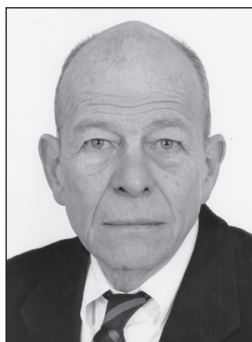


## Database epidemiology

In the current issue of the *Canadian Respiratory Journal*, Gershon et al (1) (pages 183-188) tackle a familiar problem. Our single-payer, largely fee-for-service health care system essentially guarantees that provincial health care authorities have records of physician contacts replete with diagnoses that can be attributed to (anonymous) individuals. These records potentially enable investigators to track physician and hospital contacts for patients with a given diagnosis at relatively low cost and, therefore, are highly attractive to epidemiologists, who would otherwise have to find the requisite patients, then follow them individually for years in the case of chronic diseases such as asthma. However, it is often not clear what exactly is being tracked. In the case of asthma, 'doctor-diagnosed asthma' simply means that a physician labelled a patient with asthma at some point in time. This is problematic because there is no accepted objective diagnostic criterion for asthma and, even if there was, the investigator could not be sure that it was applied in the cases considered. In other words, the investigator has no real idea whether the diagnoses he or she are analyzing are accurate.

Gershon et al (1) report a detailed and difficult study in an attempt to clarify these issues. They selected 13 primary care physician (PCP) practices in Ontario that had electronic billing systems, work space for an auditor and saw at least 30 adult asthmatic patients in 2003. A total of 518 charts from these PCPs were audited and classified into four categories of equal size: those with asthma, those with 'related diseases', such as sinusitis, acute bronchitis and rhinitis, those with chronic obstructive pulmonary disease, and control subjects who had hypertension or musculoskeletal diseases. These charts were submitted to respirologists who essentially classified them into 'asthma' or 'nonasthma' groups. These expert diagnoses were compared with database diagnoses. The PCP diagnoses underwent the same comparison. Comparison with database diagnoses yielded some results that may have been expected – and some that were of greater interest. Of the latter, there was the finding that the expert diagnosis was no better than that of the PCP. For both groups, if the database criterion for asthma was one physician visit for the disease, sensitivity was excellent: more than 90% of people who were believed to have asthma by either experts or PCP had such a visit, but the specificity of a single visit for asthma was low – approximately 60%; many people who did not have asthma apparently experienced such a visit. If the criterion was two ambulatory care visits or one hospitalization for asthma over two years, the sensitivity decreased to approximately 80%, with the specificity increasing to approximately the same figure for both experts and PCPs. Making the criterion three ambulatory care visits for asthma over two years decreased sensitivity but increased specificity. The authors concluded that two ambulatory care visits and/or one hospitalization for asthma over two years would be satisfactory database criteria for asthma.



Nick R Anthonisen

## L'épidémiologie par base de données

Dans le présent numéro de la *Revue canadienne de pneumologie*, Gershon et coll. (1) (pages 183-188) se penchent sur un problème familier. Notre système de santé à payeur unique, largement rémunéré à l'acte, garantit essentiellement que les autorités de santé provinciales disposent de dossiers de médecins remplis de diagnostics attribuables à des individus (anonymes). Ces dossiers permettent aux chercheurs de retracer, à coût relativement faible, les coordonnées de médecins et d'hôpitaux à l'égard de patients ayant une maladie donnée. Par conséquent, ils sont très attrayants pour les épidémiologistes qui, autrement, doivent trouver les patients requis, puis les suivre indivi-

duellement pendant des années dans le cas de maladies chroniques comme l'asthme. Cependant, il arrive souvent qu'on ne sache pas ce qui est suivi au juste. Pour ce qui est de l'asthme, « l'asthme diagnostiqué par un médecin » signifie simplement qu'un jour, un médecin a déclaré un patient asthmatique. C'est problématique, car il n'existe pas de critères diagnostiques objectifs acceptés pour l'asthme; même s'il y en avait, le chercheur ne pourrait savoir avec certitude s'ils ont été appliqués aux cas à l'étude. Autrement dit, le chercheur ne sait pas vraiment si les diagnostics qu'il analyse sont exacts.

Dans une tentative pour clarifier ces questions, Gershon et coll. (1) rendent compte d'une étude détaillée et difficile. Ils ont sélectionné 13 cabinets de soins primaires (CSP) de l'Ontario disposant d'un système de facturation électronique et d'espace de travail pour y installer un vérificateur, puis ils ont vu au moins 30 patients asthmatiques adultes en 2003. Au total, ils ont vérifié 518 dossiers de ces CSP, qu'ils ont classés dans quatre catégories d'égale dimension : les patients asthmatiques, les patients atteints d'une « maladie connexe » comme la sinusite, la bronchite aiguë et la rhinite, les patients atteints d'une maladie pulmonaire obstructive chronique et des sujets témoins atteints d'hypertension ou d'une maladie musculosquelettique. Ces dossiers ont ensuite été soumis à des pneumologues, qui les classaient essentiellement dans le groupe des « asthmatiques » ou celui des « non-asthmatiques ». On comparait ensuite le diagnostic de ces experts à ceux des bases de données. Les diagnostics des CSP ont subi la même comparaison. La comparaison avec les diagnostics des bases de données ont donné certains résultats auxquels on aurait pu s'attendre, ainsi que d'autres, plus intéressants. Parmi ces derniers, on a pu constater que le diagnostic des experts n'était pas plus précis que celui des CSP. Dans les deux groupes, si le critère de la base de données pour l'asthme était une consultation auprès d'un médecin à cause de la maladie, on obtenait une excellente sensibilité : plus de 90 % des personnes que les experts ou les CSP croyaient asthmatiques avaient eu une telle consultation, mais la spécificité d'une seule consultation pour l'asthme était faible, à environ 60 %. En effet, de nombreuses personnes non asthmatiques en avaient vécu une. Si le critère était de deux visites en consultations externes ou

