

Véronique Koehn¹, Fred Paccaud²

¹ Service cantonal de recherche et d'information statistiques, Lausanne

² Institut universitaire de médecine sociale et préventive, Lausanne

Variations régionales du recours aux soins: Quelques problèmes méthodologiques

Résumé

De nombreuses études mettent en évidence la variabilité de recours aux soins entre les régions; elles expliquent ensuite cette variation par des facteurs liés à la structure et au fonctionnement du système de soins, par exemple par l'incertitude de la décision clinique ou par le mode de paiement des soins. En fait, l'interprétation de ces variations pose des problèmes strictement statistiques qui ne sauraient être négligés: le recours aux soins est un événement rare, observé dans des populations de tailles réduites, et qui sont le plus souvent de tailles inégales entre elles. En fait, la probabilité que la variabilité observée soit due au hasard est plus importante que ce que la seule intuition conduirait à le penser. D'autre part, la modélisation de cette variabilité, qui permet de construire le test statistique approprié, est difficile. Cette présentation montre quelques-uns des problèmes d'interprétation statistique des indicateurs de variabilité. Elle propose également une méthode d'interprétation des variations basée sur la simulation des variations, adaptée à chaque étude particulière, qui permet d'étalonner les valeurs des indicateurs de dispersion usuels.

Les variations régionales du recours aux services de soins retiennent l'attention en santé publique depuis qu'existe l'information permettant de les mettre en évidence^{1,2}. Classiquement, le territoire est décomposé en domaines géographiques (districts sanitaires ou bassins de clientèles, par exemple), dans lesquels sont calculées des incidences régionales de recours aux soins. Les différences éventuellement observées sont mises en relation avec différents facteurs,

tels que l'incertitude médicale³⁻⁵, l'offre et l'accessibilité aux soins⁶⁻⁸, le mode de financement^{9,10}, l'attitude des patients¹¹ ou les circonstances socio-économiques^{12,13}.

Quels que soient l'ampleur des variations observées et l'intérêt des hypothèses que suggèrent ces variations, la validité des indicateurs mesurant ces variations régionales doit être discutée. Trois groupes d'erreurs doivent être considérés. Le premier concerne les erreurs de dénombrement des

recours aux soins (c'est-à-dire le numérateur de l'incidence); ainsi par exemple, les statistiques hospitalières enregistrent des séjours et non pas des patients, si bien que l'incidence annuelle est surestimée lorsqu'il s'agit de soins susceptibles d'être consommés plusieurs fois durant la même année¹⁴.

Un deuxième groupe d'erreurs concerne le dénombrement de la population «à risque» pour le soin étudié (c'est-à-dire le dénominateur de l'incidence). La population n'est recensée que tous les dix ans, si bien que seules des estimations inter-censitaires sont disponibles¹⁵: leur précision est faible, en particulier lorsque la population étudiée est de petite taille. D'autre part, les domaines géographiques ne sont jamais imperméables, si bien qu'il existe une hétérogénéité entre dénominateur et numérateur.

Une troisième source d'erreurs est liée à la composante aléatoire des variations observées¹⁶. Cette composante est importante lorsque les événements sont rares¹⁷, ce qui est fréquent puisque les populations sont de petites tailles. De plus ces populations sont souvent inégales entre elles. Il n'est donc pas exclu que cette variabilité, même si elle est forte, provienne du hasard. Lorsque le but d'une analyse de variation régionale est descriptif,

peu importe que les données observées soient le résultat d'épreuves aléatoires. En revanche, lorsqu'il s'agit de faire une inférence, il faut répondre à une question plus précise: les variations d'incidence entre régions sont-elles dues au hasard, c'est-à-dire au fait que chaque individu n'a qu'une probabilité de recourir à un service de soins? Dans ce cas, l'inférence ne peut être faite qu'après avoir vérifié la stabilité des indicateurs de variation.

L'objet de cet article est d'explorer les implications de ce troisième groupe d'erreurs.

Probabilité de recours aux soins

Soit k régions ($i=1\dots k$), dont les effectifs de population sont $n_1\dots n_k$, et dont la somme est:

$$n = \sum n_i \quad (1)$$

Soit une intervention qui ne peut être effectuée qu'une seule fois pendant la période d'examen. Les incidences observées dans chacune des k régions sont $p_1\dots p_k$, et l'incidence moyenne p est:

$$p = \frac{\sum n_i \times p_i}{n} \quad (2)$$

Ces incidences peuvent être stratifiées (par âge et par sexe en particulier), mais cette décomposition est ignorée ici afin d'alléger les notations.

