



E01 : PAPY : Pré-Screening Automatique par Python® : une méthode informatique pour détecter les patients éligibles

Titre

Français : PAPY : Pré-Screening Automatique par Python® : une méthode informatique pour détecter les patients éligibles

Anglais : PAPY : automatique pre-screening by Python : an informatic method to select eligible patients

Auteurs

R SCHIAPPA (1), C GIROUD (2), Y CHATEAU (1), J GAL (1), E BARRANGER (3), E CHAMOREY (1)

(1) Département d'Epidémiologie, de Bioinformatique et des Données de Santé (DEBDS), Centre Antoine Lacassagne, 33 avenue de Valombrose, 06189, Nice, France

(2) Délégation de la Recherche Clinique et de l'Innovation (DéIRCI), Centre Antoine Lacassagne, 33 avenue de Valombrose, 06189, Nice, France

(3) Direction Générale, Centre Antoine Lacassagne, 33 avenue de Valombrose, 06189, Nice, France

Responsable de la présentation

Nom : SCHIAPPA

Prénom : Renaud

Adresse professionnelle : 33 Avenue de Valombrose

Code postal : 06189

Ville : Nice

Pays : France

Newsletter :

Mots clés

Français : fouille textuelle, oncologie, détection automatique, études cliniques

Anglais : text mining, oncology, automatic detection, clinical studies

Spécialité

Principale : Oncologie - Fertilité

Secondaire : Autres

Texte

Contexte et objectifs

A l'heure de la médecine personnalisée, les études cliniques sont une opportunité pour les patients et un challenge pour les investigateurs. Les critères de sélection étant de plus en plus complexes, l'identification des patients éligibles est chronophage, coûteuse et souvent infructueuse. L'informatisation des dossiers médicaux devrait permettre l'utilisation en routine d'outils automatisés de pré-screening. L'objectif de ce travail était de développer une application informatique permettant de fouiller les dossiers médicaux afin d'émettre une alerte automatique lorsqu'un patient éligible est détecté.

Méthodes

Au Centre Antoine Lacassagne (CAL) les dossiers médicaux (DM) sont informatisés et stockés dans le logiciel Clinicom®. Ils sont exportés quotidiennement sous forme de fichiers texte. PAPY (Pré-Screening par Python®), a été développé afin de récupérer ces DM et les indexer. Il va ensuite rechercher dans les DM, à l'aide d'expressions régulières, les patients éligibles. La performance de PAPY a été évaluée sur l'étude Iodine Breast : NCT02759133. Toutes les patientes venues au CAL d'avril à juillet 2018 pour un cancer du sein ont été évaluées (population source). Trois cohortes de patientes ont été constituées et comparées. Cohorte C1 : toutes les patientes éligibles évaluées manuellement sur la population source (5 minutes de relecture par dossier). Cohorte C2 : patientes incluses dans l'étude en pratique standard (patientes identifiées lors des consultations). Cohorte C3 : patientes détectées par PAPY. La performance a été évaluée par le calcul de la sensibilité, spécificité, vrais positifs, faux positifs, gain de temps.

Résultats

La population source était constituée de 184 patientes, la cohorte C1 était de 23 patientes (100%). La pratique standard (C2) inclus 15/23 (65%) patientes éligibles. PAPY (C3) identifie 26 patientes éligibles dont 14/23 (60%) appartiennent à C1 et 9/15 (60%) appartiennent à C2. 5/26 (20%) patientes de C3 n'ont pas été détectées dans C2 mais sont éligibles (C1) et 12/26 (35%) patientes sont des faux positifs. Trois (13%) des patientes de C1 ne sont ni incluses en pratique standard ni identifiées par PAPY. L'analyse montre que la pratique standard a une sensibilité de 0.65 [0.45, 0.80] et une spécificité de 100%, PAPY est proche 0.60 [0.41, 0.77] avec une spécificité de 0.92 [0.87, 0.95]. Le temps nécessaire pour analyser les 184 dossiers pour établir la cohorte C1 est de 46 heures, le temps de contrôle manuel des 23 dossiers sélectionnés par PAPY a été estimé à 1 heure 55 minutes.

Discussion et Conclusion

PAPY est proche de la pratique standard (sensibilité : 0.60 vs 0.65). Malgré 35% de faux positifs, le gain de temps reste considérable, seules les patientes détectées étant évaluées pour l'inclusion. PAPY détecte 20% de patientes éligibles mais non incluses et peut donc venir en complément de la pratique courante. PAPY se heurte aux problèmes connus des systèmes automatisés de recrutement tels que le temps de validation des comptes rendus médicaux et la disponibilité du personnel hospitalier pour les inclusions. PAPY a été utilisé en routine pour 7 études, des emails automatiques sont envoyés à chaque détection de patient éligible, 3 patients ont été inclus grâce à cette méthode. Afin de diminuer le taux de faux positifs, l'intégration des critères de non inclusion à l'aide d'expressions négatives est en cours. Une méthode de deep learning est prévue en partenariat avec l'INRIA.