

Observatoire épidémiologique des données sur la santé du Brésil

Raphael de Freitas Saldanha¹, Ronaldo Rocha Bastos²

Résumé

Le Système de santé unifié brésilien (SUS) a été créé en 1988 et, afin d'organiser les systèmes d'information sur la santé et les bases de données déjà utilisées, une banque de données unifiée (DataSUS) a été créée en 1991. Les fichiers de DataSUS sont facilement disponibles sur Internet. L'accès à ces données et leur consultation se font actuellement au moyen d'un nombre limité de tableaux personnalisés et de diagrammes simples, qui ne répondent pas entièrement aux besoins des gestionnaires de la santé et des autres utilisateurs, qui souhaitent un outil souple et facile à utiliser, qui peut tenir compte des différents aspects de la santé qui les concernent, en vue de la recherche de connaissances et de la prise de décisions. Le projet actuel propose la production mensuelle interactive de rapports épidémiologiques synthétiques, qui sont non seulement facilement accessibles, mais aussi faciles à interpréter et à comprendre. L'accent est mis sur la visualisation des données, grâce à des diagrammes et des cartes plus informatifs.

Mots-clés : diffusion des données sur la santé; algorithmes dynamiques; données sur le Web.

1. Introduction

La production et la disponibilité de données sur la santé sont essentielles pour améliorer les systèmes de santé, ainsi que pour guider le processus de prise de décisions et de recherche (Handley et coll., 2015). La disponibilité de données sur la santé dans les pays en développement est souvent inappropriée (Boerma et Stansfield, 2007), même si le Brésil investit dans ce domaine depuis les années 1990. Suite au processus de redémocratisation qui a commencé en 1988, la nouvelle constitution brésilienne a permis l'établissement d'un système de santé unifié, et en 1991, le service informatique du système de santé brésilien (DataSUS) a été créé.

Depuis sa création, DataSUS recueille des données sur la mortalité, la morbidité, la fourniture des services et les ressources en santé, en vue de la prise de décisions et de la recherche. À l'heure actuelle, toutes les données recueillies par DataSUS sont disponibles dans le site Web du ministère (<http://datasus.saude.gov.br>).

Parmi les grandes préoccupations concernant les données sur la santé dans les pays en développement, mise à part la production des données proprement dites, figure l'utilisation de données déjà produites, mais pas encore utilisées, ce qui mène à la nécessité de raffermir l'analyse des données (Boerma et Stansfield, 2007). On fait aussi face à ce problème au Brésil : compte tenu de son propre objectif, DataSUS est davantage axé sur la production et la diffusion de données que sur l'analyse, ce dont rend compte son site Web. L'accès aux données nécessite de la formation et des connaissances sur l'épidémiologie, étant donné que l'on ne peut produire que des fichiers ou des tableaux bruts, ce qui exige de l'utilisateur une deuxième étape d'analyse pour transformer les données en information.

Les données de 2014 montrent qu'au Brésil, 85 % des services de santé publique ont accès à Internet, et 99 % des médecins avaient eu accès à Internet au cours des 12 mois précédents (Comité Gestor da Internet no Brasil, 2015), mais les statistiques figurant dans la page Web de DataSUS montrent une moyenne de 210 000 accès au cours de la même année (DataSUS, 2015). Si l'on tient compte du fait que le Brésil compte plus de 200 000 professionnels administratifs dans le système de santé (MINISTÉRIO DA SAÚDE, 2016), le nombre d'accès au site Web de DataSUS est beaucoup moins grand que prévu.

1 UFJF – Programa de Pós-Graduação em Saúde Coletiva, Rua José Lourenço Kelmer, s/n. Campus Universitário – São Pedro – Juiz de Fora – Brésil, code postal : 36036-900 (rfsaldanha@gmail.com);

2 UFJF – Departamento de Estatística, Rua José Lourenço Kelmer, s/n. Campus Universitário – São Pedro – Juiz de Fora – Brésil, code postal : 36036-900 (ronaldo.bastos@ufjf.edu.br).