Si l'on suppose que tous les individus ont la même probabilité P de subir cette intervention et sous certaines hypothèses d'indépendance (pas de réhospitalisations, pas de phénomène d'épidémie), le nombre d'événements observés dans chacune des k régions de taille n_i suit une loi binomiale $B(n_i, P)$ (ou, par approximation, une loi de Poisson de paramètre $\lambda_i = n_i P$) dans l'ensemble des régions avec k tirages selon cette distribution.

La Figure 1 présente les résultats d'une simulation basée sur ces principes. Soit 200 régions de 20000 habitants, dans lesquelles on étudie une intervention dont l'incidence annuelle moyenne est de 3 pour mille. Le nombre attendu d'opérations dans chacune des 200 régions suit une loi binomiale de forme $B(n_i=2000, P=0.003)$. De façon à stabiliser les résultats, dix tirages ont été effectués pour chaque région.

La Figure 1 fait constater qu'il existe une importante variabilité autour de la valeur «vraie» (3 pour mille): les valeurs oscillent dans un intervalle de 2 pour mille, correspondant à une variation absolue du nombre d'interventions allant de 40 à 80.

Indicateurs de variabilité

Les indicateurs couramment utilisés pour mesurer les variations régionales sont le coefficient de variation, le quotient extrémal et la composante systématique de variation.

Le coefficient de variation (CV) divise l'écart-type (ECT), qui est l'indice de dispersion par excellence, par l'incidence moyenne empirique, de façon à corriger l'effet de l'unité de mesure choisie:

$$CV = \frac{ECT}{p} \quad (3)$$

L'écart-type des taux entre régions (ECT) est ici pondéré par les populations.

Le quotient extrémal (QE) est le rapport entre la plus grande et la plus petite des valeurs observées. Bien qu'il n'utilise qu'une petite partie de l'information disponible (2 des k valeurs disponibles), le QE est souvent utilisé à cause de sa simplicité¹⁸ (QE=2 signifie que l'incidence de l'intervention varie du simple au double):

$$QE = \frac{\max(p_i)}{\min(p_i)} \quad (4)$$

La composante systématique de variation (CSV) est une mesure propre aux analyses de variations régionales, introduite en 1982 par McPherson et collaborateurs³. Bien que peu intuitive, elle est très répandue, car elle est supposée éliminer l'effet des variations dues au hasard.

Elle se fonde sur les différences entre le nombre d'interventions observées, O_i , et le nombre d'interventions attendues E_i :

$$E_i = n_i \times P \quad (5)$$

La CSV est définie de la façon suivante:

$$CSV = \frac{\sum \left(\frac{O_i - E_i}{E_i} \right)^2 - \sum \frac{1}{E_i}}{k} \quad (6)$$

Le χ^2 ne mesure pas la variabilité observée entre régions, mais répond très exactement à la question de l'homogénéité des incidences entre les régions: le χ^2 est par excellence la mesure de la dépendance entre deux variables qualitatives, soit ici la région d'appartenance et le fait d'être opéré ou pas:

$$\chi^2 = \frac{1}{1-P} \sum \frac{(O_i - E_i)^2}{E_i} \quad (7)$$

Chacun de ces quatre indicateurs de variations mesure la quantité de variations, mais ne renseigne pas sur l'élément aléatoire de ces variations. Or, la Figure 1 suggère que les effets du hasard sont substantiels: dans l'exemple développé, le QE atteint la valeur de 2 (puisque l'incidence varie du simple au double) par le simple effet des variations aléatoires. Avant de conclure à l'existence de vraies différences entre les régions, il faut donc disposer d'un test de signification statistique.

