

Recommandations pour la gestion des données territoriales

Recommandations pour la gestion des données territoriales dans les systèmes d'information en santé et en services sociaux

Rappel du contexte

Nous possédons un grand nombre d'informations associées à des individus, à des clients, à des services rendus ou à des ressources (désignées ici *données sociosanitaires*) qui sont dispersées dans plusieurs banques de données. Ces diverses entités sont, en général, localisables géographiquement grâce à des *informations de localisation*, comme l'adresse, le code postal, la municipalité, l'aire de diffusion (ou d'autres types de territoires plus ou moins grands), qui sont saisies à la source primaire de données. Parfois, les données sont associées simplement à un établissement (à un lieu de dispensation plutôt qu'à un lieu de résidence) et c'est alors l'établissement qui possède des informations de localisation.

Typiquement, pour nos fins de gestion, surtout en planification et en évaluation, nous devons analyser ces données sociosanitaires *dans le temps* et *dans l'espace*. Souvent, les *territoires d'analyse* sont différents de ceux qui correspondent aux informations de localisation. On utilise par exemple la région sociosanitaire (RSS), le réseau territorial de service (RTS), le territoire du réseau local de service (RLS) ou le territoire de CLSC. Des *tables de correspondance* servent donc à déterminer le territoire d'analyse à partir des informations de localisation et le résultat est généralement mémorisé comme une variable territoriale additionnelle dans la banque. Certaines tables de correspondance constituent en fait des *définitions territoriales*. Ainsi, les territoires de CLSC sont actuellement définis à l'aide des codes postaux et des codes municipaux qui leur sont reliés. Les RLS, les RTS et les RSS sont simplement des regroupements de territoires CLSC.

Les choses sont cependant plus compliquées qu'il n'y paraît à prime abord. Premièrement, les informations de localisation ne sont pas toujours bien validées à la source. Deuxièmement, les correspondances sont parfois imparfaites et l'utilisation des tables est délicate de sorte qu'elle devrait être supervisée par des experts pour donner les meilleurs résultats. Troisièmement et principalement, les entités territoriales impliquées subissent des changements dans le temps, à fréquence variable et imprévisible. Par exemple, les territoires desservis par les codes postaux peuvent changer et de nouveaux territoires sont constamment créés par Postes Canada. Les municipalités subissent parfois des modifications de frontières, de même que les territoires de CLSC. Ces changements sont souvent non apparents car les codes territoriaux ne changent pas automatiquement. Ainsi, la table de correspondance utilisée l'an dernier n'est probablement plus bonne cette année. Les conséquences sont évidentes pour ceux qui désirent analyser l'évolution temporelle d'un phénomène pour un territoire donné. Ils ne peuvent pas se fier aux codes territoriaux présents dans l'historique des banques, s'ils n'ont pas été mis à jour récemment.

Recommandations pour les informations de localisation

Les informations de localisation saisies à la source devraient être le *code postal* complet de 6 positions et le *code municipal*.

Ce dernier ne doit pas être simplement déduit du code postal, puisque ce n'est pas toujours possible. Au Québec, plusieurs codes postaux (près de 600) sont effectivement reliés à plus d'une municipalité, surtout en milieu rural. Le problème peut sembler négligeable globalement et être en déclin, mais il est encore sérieux dans certaines parties du Québec.

En fait, le *code municipal* n'est pas à proprement parlé « saisi », il devrait toujours être déduit à la fois du code postal et du nom de la municipalité faisant partie de l'adresse. Idéalement, un système informatique devrait supporter cette tâche au moment ou quelqu'un est disponible pour fournir l'information, par exemple l'admission à l'hôpital. Le code postal est d'abord saisi, puis le système fournit une liste de municipalités possibles pour ce code (l'ensemble des toponymes connus, soit les différentes appellations courantes de ces municipalités). On doit alors choisir la bonne municipalité, en distinguant même le village de la paroisse portant le même nom dans certain cas.