Suite aux efforts de collecte des données, on doit accorder davantage d'attention à la communication des données et des indicateurs de la santé. La communication de données sur la santé destinées aux décideurs devrait aller au-delà d'indicateurs et de chiffres bruts, et mettre l'accent sur le rapport entre les données actuelles et les tendances passées et d'autres indicateurs, en établissant des relations entre les sous-groupes et les régions géographiques et en utilisant des graphiques et des cartes pour aider à fournir l'information (AbouZahr, Adjei et Kanchanachitra, 2007).

Le projet actuel vise à contribuer aux besoins des gestionnaires de la santé, des chercheurs et de la population en général qui souhaitent obtenir un outil souple et facile à utiliser, qui peut tenir compte des différents aspects de la santé qui les concernent, en vue de la recherche de connaissances et de la prise de décisions. À cette fin, nous proposons la production mensuelle interactive de rapports épidémiologiques synthétiques, qui sont non seulement facilement accessibles, mais aussi faciles à interpréter et à comprendre. L'accent est mis sur la visualisation des données, grâce à des diagrammes et des cartes plus informatifs, comme ceux figurant dans la section suivante.

2. Méthodologie

La méthodologie de ce projet concerne l'acquisition des données, ainsi que le traitement et la production de rapports sur une base mensuelle, comme il est indiqué ci-après.

2.1 Données

Le ministère de la Santé du Brésil, par l'entremise de DataSUS, diffuse des données agrégées et des données brutes dans son site Web. Nous utilisons les données brutes disponibles en format DBC pour créer des rapports utilisant le niveau souhaité d'agrégation. Dans ces fichiers de données brutes, chaque enregistrement présente des renseignements sur une occurrence, comme une admission à l'hôpital, par exemple.

Les fichiers DBC sont téléchargés sur une base mensuelle, convertis en DBF et importés en langage R pour l'environnement de traitement statistique (R Core Team, 2014). Après avoir consulté les données brutes, les bases de données sont préparées et épurées en supprimant des variables sans intérêt, en convertissant des unités et en étiquetant les niveaux de variables catégoriques, selon les fichiers de documentation qui sont aussi disponibles dans le site Web de DataSUS. Toutes les étapes sont exécutées de façon automatisée, ce qui crée des ensembles de données prêts à être utilisés qui sont séparés par catégories d'intérêt.

Ces catégories d'intérêt permettent de faire une distinction entre les enregistrements selon le Code international des maladies (CIM) éclairée par les domaines de diagnostic, en suivant les mêmes 22 critères d'agrégation que ceux utilisés dans le cadre du projet sur le fardeau global de la maladie (Department of Health Statistics and Information System WHO, 2013).

À l'heure actuelle, nous avons utilisé les fichiers DBC concernant les admissions à l'hôpital, mais d'autres dimensions, comme les soins ambulatoires et la mortalité, sont aussi disponibles.

2.2 Traitement

Cette étape sert à créer des rapports pour chaque catégorie d'intérêt et niveau d'agrégation. À cette fin, quatre algorithmes ont été élaborés aux niveaux national, de l'état, macro-régional et micro-régional au Brésil.

La création de rapports en *html* est effectuée au moyen de *knitr* (Yihui, 2015) et *markdown* (Allaire et coll., 2014) de la bibliothèque R, ce qui constitue l'étape la plus longue du processus. Dans la mesure du possible, des stratégies à multiples facettes sont appliquées au moyen du logiciel *doParallel* (Revolution Analytics et Weston, 2014).