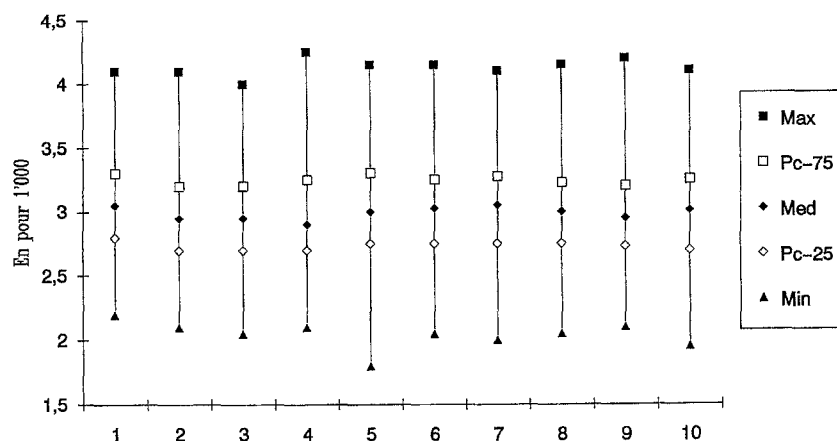


Figure 1. Variations dues au hasard des taux d'incidences sur 200 régions (10 simulations).

Test de signification statistique

Un tel test doit être construit en référence à un modèle, caractérisé par l'ensemble des valeurs possibles des observations. Les modèles paramétriques sont les plus simples (ils varient entre eux par quelques paramètres) dont le choix conditionne la manière d'aborder le problème. L'hypothèse de base suppose que tous les individus d'une région i ont la même probabilité P_i de subir une intervention. Le nombre d'interventions X_i dans la région i de population n_i suit une loi de Poisson de paramètres $\lambda_i = n_i P_i$.

L'hypothèse nulle est rejetée s'il existe au moins une région où la probabilité de subir une intervention est différente de celle des autres régions. Par conséquent, l'hypothèse nulle est:

$$H_0: P_1 = P_2 = \dots = P_k = P \quad (8)$$

Un autre modèle a été proposé par McPherson et collaborateurs³: il n'est pas basé sur les comportements individuels (la probabilité de subir l'intervention, comme dans l'exemple précédent), mais sur le nombre d'opérations survenant dans chaque région. Ce nombre suit une loi de Poisson de para-

mètres $l_i = n_i P$. Les l_i sont eux-mêmes des réalisations d'une variable aléatoire L , qui suit une loi normale $N(1, s)$. En d'autres termes, dans chaque région les probabilités individuelles (P_i) sont identiques, de même que le nombre d'interventions est partout distribué suivant une loi de Poisson; en revanche, il existe un «facteur de dilatation», I_i qui modifie le paramètre, c'est-à-dire la forme de la distribution de Poisson, ce facteur étant propre à chaque région (et donc à chaque intervention dans cette région). L'hypothèse nulle est dès lors:

$$H_0: s = 0 \quad (9)$$

La CSV est un estimateur sans biais de s .

Beaucoup d'autres modèles paramétriques sont envisageables, choisissant d'autres distributions comme les lois beta¹⁸.

Valeurs-seuils de l'indicateur de variabilité

Lorsqu'un modèle est choisi, il faut déterminer la valeur-seuil de l'indicateur de variabilité à partir de laquelle l'hypothèse nulle sera rejetée. Dans les tests classiques

(test de Student pour la comparaison de moyennes d'échantillons, χ^2 d'indépendance, etc.), ces statistiques de test suivent (si l'hypothèse nulle est vraie) des lois statistiques connues et les valeurs seuils sont données dans des tables statistiques.

Dans le cas des indicateurs de variation, aucune loi n'est connue. L'élaboration et l'utilisation de modèles théoriques sont ici limitées par la faible quantité d'informations disponibles: un seul chiffre dans chaque région masque toute la variabilité naturelle du phénomène. Dans ce type de situation, la simulation est l'approche la plus adéquate pour déterminer ces valeurs-théoriques. C'est ce qui a été fait dans l'exemple proposé plus haut (Fig. 1); il montrait qu'il est statistiquement acceptable que les taux varient du simple au double, et l'on rejettera l'hypothèse nulle (pas de différence d'incidence de l'intervention) dès que le QE dépassera 2.

Pour une utilisation générale, cet exemple est limité parce qu'une seule combinaison a été explorée (incidence de 3 pour mille dans 200 régions dont la population est strictement identique). De plus, le nombre de simulations (10) est très faible. Il faut donc explorer un plus grand nombre de situations pour développer une table statistique.