When I served in a large federal institution, approximations were characterized as 'close enough for government work', the implication being that they might not be close enough for anyone else. The Gershon et al (1) results fall into this category. There were 160 cases of asthma diagnosed by the experts; the PCP did not diagnose more than 20% of these cases. On the other hand, the PCP identified 196 of the 518 patients as being asthmatic. These differences were not trivial; by my calculation, of 228 patients, there were approximately 68 whose diagnosis differed between the two sets of physicians. There are several potential causes of such discrepancies, which are addressed by the authors. It is perhaps remarkable that the sensitivity and specificity of a given database definition did not differ between the two groups, and shows that these numbers contained a substantial degree of variation.

Thus, database epidemiology is, at best, an approximation and must always be recognized as such, as Gershon et al indicate. In the case of asthma, others have examined the issue of database accuracy (2-4), with somewhat similar results. The fewer the potentially competing diagnoses, the better the results, the best being in children and young adults who are least likely to have chronic obstructive pulmonary disease. Database epidemiology is best confined to these groups, when possible.

#### REFERENCES

1. Gershon AS, Wan C, Guan J, Vasilievska-Ristovska J, Cicutto L, To T. Identifying patients with physician-diagnosed asthma in health administrative databases. *Can Respir J* 2009;16:183-8.
2. To T, Dell S, Dick PT, et al. Case verification of children with asthma in Ontario. *Pediatr Allergy Immunol* 2006;17:69-76.
3. Blais L, Lemiere C, Menzies D, Berbiche D. Validity of asthma diagnoses recorded in the Medical Service Database of Quebec. *Pharmacoepidemiol Drug Saf* 2006;15:245-52.
4. Huzel L, Roos LL, Anthonisen NR, Manfreda J. Diagnosing asthma: The fit between survey and administrative database. *Can Respir J* 2002;9:407-12.

d'une hospitalisation pour l'asthme sur une période de deux ans, la sensibilité fléchissait à environ 80 %, mais la spécificité passait environ au même pourcentage pour les experts et les CSP. Si le critère était plutôt de trois visites en consultations externes pour l'asthme sur une période de deux ans, la sensibilité diminuait, mais la spécificité augmentait. Les auteurs ont conclu que, sur une période de deux ans, deux visites en consultations externes ou une hospitalisation, ou les deux, constitueraient des critères satisfaisants de l'asthme dans les bases de données.

Lorsque je travaillais dans un grand établissement fédéral, les approximations étaient qualifiées d'« assez justes pour un travail fédéral », laissant entendre qu'elles ne le seraient pas pour qui que ce soit d'autre. Les résultats de Gershon et coll. (1) s'inscrivent dans cette catégorie. En effet, 160 cas d'asthme avaient été diagnostiqués par des experts, et les CSP n'en avaient pas diagnostiqué plus de 20 %. Par contre, les CSP avaient permis de repérer que 196 des 518 patients étaient asthmatiques. Ces différences n'étaient pas négligeables. Selon mes calculs, des 228 patients, environ 68 étaient diagnostiqués différemment par les deux groupes de médecins. Ces différences s'expliquent de plusieurs façons, abordées par les auteurs. Il est peut-être remarquable que la sensibilité et la spécificité de la définition tirée d'une base de données ne diffèrent pas entre les deux groupes. On découvre également que ces chiffres s'associent à une variation substantielle.

Ainsi, comme l'indiquent Gershon et coll., dans le meilleur des cas, l'épidémiologie par base de données est une approximation et doit toujours être perçue comme telle. Pour ce qui est de l'asthme, d'autres ont analysé la question de l'exactitude des bases de données (2-4) et ont obtenu des résultats quelque peu similaires. Le nombre de diagnostics divergents possible est inversement proportionnel aux résultats, les meilleurs résultats étant observés chez les enfants et les jeunes adultes les moins susceptibles d'être atteints d'une maladie pulmonaire obstructive chronique. Dans la mesure du possible, il est préférable de confiner l'épidémiologie par base de données à ces groupes.

#### RÉFÉRENCES

1. Gershon AS, Wan C, Guan J, Vasilievska-Ristovska J, Cicutto L, To T. Identifying patients with physician-diagnosed asthma in health administrative databases. *Can Respir J* 2009;16:183-8.
2. To T, Dell S, Dick PT et coll. Case verification of children with asthma in Ontario. *Pediatr Allergy Immunol* 2006;17:69-76.
3. Blais L, Lemiere C, Menzies D, Berbiche D. Validity of asthma diagnoses recorded in the Medical Service Database of Quebec. *Pharmacoepidemiol Drug Saf* 2006;15:245-52.
4. Huzel L, Roos LL, Anthonisen NR, Manfreda J. Diagnosing asthma: The fit between survey and administrative database. *Can Respir J* 2002;9:407-12.



**Hindawi**  
Submit your manuscripts at  
<http://www.hindawi.com>