Si le nom ne semble pas être dans la liste, il y a trois possibilités :

- le code postal a été *mal saisi*
 - on doit le corriger;
- le nom de municipalité (de l'adresse) est *incorrect*
 - on peut le remplacer par un autre, tiré de la liste;
- les tables de correspondance utilisées sont *incomplètes*
 - on saisit le nouveau nom de municipalité et le système fournit une liste de codes municipaux possibles (parmi un ensemble de toponymes accompagnés de codes postaux, nom de MRC et nom de région qui aideront à choisir le bon);
 - le *code postal* lui-même peut être *absent de la table* s'il est nouveau, dans ce cas, on pourrait alors accepter ce code si la personne qui le fournit le confirme après un *avertissement* et si ce code respecte les normes (dont le format ANA NAN);
 - le *code postal* peut être *présent dans la table mais inactif* selon Postes Canada et cette validation additionnelle permet d'exiger un code postal actif s'il est important pour la livraison efficace de courrier.

Pour déterminer le code municipal, il est très utile de consulter le code postal, car le même nom de municipalité peut souvent correspondre à des codes différents dans diverses régions. De plus, des gens utiliseront un nom comme Montréal, alors que le code postal précisera qu'il s'agit de Montréal-Nord, par exemple.

S'il n'est pas possible de faire le travail en présence du client, il faut du moins saisir le nom de la municipalité faisant partie de l'adresse (en plus du code postal) pour éventuellement s'en servir pour déterminer le code municipal. Il sera cependant plus difficile de résoudre les incohérences occasionnelles entre les deux informations de localisation.

Il faut absolument éviter d'utiliser seulement l'identifiant du lieu de service pour localiser les clientèles. Celles-ci peuvent résider très loin de ce lieu. Même si un établissement comme un CLSC semble avoir un bassin de desserte assez bien défini et stable, ce n'est pas toujours le cas et surtout pas à long terme. Les données ainsi rattachées à un simple code d'établissement peuvent difficilement servir à des analyses d'évolution de performance par rapport aux besoins d'une population et à ses facteurs de risque. D'une année à l'autre, on ne peut pas constituer un bassin comparable de clients.

Remarquez finalement *qu'une localisation doit être associée à une date*. Les individus (même les établissements) déménagent et il ne suffit pas de garder leur dernière adresse en mémoire. C'est l'adresse lors de chaque « événement » qui nous sera utile.

Recommandations pour les définitions territoriales de CLSC (ou autres territoires spécifiques)

Pour délimiter les frontières de territoires spécifiques, comme les CLSC, il est courant de se baser sur les limites de territoires bien reconnus comme les municipalités, les MRC et les régions administratives. C'est une pratique hautement recommandable et on devrait aller jusqu'à préciser si cette *dépendance territoriale* est *continue ou occasionnelle*. Est-ce que la frontière du CLSC doit suivre automatiquement toute modification de celle de la municipalité désignée ? (ceci faciliterait grandement les procédures de mise à jour) Ou bien, doit-on considérer les frontières indépendantes dans le futur ? Il faudrait alors établir une procédure claire de révision, d'approbation et de mise en vigueur des changements de frontières du CLSC.

Assez souvent, il faut diviser une municipalité en plusieurs territoires. On utilise alors naturellement le réseau routier, d'autres éléments physiques repérables (cours d'eau, ligne d'hydroélectricité...) ou des codes postaux (à 3 ou 6 positions). Il vaut mieux éviter de définir ces *frontières intra-municipales* en termes d'entités territoriales qui n'ont pas de reconnaissance généralisée (ou risquent de la perdre) et n'ont pas de frontières précisées et gardées en vigueur. À éviter donc, les anciens quartiers et anciennes localités, de même que les territoires du recensement canadien (AD SR), à moins qu'on ne compte travailler qu'avec les données d'un seul recensement.