C'est ce qu'ont fait Diehr et al.¹⁴, en exploitant une situation réelle: celle de l'Etat de Washington, divisé en 39 comtés dont la population varie entre 2600 et 1,3 millions d'habitants (moyenne par comté = 114000). Les autres ont considéré 8 incidences d'interventions entre 50 et 10000 pour cent mille, avec une distribution binomiale dans chacun des comtés de l'Etat de Washington. Mille simulations ont été effectuées. Les valeurs-seuils (au percentile 95) des quatre indicateurs de variation (CV , QE , CSV et χ^2) ont été calculées.

Les résultats sont présentés dans le Tableau 1. Par exemple, pour une

	Incidence pour 100 000 hab.	39 régions de taille identique (114 000 hab.)	39 régions avec leur taille réelle	les 31 régions de plus de 10 000 hab.	10 régions de taille identique (114 000)
CV	50	0.16	0.16	0.14	0.19
	100	0.11	0.11	0.10	0.13
	250	0.07	0.07	0.06	0.08
	500	0.05	0.05	0.05	0.06
	1000	0.04	0.04	0.03	0.04
	2500	0.02	0.02	0.02	0.03
	5000	0.02	0.02	0.01	0.02
	10000	0.01	0.01	0.01	0.01
QE	50	2.2	11.1	6.8	1.9
	100	1.7	8.9	3.1	1.5
	250	1.4	4.7	2.0	1.3
	500	1.3	2.6	1.6	1.2
	1000	1.2	2.0	1.4	1.1
	2500	1.1	1.5	1.2	1.1
	5000	1.1	1.3	1.1	1.1
	10000	1.1	1.2	1.1	1.0
CSV	50	7.1	100.3	27.3	15.1
	100	3.5	41.2	13.3	6.8
	250	1.5	17.4	6.1	2.8
	500	0.7	8.1	2.7	1.6
	1000	0.3	4.4	1.4	0.7
	2500	0.1	1.6	0.6	0.3
	5000	0.1	0.7	0.3	0.1
	10000	0.0	0.3	0.1	0.1
$\chi^2(*)$	50	54.7	55.3	54.6	55.4
	100	54.6	55.5	54.6	52.8
	250	55.7	53.9	57.2	53.6
	500	55.4	54.9	55.0	57.2
	1000	54.0	55.1	53.5	52.4
	2500	54.1	55.0	55.0	55.1
	5000	53.4	53.7	55.0	56.6
	10000	54.1	53.9	53.0	54.8

(*) le percentile 95 d'un χ^2 à 39 degrés de liberté est de 54.57. Pour les deux dernières colonnes, les auteurs ont choisi de redresser les seuils obtenus par simulation par le rapport entre cette valeur et le percentile 95 d'un χ^2 à, respectivement, 31 et 10 degrés de liberté. Ceci permet de tenir compte de la différence entre le nombre de régions.
Source: Diehr et al 1992.

Tableau 1. Valeurs-seuils de 4 indicateurs de dispersion, selon l'incidence, le nombre de régions et leur taille.

incidence de 10000 pour 100000, 95% des simulations donnent un CV inférieur à 0.01, quelle que soit la combinaison envisagée du nombre des régions et de la taille de chacune des régions: tout CV supérieur

à 0.01 viole donc l'hypothèse nulle. Ce tableau appelle plusieurs commentaires. Le CV résiste bien aux variations de la taille des régions, mais varie fortement selon l'incidence: ainsi, diviser l'écart-type

par la moyenne n'élimine pas les effets d'échelle. Le QE est très sensible aux variations de la taille des régions étudiées, en particulier lorsque l'incidence moyenne du recours aux soins est faible. La

CSV est sensible à la fois à l'incidence, à la taille des régions et au nombre de régions étudiées: l'utilisation de routine de la CSV pour comparer différentes procédures chirurgicales ou différents pays est en partie abusive. Enfin, il faut remarquer que le χ^2 est le seul indicateur qui reste stable dans les quatre situations décrites.

Conclusion

L'analyse des variations régionales du recours aux services de soins ouvre des perspectives considérables de recherches^{19–23} et de gestion^{10,24}, mais les nombreux problèmes techniques doivent être abordés et résolus. Ils découlent des imprécisions du numérateur et du dénominateur des incidences de recours, mais aussi des variations aléatoires des indicateurs de variabilité. Les résultats présentés ici suggèrent qu'aucune méthode simple ne permet de décider «d'un coup d'oeil» si les variations constatées sont dues au hasard plutôt qu'à des différences réelles, et que les indicateurs utilisés dans les analyses de variations régionales sont peu stables. On retiendra en particulier l'impact des variations de la taille des régions sur les indicateurs^{25,26}.

