ. The prevalence of neuropathic pain: clinical evaluation compared with screening tools in a community population. Pain Med 2009;10:586–93.