Methodological Considerations in the Assessment of Effectiveness of Homeopathic Care: A Critical Review of the EPI3 Study

Yola Moride[1,2,3]

[1] Faculty of Pharmacy, Université de Montréal, Montreal, Quebec, Canada
[2] Center for Pharmacoepidemiology and Treatment Science, Rutgers, The State University of New Jersey, New Brunswick, New Jersey, United States
[3] YolaRX Consultants, Montreal, Quebec, Canada

Address for correspondence Yola Moride, PhD FISPE, Faculty of Pharmacy, Université de Montréal, CP 6128, succ. Centre-ville, Montreal, QC, H3C 3J7, Canada (e-mail: yola.moride@umontreal.ca).

Abstract

Background   EPI3 is an observational study of a representative sample of general practitioners (GPs) and patients in France, demonstrating that patient characteristics differ according to the prescribing preferences of their GPs for homeopathy. For selected conditions (musculoskeletal disorders, sleep disorders, anxiety/depression, upper respiratory tract infections), progression of symptoms and adverse events over follow-up in the homeopathy preference group did not significantly differ from other practice preferences, but there was a two-fold to four-fold lower usage of conventional medicines. The EPI3 study's validity was challenged due to absence of head-to-head comparison of medicines to conclude on a causal association between homeopathy and outcomes.

Methods   A critical review of the nine EPI3 publications was conducted, focusing on generalizability, selection bias, outcome measurements and confounding.

Results   The conceptual framework of EPI3 rests on a systemic construct, i.e., the homeopathic treatment concept assessed using the type of GP prescribing preference, taking into account the clinical, human and social aspects. The enrollment process enhanced the generalizability of findings. Validated instruments for outcome measurements were used for three conditions, and control of confounding was rigorous.

Conclusion   EPI3 was conducted according to best practices. Homeopathy prescribing preference met specific patient needs with less use of conventional medicines and without an apparent loss in therapeutic opportunity.

Keywords
► homeopathy
► cohort study
► care pathway
► pharmacoepidemiology
► observational study

Introduction

EPI3 ("Etude Pharmacoépidémiologique de l'Impact de santé publique des modes de prise en charge pour 3 groupes de pathologies") is a nationwide observational study that was conducted in France between March 2007 and July 2008 in the community-based general practice setting. Using a cross-sectional design, the study aimed at comparing the characteristics of 6,379 patients who regularly attended three types of general practice, defined according to prescribing preferences for homeopathy (homeopathy; strictly conventional; mixed).[1] The 804 participating general practitioners (GPs) were a representative sample of the primary care practice in France, and patients enrolled were those seen at the selected

GP practices on a random practice day (maximum 15 patients per site). Study patients who consulted their GP for either musculoskeletal disorders ($n = 1,153$), sleep disorder ($n = 346$), anxiety/depression ($n = 710$) or upper respiratory tract infections (URTIs) ($n = 518$) were enrolled in a prospective cohort study, with a follow-up of 12 months (with contacts at baseline, 1 month, 3 months and 12 months), that aimed at comparing the three prescribing preference groups on the basis of clinical and symptomatic outcomes, patient-reported outcomes (quality of life), drug utilization and adverse events that indicated loss of therapeutic opportunity. The study found slight differences in the sociodemographic characteristics of patients who are followed in the homeopathy preference group, consisting of a greater proportion of females and individuals with higher education, but no differences in co-morbidities and quality of life.[1,2] However, they had markedly healthier lifestyle and positive attitude toward complementary and alternative medicines.

In each of the four disease cohorts, patients followed in the homeopathy preference practices used fewer conventional medicines than those followed in the conventional practices, ranging from a more than two-fold decrease in the use of antibiotics in the URTI cohort[3] and in the use of NSAIDs in the musculoskeletal disease cohort[4] to a four-fold decrease in psychotropic drug use in the sleep disorder[5] and anxiety/depression[6] cohorts. Differences between practice preferences regarding symptoms improvement or progression were not statistically significant. Criticisms were raised by the French Health Authority (Haute Autorité de Santé [HAS])[7] regarding the validity of the EPI3 study: namely the absence of a head-to-head comparison of medicines to conclude on a causal association between homeopathy and outcomes, as well as the non-comparability of patients followed in the different types of practice preferences.