Il faut justement penser à quelles fonctions principales devront servir ces définitions territoriales. Les territoires CLSC sont définis dans le but premier d'assigner une population à desservir à chaque établissement CLSC. En plus, ils servent de territoires d'analyse pour toutes sortes de travaux qui demandent d'agréger des données liées généralement à des individus ou à des lieux de service. Dans ces multiples sources, la localisation part donc d'une adresse qui est représentée numériquement par un code postal et un code municipal. On voit alors l'intérêt à définir les CLSC avec ces deux informations. C'est ce qui est nécessairement fait dans nos tables de correspondance. On peut parfois décrire un territoire en faisant une liste d'adresses ou en traçant une carte de rues, mais sa *définition opérationnelle* est bien en codes postaux et municipaux (quand on travaille avec des données du recensement, on utilise une autre définition en termes de DRSDR, SR et CEFSR). Cette définition implique aussi une dépendance continue des CLSC avec l'évolution des codes postaux et des municipalités. Si une municipalité frontalière annexe une partie d'une voisine située dans l'autre CLSC, cette partie est automatiquement transférée de CLSC, à moins que l'on n'intervienne en identifiant les codes postaux impliqués et en modifiant les tables de correspondance.

Un cas particulier survient quand deux municipalités de *CLSC différents* sont fusionnées. Il faut alors décider à quel CLSC attribuer la résultante, ou bien si l'on doit la fractionner en utilisant les codes postaux. Parfois ce fractionnement est difficile quand il y a peu de codes postaux (ou un seul). Il peut arriver aussi qu'une municipalité change de MRC et que notre frontière de CLSC suive traditionnellement celle de la MRC. Devons-nous modifier le CLSC ? À quelle date ? Par quel mécanisme devons nous prendre ces décisions ?

En résumé, il est recommandé de définir les CLSC selon *une dépendance continue avec les codes municipaux et les codes postaux*. Une frontière intra-municipale devrait être définie par des codes postaux entiers. Il est possible d'attribuer une fraction à un code postal pour indiquer qu'il est partagé entre deux CLSC, mais cela doit être évité puisqu'un individu relié à ce code devra être assigné aléatoirement à un CLSC (quelques cas semblables existent actuellement dans nos définitions, parce qu'on a rattaché une frontière au réseau routier ou à une ancienne localité, sans égard aux limites effectives des codes postaux). Enfin, une procédure claire doit être établie pour régler efficacement les cas particuliers reliés à des événements territoriaux affectant nos frontières.

Recommandations pour l'attribution des territoires d'analyse

Les territoires d'analyse (ex. CLSC) doivent être attribués à l'aide des *plus récentes tables* de correspondances. Celles-ci sont modifiées normalement chaque année et entrent en vigueur le 1er avril pour l'ensemble des intervenants du réseau de la santé et des services sociaux. Au moment de l'attribution, il faut avoir accès aux informations détaillées et valides de localisation (code postal complet et code municipal).

Un outil informatisé doit être développé pour faciliter la diffusion de ces tables et leur utilisation. Il doit servir à *standardiser les résultats* des attributions et ainsi obtenir les mêmes territoires d'analyse officiels dans toutes les banques de données. Si cet outil n'est pas utilisé selon les spécifications, on ne peut pas être sûr de la qualité de l'attribution territoriale. C'est pourquoi, chaque «système» référant à nos territoires (CLSC, RLS, RTS, RSS, RUIS) devra s'adapter au plus tôt à cet outil, ce qui demande une planification et un financement adéquats.

Avant d'utiliser des données sociosanitaires selon un certain territoire d'analyse, il faut donc s'assurer que l'attribution a été réalisée adéquatement, dans chacune des sources de nos données. Si celles-ci proviennent d'un historique, l'attribution doit avoir été *actualisée pour toute la période couverte*. On ne peut se fier à une attribution qui aurait été réalisée à l'aide d'une ancienne version des tables de correspondance. L'objectif est de répartir toutes les données (même historiques) dans la définition actuelle du territoire d'analyse.

Il est possible d'identifier quels territoires sont modifiés dans chaque nouvelle version des tables et il peut arriver qu'il n'y ait pas de changements significatifs pendant une période donnée. Les pilotes de systèmes en seront informés et pourront donc décider quelles banques de données sociosanitaires doivent être actualisées, compte tenu de la partie du Québec qu'elles couvrent. Ils devront également décider si, dans leur contexte, il faut mémoriser les territoires d'analyse dans la banque (et pour combien d'années de l'historique), ou s'ils seront plutôt attribués à partir des tables lors de chaque exploitation demandant ces territoires.