

MODÈLE DE SURVEILLANCE POUR L'IDENTIFICATION DE PICS SUR DONNÉES LONGITUDINALES : LE CAS DES ARRÊTS MALADIE

Tom DUCHEMIN (1,2), Angela NOUFAILY (3), Mounia N. HOCINE (2)

(1) Laboratoire MESuRS, Le Cnam, 292 Rue Saint-Martin, 75003 Paris; (2) Malakoff Humanis, 21 Rue Laffitte, 75009 Paris; (3) Warwick Medical School, Coventry, UK.

Introduction La surveillance d'épidémies de maladie infectieuse doit souvent être mise en place sur plusieurs sites pour pouvoir détecter des événements localisés (par villes ou régions par exemple). Les arrêts maladie peuvent être analysés comme des données similaires puisque le lieu de travail est un lieu de diffusion de maladies infectieuses mais aussi de troubles physiques ou mentaux. Chaque entreprise est ainsi un potentiel site de diffusion de ces troubles et l'identification des entreprises en "situation épidémique" pourrait permettre d'agir pour le bien-être des salariés.

Méthode Nous proposons une adaptation du modèle de Farrington [1], un algorithme basé sur une régression de Quasi-Poisson, pour la surveillance simultanée de plusieurs sites. L'adaptation proposée est une régression Négative Binomiale à effet mixte ajusté sur la saisonnalité, la tendance et diverses covariables. Comme dans l'algorithme original, le modèle est estimé en repondérant les observations pour prendre en compte les alertes passées et améliorer la sensibilité du modèle. La borne d'alerte du modèle est construite à partir des quantiles d'une loi Binomiale Négative. Nous avons évalué par simulations la sensibilité et la spécificité du modèle pour différents scénarios. Une application à des données d'arrêts maladie de 1376 pendant les premières semaines de l'épidémie de la COVID-19 en mars et avril 2020 est proposée.

Résultats Le modèle présente des résultats corrects en termes de Taux de Faux Positif (entre 0,015 et 0,025 selon les scénarios pour une valeur attendue de 0,025) et de Probabilité de Détection (entre 0,368 et 0,616). L'application aux données d'arrêt maladie montre que 5.9% des entreprises présentaient des volumes d'arrêts maladie anormalement hauts en Janvier 2020 et 56,8% en Mars 2020. Ces résultats ne sont pas causés par des cas avérés de COVID-19 mais principalement par des arrêts de travail nécessaires à la garde d'enfant à domicile.

Discussion L'analyse proposée permet de valider un modèle de surveillance multi-site qui pourrait être utilisé pour d'autres cas d'usages que les arrêts maladie. Notamment, des méthodes de surveillance avec effets mixtes sont utilisés dans le cadre de la surveillance syndromique à Public Health England [2] mais n'ont pas été validés et ne prennent pas en compte les améliorations de l'algorithme de Farrington (saisonnalité, sous-pondération des alertes passées).

References

- [1] C. P. Farrington, N. J. Andrews, A. D. Beale, and M. A. Catchpole. A Statistical Algorithm for the Early Detection of Outbreaks of Infectious Disease. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 159(3):547–563, 1996.
- [2] Roger A. Morbey, Alex J. Elliot, Andre Charlett, Neville Q. Verlander, Nick Andrews, and Gillian E. Smith. The application of a novel 'rising activity, multi-level mixed effects, indicator emphasis' (RAMMIE) method for syndromic surveillance in England. *Bioinformatics (Oxford, England)*, 31(22):3660–3665, November 2015.