The present Commentary briefly addresses the following: what evidence has been generated by the EPI3 study with regard to the study findings, the quality of evidence including methodological robustness, and the generalizability of findings?

Methods

A critical review of the methodological quality of the EPI3 study was conducted. Sources of information were the nine papers published in the literature by the co-investigators. Three follow-up studies consisting of economic evaluations were not considered in the evaluation[8–10] given that they involved secondary analyses of existing data, were not planned *a priori*, and were not conducted by the original study investigators. The review was conducted against the methodological criteria described in the Guidelines for Good Pharmacoepidemiology Practices (version 4, 2015),[11] the ENCePP Guide on Methodological Standards in Pharmacoepidemiology (Version 8, 2020),[12] and the Joanna Briggs Institute (JBI) Critical Appraisal Tools.[13] More specifically, the following four dimensions of study quality were assessed: (1) generalizability of study findings to the French

population; (2) validity of the selection process for the enrollment of GPs and patients in the study; (3) quality of data collected, including validity of instruments to measure changes in clinical and humanistic outcomes over time; (4) analytical methods applied to take into account the non-comparability of patients followed in the various prescribing preference practices (homeopathy, conventional, mixed).

Results

Conceptual Framework

Health care practice needs to be based on robust theory that guides efforts in practice and research. Evaluation is key to inform best practices, using proximal and/or distal outcomes.[14] Although not explicitly stated in any of the EPI3 papers, the study reports on an evaluation conducted after the GPs selected and implemented their prescribing preference. The main construct of the conceptual framework that was evaluated was the effectiveness of the homeopathic treatment concept assessed using the type of GP prescribing preference (homeopathy, strictly conventional, or mixed), taking into account the clinical, human and social aspects. Proximal outcome consisted of conventional drug usage during follow-up, whilst distal outcomes included symptoms progression, humanistic outcomes (quality of life) and loss of therapeutic opportunity. It is important to mention that the framework does not include the evaluation of the efficacy of the homeopathic medicinal product compared to conventional medicine as one of the constructs; hence, the absence of head-to-head comparison between products. A more explicit description of the conceptual framework of the EPI3 study in the original papers would have been helpful.

Generalizability of Study Findings

The quality of evidence generated by the cross-sectional study on patient characteristics according to type of prescribing preference relies on the appropriateness of the sampling strategy. In EPI3, the sampling unit was the GP and the unit of observation was the patient. A two-stage sampling process was used whereby, first, a random sample of GPs stratified by prescribing preference (homeopathy-certified vs. non-homeopathy-certified) was drawn from the French national directory of physicians. The sample was stratified according to practice preference (homeopathy, strictly conventional, mixed). Second, a random practice day was selected for each sampled GP in order to enroll patients who attended the clinic on that day (up to a maximum of 15 per GP). Unlike many population samples based on primary data collection, the EPI3 study benefited from the availability of the French national directory of physicians, which provided a well-defined roster of practicing GPs in the country, with selected characteristics such as region and practice type (homeopathy vs. no homeopathy). Furthermore, participating patients were compared to non-participating eligible patients on the basis of gender, age, length of time attending the GP practice, type of health insurance and main reason for consultation, which allowed for the calibration of the patient sample. Results on the

characteristics of patients attending the three types of practices based on their prescribing preferences for homeopathy (homeopathy, strictly conventional, mixed) were thus weighted according to patient and GP distribution, using CALMAR (a well-known procedure in demographic studies designed to enhance the representativeness of results to the source population).[15] Generalizability of findings was thus an important strength of the EPI3 study because it aimed at describing practice and patient characteristics.