En pratique, il faut utiliser simultanément plusieurs indicateurs de variation afin d'étayer la décision finale. La démarche applicable dans chaque situation est la simulation. Ceci a d'ailleurs déjà été appliqué dans les analyses de la mortalité hospitalière, en discutant par exemple les avantages et les inconvénients liés à l'allongement de la période d'analyse²⁷. Les ordinateurs et les logiciels disponibles permettent de simuler ce que l'on devrait observer s'il n'existait aucune différence de pratique régionale. La marche à suivre est rappelée dans l'Annexe.

Zusammenfassung

Regionale Unterschiede in der Inanspruchnahme medizinischer Dienstleistungen: Einige methodologische Probleme

Zahlreich sind die Studien, die regionale Unterschiede in der Inanspruchnahme medizinischer Dienstleistungen aufzeigen. Die Autoren erklären diese Unterschiede jeweils mit Faktoren, die durch die Struktur des Gesundheitswesens und das Zusammenspiel seiner Protagonisten beeinflusst werden. Das jeweilige System zur Vergütung medizinischer Leistungen ist ein Beispiel dafür. Die Interpretation der beobachteten Unterschiede bringt statistische Probleme mit sich, die nicht vernachlässigt werden dürfen: die Inanspruchnahme einer medizinischen Dienstleistung ist ein seltenes Ereignis, das in beschränkten und meist unterschiedlich grossen Bevölkerungen beobachtet wird. Die Wahrscheinlichkeit, dass die beobachteten Unterschiede durch den Zufall bedingt sind, ist in der Tat grösser, als man glauben möchte. Andererseits ist es schwierig, die Variablen so zu transformieren, dass die beobachteten Unterschiede mit einem adequaten statistischen Test überprüft werden können. Diese Arbeit zeigt einige der Probleme der statistischen Interpretation der beobachteten Unterschiede auf. Schliesslich wird eine Interpretationsmethode vorgeschlagen, die auf einer Simulation der Variabilität in der jeweiligen Studiensituation beruht und erlaubt, die üblichen Indikatoren der statistischen Streuung entsprechend anzupassen.

Summary

Assessment of regional differences in utilisation of health services: Inherent statistical problems

Many studies show strong variation of health consumption between regions, suggesting that these variations are related to the uncertainty of medical practice or to other factors related to health services or patients attitude. However the statistical interpretation of these variations is far from easy: apart from usual and specific information bias, there are statistical problems when observing incidence of events like health care consumption: it is in fact a rare event, which is observed within small population, and among regions with unequal number of person. Therefore, most of the variation reported might be well explained by a purely statistical phenomenon. This paper presents some aspects of this variability for three common indicators of variation, and suggest the use of ad hoc simulation to get statistical criterias.