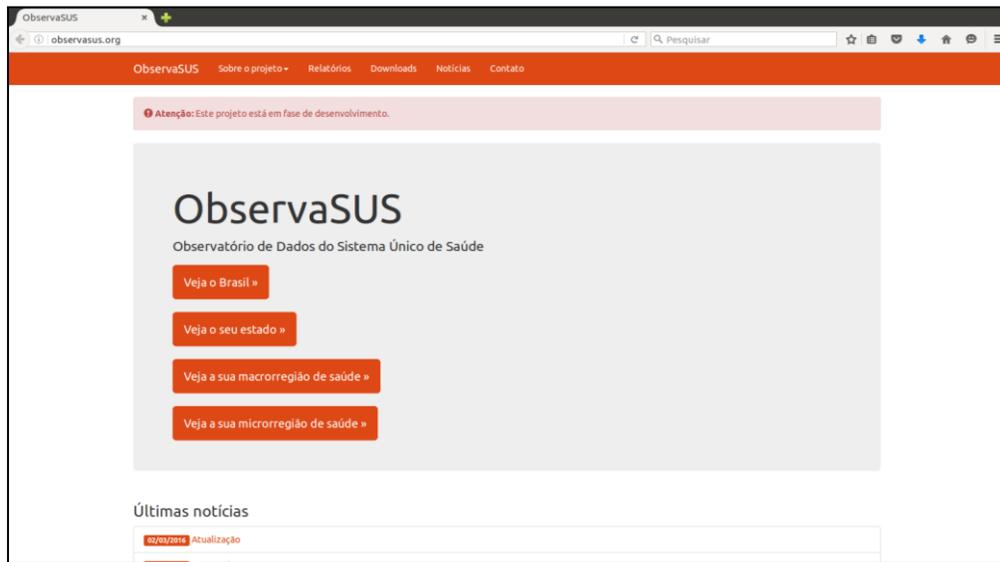
Tous les algorithmes que nous utilisons sont gérés par un répertoire *subversion*, qui permet l'élaboration, la gestion et la récupération rapides et fiables de codes pour les développeurs.

2.3 Page Web

Les rapports créés sont disponibles dans la page Web de projet (<http://observasus.org>). L'architecture du site Web a été conçue en vue d'une interface simple et directe, qui exige de l'utilisateur un niveau minimum de connaissances de l'épidémiologie.

Les rapports peuvent être consultés directement à partir de la première page (voir la figure 2.2-1). Après avoir sélectionné le niveau d'agrégation, l'utilisateur doit choisir la catégorie d'intérêt présente dans l'ensemble du rapport produit, ce qui nécessite de deux à quatre clics pour obtenir un rapport à l'écran.

Figure 2.2-1
Page principale d'ObservaSUS



Les niveaux d'agrégation pour les régions macro-économiques et micro-économiques s'accompagnent d'un lien qui aide l'utilisateur à sélectionner la région souhaitée. La même stratégie s'applique aux catégories d'intérêt, l'utilisateur pouvant indiquer le nom d'une maladie et choisir la catégorie.

Le site Web comprend d'autres sections : 1) un blogue utilisé pour informer les utilisateurs des mises à jour et des améliorations apportées; 2) une section « au sujet du projet », qui comprend des renseignements généraux, l'équipe, les technologies et la méthodologie utilisées, 3) une zone de téléchargement, où les utilisateurs peuvent télécharger les ensembles de données mentionnés dans la section 2.1, et 4) un système de suivi des questions, où les utilisateurs peuvent envoyer des suggestions et rendre compte des erreurs.