Selection Process

Selection bias is a threat to observational studies conducted using primary data collection as it is necessary to sample and contact participants. As mentioned in the previous section, great care was taken in the EPI3 study to obtain a representative sample of GPs and patients. In addition, at the time of patient enrollment, broad objectives were presented to potential participants without disclosing the study hypotheses; this methodological feature was important in order to avoid influencing participation and the self-reported information. Furthermore, unique to the EPI3 study was the creation of a registry of patients who declined participation based on data provided by the GP (patient age, gender, co-morbidities, reason for visiting the GP on that day). This allowed for the comparison of participants with non-participants as an attempt to rule out a selection bias. Attrition bias in a cohort study is a threat to validity, given that participants lost to follow-up may differ from those who remain in the risk set.[16] It would have been relevant to compare the baseline characteristics of patients lost to follow-up at each contact point (1, 3, 12 months) with those who were still in the cohort, in order to determine whether loss to follow-up was random or selective as it may affect both the absolute and the relative measures. Nevertheless, it is important to highlight that 80% of the patients initially enrolled completed the 12-month follow-up, which is within the expected range based on a meta-analysis of attrition rate in cohort studies with primary data collection,[17] and there was no difference in the proportion of loss to follow-up between the various GP prescribing preference types, which was the main exposure of interest.

Data Collection Tools

The EPI3 study was based on primary data collection, using questionnaires to GPs and patients. Across the four cohort studies, two types of data were collected from patients: 1) self-reported information collected at enrollment (i.e., socio-demographics, education, smoking, body mass index, employment status, lifestyle, history of hospitalization in the previous 12 months, whether the GP was the regular physician or not, number of GP consultations in the past year, quality of life [SF-12], attitude and beliefs regarding complementary and alternative medicines [CAMBI score]); and 2) self-reported data at each follow-up (history of drug utilization using the standardized Progressive Assisted Backward Active Recall [PABAR]). One of the advantages of patient interviews is the availability of information on over-the-counter drug usage. This instrument has been validated

against prescription data only for cardiovascular drugs[18] and drugs used for musculoskeletal disorders,[19] the latter being one of the cohorts of interest in EPI3. Because of over-the-counter and use on an as-needed basis, agreement between self-report and physician's prescription was shown to be lower than for cardiovascular drugs.[19] In addition, for each cohort, disease-specific instruments were used to enroll patients at baseline and monitor disease progression and symptoms over time as study outcomes. The robustness of the outcome measures varied between the cohorts, and the most robust were those used in the musculoskeletal disease cohort, whereby functional status was measured using the Roland-Morris score (back pain),[20] Quick Dash (upper limbs),[21] and Lequesne (lower limbs).[22] Scores were standardized to 100 and improvement was defined by an increase of 12.5 points relative to baseline, this threshold having been validated in the literature.[23]

In the anxiety/depression cohort, the Hospital Anxiety and Depression Scale (HADS) was used to ascertain cases, using a threshold of 9. This threshold has been validated in primary care against DSM-III-defined psychiatric morbidity for use as a case finder.[24] However, study outcome of improvement in symptoms, defined by a HADS score less than 9, does not appear to have been validated to measure improvement, though both are related. In EPI3, the threshold was thus determined *a priori* based on clinical expert assessment as opposed to a formal validation of the HADS scale to detect improvement. Because the mean HADS score at baseline was similar across the three prescribing preference groups (ranging from 11.8 to 12.0),[5] one would not expect a differential misclassification of outcome across GP types. In the sleep disorder cohort, the Pittsburgh sleep quality index (PSQI) ≥ 1 was used to find cases as well as to define outcome (persistence of sleep disorder over the follow-up). The PSQI has been validated in the literature.[25] In addition, self-reported occurrence of any injury resulting from a fall, motor vehicle collision, sports or employment over the 12-month follow-up was another study outcome. In the URTI cohort, outcome consisted of self-reported change in URTI symptoms (cleared, much improved, slightly improved, no change, or worsened) at one month and associated infections (otitis or sinusitis) in the 12-month follow-up.

Statistical Analysis

In order to compare study outcomes (conventional drug use, symptoms progression) across the prescribing preference types, it was necessary to adjust for differences in patient characteristics. Those characteristics included gender, age group, body mass index, smoking habit, alcohol consumption, physical activity, education, employment, complementary health insurance, CAMBI (Complementary and Alternative Medicine Beliefs Inventory) score, as well as baseline value of disease-specific assessment scores. Propensity scores were derived with those covariates, using multivariate logistic regression. Two sets of propensity scores were derived, indicating the probability of patients to belong to homeopathy versus conventional practice type, and mixed versus conventional practice type, conditional on all other

baseline variables. Propensity scores have become the cornerstone of comparative safety and effectiveness research and, when used for matching, are referred to as pseudo-randomization.[26] For binary outcome variables (e.g., disease progression, conventional drug use), multivariate logistic regression analyses were conducted adjusting for baseline disease characteristics, age, gender and propensity scores. For continuous outcome variables (e.g., functional scores), multivariate ANOVA for repeated measures was used. Non-independence of patients within a given GP type, as well as autocorrelation between responses to the four consecutive interviews, were addressed through generalized estimating equations in the multivariate models. Furthermore, stratification by incident versus prevalent disease status at baseline was conducted.