Références

- 1 Paul-Shaheen P, Clark JD, Williams D. Small area analysis: a review of the North American literature. *J Health Polit Policy Law*; 741–809.
- 2 Sanders D, Coulter A, McPherson K. Variations in hospital admission rates: a review of the literature. In: The evidence for geographical variations in admission rates. London: King's Fund Publishing Office, 1989:10–34.
- 3 McPherson K, Wennberg JE, Hovind OB, Clifford P. Small-area variations in the use of common surgical procedures: an international comparison of New England, England and Norway. *N Engl J Med* 1982; 307:1310–1314.
- 4 Bunker P. Surgical manpower: a comparison of operations and surgeons in the US and in England and Wales. *N Engl J Med* 1970; 282: 135–144.
- 5 Eckerlund I, Hakansson S. Variation in resource utilization – the role of medical practice and its economic impact. *Soc Sci Med* 1989; 165–173.
- 6 McPherson K, Strong PM, Epstein A, Jones L. Regional variations in the use of common surgical procedures: within and between England and Wales, Canada and the United States. *Soc Sci Med* 1981; 15A: 273–288.
- 7 Lewis CE. Variations in the incidence of surgery. *N Engl J Med* 1969; 281:880–884.
- 8 Lagoe RJ, Millirem JW. A community-based analysis of ambulatory based utilization. *Am J Public Health* 1986; 76:150–153.
- 9 Vayda E. A comparison of surgical rates in Canada and in England and Wales. *N Engl J Med* 1973; 289: 1224–1229.
- 10 Wennberg JE, McPherson K, Casper P. Will payment based on DRGs control hospital costs? *N Engl J Med* 1984; 311:295–300.
- 11 Drury P. ReViews: The consumer. *Health Affairs* 1984; 3:46–49.
- 12 McLaughlin CG, Normolle DP, Wolfe RA, McMahon LF, Griffith JR. Small-area variation in hospital discharge rates: Do socioeconomic variables matter? *Med Care* 1989; 27:507–521.
- 13 Tedeschi PJ, Wolfe RA, Griffith JR. Micro-area variation in hospital use. *Health Serv Res* 1990; 24: 729–740.
- 14 Diehr P, Cain K, Connel F, Volinn MD, Volinn E. What is too much variations? The null hypothesis in small-area analysis. *Health Serv Res* 1990; 24:741–771.
- 15 Verma SBP, Basavarajappa KG. Recent developments in the regression method for estimation of population for small areas in Canada. In: Platek R, Rao JNK, Särndal CE, Singh M, eds. *Small Area Statistics*. New York: John Wiley, 1987:46–61.
- 16 Diehr P, Cain KC, Kreuter W, et al. Can small-area detect variation in surgery rates? The power of small-area variation analysis. *Med Care* 1992; 30:484–502.
- 17 Anonymous. Dicing with death rates. [Editorial]. *Lancet* 1993; 341: 1183–1184.
- 18 Kazandjian VA, Durance PW, Schork MA. The extremal quotient in small-area variation analysis. *Health Serv Res* 1989; 24: 665–684.
- 19 Chassin MR. Does inappropriate use explain geographic variations in the use of health care service? A study in three procedures. *JAMA* 1987; 258:2533–2537.
- 20 Wennberg JE, Mulley AG, Hanley D, et al. An assessment of prostatectomy for benign urinary tract obstruction: geographic variations and the evaluation of medical care outcomes. *JAMA* 1988; 259:3027–3030.
- 21 Eddy DM. Variations in physician practice: the role of uncertainty. *Health Affairs* 1984; 3: 74–89.
- 22 Input and outcome. [Editorial]. *Lancet* 1987; 1:1182–1183.
- 23 Brook RH, Lohr KN, Chassin M, Kosecoff J, Fink A, Solomon D. Geographic variations in the use of services: do they have any clinical significance? *Health Affairs* 1984; 3:64–73.
- 24 Bevan G, Ingram R. Variations in admission rates: implications for equitable allocation of resources. *BMJ* 1987; 295:1039–1042.
- 25 Cislighi C, Braga M, Danielli A, Luppi G. An analysis of the spatial association between cancer mortality and risk factor: the role of geographical scale. *Escape Population Sociétés* 1990; 3:407–416.
- 26 Noin D. L'étude géographique de la mortalité: bilan et problèmes. *Espace Population Sociétés* 1990; 3:367–376.
- 27 Park RE, Brook RH, Kosecoff J, et al. Explaining variations in hospital death rates: randomness, severity of illness, quality of care. *JAMA* 1990; 264:484–490.

Adresse pour correspondance

Véronique Koehn
Service cantonal de recherche
et d'information statistiques
Rue de la Paix 6
CH-1014 Lausanne

Annexe:
Marche à suivre pour réaliser
une simulation

Les incidences p_i sont observées dans k régions de tailles n_i .

1. Déterminer l'incidence moyenne p dans les k régions (ou, éventuellement, plusieurs valeurs dans une fourchette raisonnable de taux);

2. Générer 1000 fois k nombres suivant une loi binomiale de paramètres $B(n_i, p)$;

3. Pour chacune de ces 1000 simulations, calculer la valeur des 4 indicateurs de dispersion pour déterminer le percentile 95 de la distribution de chacun de ces indicateurs;

Répéter les étapes 2 et 3. Tant que les résultats divergent sensiblement augmenter le nombre de simulations;

4. Confronter les résultats observés et simulés: l'hypothèse d'une «variation due au hasard» n'est rejetée que si les valeurs observées des indicateurs de variation sont au-delà des percentiles 95 obtenus par simulation.

Si une ou plusieurs régions présentent des valeurs extrême du p_i , il est raisonnable de recommencer la démarche en éliminant la ou les valeurs extrêmes: il est en effet possible que de grandes valeurs d'un indicateur soient attribuables au comportement atypique d'une région.