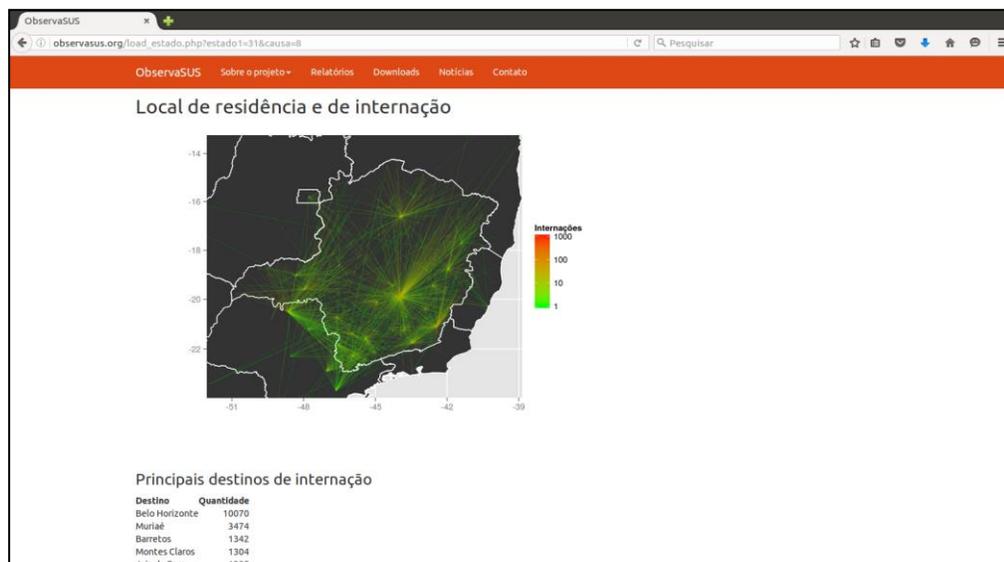
3. Résultats

Dans le cadre du projet actuel, on peut fournir des rapports d'information pour plusieurs catégories d'intérêt et niveaux d'agrégation, ce qui aide à combler les lacunes entre l'accès aux données et la production de l'information, sans se substituer au DataSUS, mais en abordant les données différemment. Un échantillon (une partie) d'un rapport typique est présenté à la figure 3-1.

D'autres progrès sont prévus à partir d'autres bases de données déjà disponibles, sur les soins ambulatoires et la mortalité, par exemple, afin de fournir un aperçu plus complet de l'état de santé au pays. L'augmentation qui découlera de la quantité d'information produite doit s'accompagner d'améliorations dans l'interface, en exigeant le minimum

de compétences de l'utilisateur pour accéder aux rapports et les comprendre pleinement. Ainsi, le projet, qui doit fournir des renseignements sur la santé dans le cadre de rapports d'information mis à jour, devrait constituer une source fiable d'information pour les décideurs et la population du Brésil.

Figure 3-1
Exemple de rapports ObservaSUS



Remerciements

Le présent projet a reçu le soutien de l'Université fédérale de Juiz de Fora, par l'entremise du Programme de deuxième cycle sur la santé publique, du Laboratoire d'information géoréférencée – LINGE du Service des statistiques et du Centre des études consultatives, de la formation et des études sur la santé – NATES.

Bibliographie

- AbouZahr, Carla, Sam Adjei et Churnrurtai Kanchanachitra. (2007). "From Data to Policy: Good Practices and Cautionary Tales." *Lancet* 369 (9566): 1039–46. doi:10.1016/S0140-6736(07)60463-2.
- Allaire, JJ, Jeffrey Horner, Vicent Marti et Natacha Porte. (2014). "Markdown: Markdown Rendering for R." Version 0.7.4 du progiciel R.
- Boerma, J Ties et Sally K Stansfield. (2007). "Health Statistics Now: Are We Making the Right Investments?" *Lancet* 369 (9563): 779–86. doi:10.1016/S0140-6736(07)60364-X.
- Comitê Gestor da Internet no Brasil. (2015). "TIC Saúde 2014: Pesquisa Sobre O Uso Das Tecnologias de Informação E Comunicação Nos Estabelecimentos de Saúde Brasileiros". São Paulo: CETIC.BR.
- DataSUS. (2015). "Estatísticas de Acesso Ao TabNET: 2014". Brasília.
- Department of Health Statistics and Information Systems WHO. (2013). "WHO Methods and Data Sources for Global Burden of Disease Estimates 2000-2011". Global Health Estimates Technical Paper WHO/HIS/HSI/GHE/2013.4. Genève.
http://www.who.int/healthinfo/statistics/GlobalDALYmethods_2000_2011.pdf?ua=1.

Handley, Kathleen, Ties Boerma, Cesar Victora et Timothy Grant Evans. (2015). "An Inflection Point for Country Health Data." *The Lancet Global Health*, no. 15: 437–38. doi:10.1016/S2214-109X(15)00067-4.

MINISTÉRIO DA SAÚDE. (2016). "DataSUS." <http://www2.datasus.gov.br>.

R Core Team. (2014). "R: A Language and Environment for Statistical Computing". Vienna: R Foundation for Statistical Computing. <http://www.r-project.org/>.

Revolution Analytics et Steve Weston. (2014). "doParallel: Foreach Parallel Adaptor for the Parallel Package." Version 1.0.8 du logiciel R.

Yihui, Xie. (2015). "Knitr: A General-Purpose Package for Dynamic Report Generation in R." Version 1.9 du logiciel R.