Conclusion

EPI3 is a large-scale epidemiologic study that confirmed heterogeneity in the characteristics of patients who consult GPs that differ according to their prescribing preferences for homeopathy and complementary medicine. The EPI3 study was based on a systemic construct defined by the type of GP prescribing preference, and was not designed to perform head-to-head comparisons of individual medicines. Owing to careful methods of sampling, findings are highly generalizable to the French population. Overall, evidence generated by the EPI3 study is robust, especially for musculoskeletal diseases, since validated instruments were used for both case ascertainment and study outcomes. As the EPI3 study was non-randomized, the non-comparability of patients followed in the various types of GP practices was addressed through advanced analytic techniques consisting of propensity scores, which are considered to be pseudo-randomization and part of good pharmacoepidemiology practice.

The EPI3 study showed that patients followed in homeopathy preference practice use fewer conventional medicines and without statistically significant differences in symptoms progression or adverse events compared to the other prescribing preferences, thereby indicating no loss of therapeutic opportunity. The EPI3 study highlighted differences between practice types regarding patient characteristics and beliefs. Thus, given those specificities in patient populations, homeopathic practice met specific patient needs.

Highlights
- The conceptual framework for the EPI3 study rests on a systemic construct (GP prescription preference) as opposed to head-to-head comparison of individual medicines.
- Through the enrollment process that was used, the study population was representative of the French population, and selection bias was minimized through adequate study procedures and a registry of non-participants.
- The EPI3 study was based on primary data collection, provided by GPs and self-reported by participants.

- Outcomes were measured using validated scales and instruments for three conditions of interest (musculoskeletal disorders, sleep disorder, anxiety/depression), though only the one for musculoskeletal disease had been validated as a measure of symptoms improvement.
- As the EPI3 study was non-randomized, the non-comparability of patients followed in the various types of GP practices was addressed through advanced analytic techniques consisting of propensity scores.

Conflict of Interest
Yola Moride is a consultant for the pharmaceutical industry, and played no part in the original EPI3 study or its associated discussions. Critical review was conducted at arm's length, whereby the sponsor did not influence the assessments.

References

1 Lert F, Grimaldi-Bensouda L, Rouillon F, et al; EPI3-LA-SER Group. Characteristics of patients consulting their regular primary care physician according to their prescribing preferences for homeopathy and complementary medicine. Homeopathy 2014;103:51–57

2 Grimaldi-Bensouda L, Engel P, Massol J, et al; EPI3-LA-SER group. Who seeks primary care for sleep, anxiety and depressive disorders from physicians prescribing homeopathic and other complementary medicine? Results from the EPI3 population survey. BMJ Open 2012;2:e001498

3 Grimaldi-Bensouda L, Bégaud B, Rossignol M, et al. Management of upper respiratory tract infections by different medical practices, including homeopathy, and consumption of antibiotics in primary care: the EPI3 cohort study in France 2007-2008. PLoS One 2014;9:e89990

4 Rossignol M, Begaud B, Engel P, et al; EPI3-LA-SER group. Impact of physician preferences for homeopathic or conventional medicines on patients with musculoskeletal disorders: results from the EPI3-MSD cohort. Pharmacoepidemiol Drug Saf 2012;21:1093–1101

5 Grimaldi-Bensouda L, Abenhaim L, Massol J, et al; EPI3-LA-SER Group. Utilization of psychotropic drugs by patients consulting for sleeping disorders in homeopathic and conventional primary care settings: the EPI3 cohort study. Homeopathy 2015;104:170–175

6 Grimaldi-Bensouda L, Abenhaim L, Massol J, et al; EPI3-LA-SER group. Homeopathic medical practice for anxiety and depression in primary care: the EPI3 cohort study. BMC Complement Altern Med 2016;16:125

7 Haute Autorité De Santé Commission de la Transparence. Évaluation des Médicaments Homéopathiques Soumis à la Procédure d'Enregistrement Prévue à l'Article L.5121-13 du CSP. Accessed March 24, 2021 at: https://www.has-sante.fr/upload/docs/application/pdf/201906/homeopathie_pic_avis3_cteval415.pdf

8 Colas A, Danno K, Tabar C, Ehreth J, Duru G. Economic impact of homeopathic practice in general medicine in France. Health Econ Rev 2015;5:55

9 Danno K, Joubert C, Duru G, Vetel JM. Physician practicing preferences for conventional or homeopathic medicines in elderly subjects with musculoskeletal disorders in the EPI3-MSD cohort. Clin Epidemiol 2014;6:333–341

10 Danno K, Duru G, Vetel JM. Management of anxiety and depressive disorders in patients ≥ 65 years of age by homeopath general practitioners versus conventional general practitioners, with overview of the EPI3-LASER study results. Homeopathy 2018;107:81–89

11 Public Policy Committee, International Society for Pharmaco-epidemiology. Guidelines for good pharmacoepidemiology practice (GPP). Pharmacoepidemiol Drug Saf 2016;25:2–10

12 European Network of Centers for Pharmacoepidemiology and Pharmacovigilance. ENCePP Guide on Methodological Standards in Pharmacoepidemiology. Accessed March 24, 2021 at: http://www.encepp.eu/standards_and_guidances/methodologicalGuide.shtml

13 Joanna Briggs Institute Critical Appraisal Tools. Accessed March 24, 2021 at: https://jbi.global/critical-appraisal-tools

14 McGill E, Marks D, Er V, Penney T, Petticrew M, Egan M. Qualitative process evaluation from a complex systems perspective: a systematic review and framework for public health evaluators. PLoS Med 2020;17:e1003368

15 Deville JC, Sarndal CE. Calibration estimators in survey sampling. J Am Stat Assoc 1992;87:376–382

16 Howe CJ, Cole SR, Lau B, Napravnik S, Eron JJ Jr. Selection bias due to loss to follow-up in cohort studies. Epidemiology 2016;27:91–97

17 Teague S, Youssef GJ, Macdonald JA, et al; SEED Lifecourse Sciences Theme. Retention strategies in longitudinal cohort studies: a systematic review and meta-analysis. BMC Med Res Methodol 2018;18:151

18 Grimaldi-Bensouda L, Rossignol M, Aubrun E, Benichou J, Abenhaim L; PGRx Study Group. Agreement between patients' self-report and physicians' prescriptions on cardiovascular drug exposure: the PGRx database experience. Pharmacoepidemiol Drug Saf 2010;19:591–595

19 Grimaldi-Bensouda L, Rossignol M, Aubrun E, Benichou J, Abenhaim L; PGRx Study Group. Agreement between patients' self-report and physicians' prescriptions on nonsteroidal anti-inflammatory drugs and other drugs used in musculoskeletal disorders: the international Pharmacoepidemiologic General Research eXtension database. Pharmacoepidemiol Drug Saf 2012;21:753–759

20 Coste J, Le Parc JM, Berge E, Delecoeuillerie G, Paolaggi JB. French validation of a disability rating scale for the evaluation of low back pain (EIFEL questionnaire). Rev Rhum 1993;60:335–341

21 Beaton DE, Wright JG, Katz JN. Upper Extremity Collaborative Group. Development of the QuickDASH: comparison of three item-reduction approaches. J Bone Joint Surg Am 2005;87:1038–1046

22 Lequesne MG, Méry C, Samson M, Marty M. Comparison between the WOMAC and the Lequesne indices in patients with knee and hip osteoarthritis. Osteoarthritis Cartilage 1998;6:441–442

23 Maughan EF, Lewis JS. Outcome measures in chronic low back pain. Eur Spine J 2010;19:1484–1494

24 Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. J Psychosom Res 2002;52:69–77

25 Backhaus J, Junghanns K, Broocks A, Riemann D, Hohagen F. Test-retest reliability and validity of the Pittsburgh Sleep Quality Index in primary insomnia. J Psychosom Res 2002;53:737–740

26 Austin PC, Jembere N, Chiu M. Propensity score matching and complex surveys. Stat Methods Med Res 2018;27:1240–1257