



**HAL**  
open science

## Types de risque médical et leur traitement avec des méthodes de TAL

Natalia Grabar, Frantz Thiessard

► **To cite this version:**

Natalia Grabar, Frantz Thiessard. Types de risque médical et leur traitement avec des méthodes de TAL. Risque et TAL, TALN 2016 workshop, Jul 2016, Paris, France. hal-01426823

**HAL Id: hal-01426823**

**<https://hal.science/hal-01426823>**

Submitted on 4 Jan 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Types de risque médical et leur traitement avec des méthodes de TAL

Natalia Grabar<sup>1</sup> Frantz Thiessard<sup>2</sup>

(1) CNRS UMR 8163 STL, Université Lille 3, 59653 Villeneuve d'Ascq, France

(2) ERIAS, INSERM UMR 897 ISPED, Université de Bordeaux, 33076 Bordeaux, France

natalia.grabar@univ-lille3.fr, frantz.thiessard@isped.u-bordeaux2.fr

## RÉSUMÉ

---

L'objectif de ce travail est d'étudier le risque médical et la manière dont il est traité avec les méthodes de TAL. Le domaine médical est en effet intimement lié avec le risque. D'une part, la médecine a la vocation de traiter les patients qui présentent des risques multiples de développer des maladies. Toutes les étapes de prise en charge sont entachées d'un degré plus ou moins important d'incertitude : incertitude diagnostique, incertitude de la réaction individuelle au traitement, incertitude de l'évolution, etc. D'autre part, le processus de soins médicaux peut également générer le risque chez les patients. Nous proposons une classification de risques médicaux. Nous analysons aussi plus en détail certains types de risques, en particulier ceux qui sont traités avec des méthodes de TAL.

## ABSTRACT

---

### Types of medical risk and their processing with the NLP methods

The aim of this study is the medical risk and the way it is processed by NLP methods. The medical field is indeed widely concerned by risk. On one hand, the purpose of medicine is to treat patients who present multiple risks to develop pathologies. All steps of healthcare process are prone to a certain degree of uncertainty: diagnosis, individual reaction to treatment, evolution, etc. On the other hand, the healthcare process can also generate the risk for patients. We propose a classification of the medical risk. We also analyze with more detail some types of risk, and particularly those which are processed with the NLP methods.

---

**MOTS-CLÉS** : Risque médical, méthodes de TAL, classification.

**KEYWORDS**: Medical risk, NLP methods, classification.

---

## 1 Contexte

La médecine doit traiter et prendre en compte différents types de situations à risque, émanant essentiellement du patient et de son environnement ou encore du processus de soins. Le risque, y compris médical, est défini comme un événement dommageable dont la survenance est incertaine (Cornu, 2007) et se trouve donc associé à une notion d'incertitude (Mayaux, 2011). De fait, le risque médical tient à différents facteurs et à leurs combinaisons. L'objectif de notre travail est double : (1) modéliser la notion du risque médical (section 2), (2) étudier quelle est la contribution des méthodes de TAL pour appréhender le risque médical et à quels moments cela intervient, à savoir pour la détection, la prévention ou la gestion du risque (section 3). Le matériel nécessaire à ce travail provient essentiellement de la littérature scientifique de la base bibliographique MedLine (NLM, 2015) et des

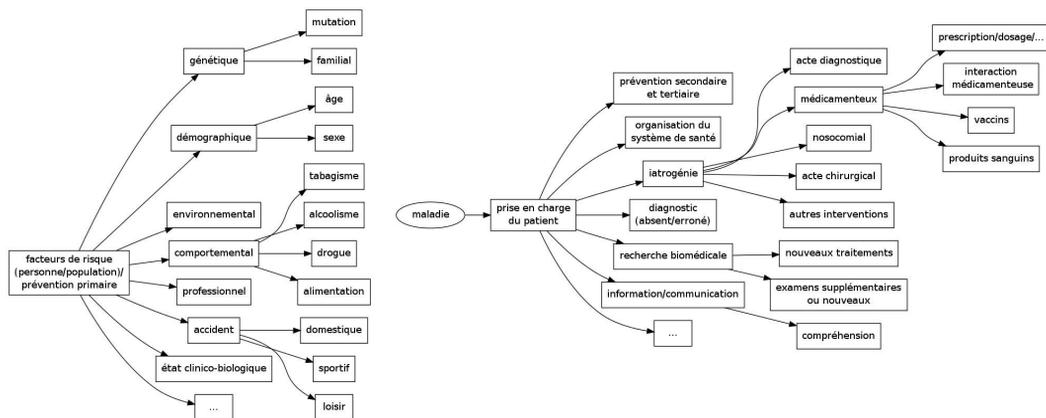


Fig. 1 – Types de risques en médecine pour le patient

informations légales du code de la santé publique. Nous proposons d’abord de cerner la notion du risque médical (section 2) et analysons quelques types de risque pour lesquels les méthodes de TAL sont exploitées (section 3). Nous proposons ensuite une conclusion à ce travail (section 4).

## 2 Cerner le risque médical

À la figure 1, nous proposons un schéma général de risques en médecine du point de vue du patient. Dans la partie gauche du schéma, il s’agit de risques existant lorsque la maladie n’est pas encore détectée. Les risques sont alors dus aux facteurs de risque du patient ou bien à la prévention primaire, qui elle aussi peut générer des risques chez le patient. Lorsque la maladie se déclare, suite aux risques de la partie gauche, d’autres risques apparaissent. Ainsi, dans la partie droite du schéma, les risques peuvent intervenir suite à une maladie déjà déclarée. Ils peuvent alors être causés par le système ou le processus de santé, la prévention secondaire ou tertiaire, l’état de la recherche biomédicale, le diagnostic ou l’information aux patients.

Nous détaillons ici quelques risques plus ou moins élémentaires (deuxième niveau de profondeur dans le schéma), que nous distinguons en fonction de leurs sources :

- le risque génétique est inhérent au patient et provient de son patrimoine génétique et de l’histoire familiale. Par exemple, dans une famille où plusieurs cas de maladies cardiaques ou de cancers sont connus, il peut être nécessaire de surveiller les descendants pour ces maladies ;
- le risque démographique est lié aux caractéristiques démographiques du patient, comme par exemple son âge ou sexe. Il existe par exemple des risques et pathologies plus fréquentes chez les personnes âgées (chutes, maladies neurodégénératives...) ou chez les enfants (angines, otites) ;
- le risque environnemental provient de l’environnement du patient, qu’il s’agisse de l’environnement géographique ou social. Par exemple, certaines pathologies sont associées avec des zones géographiques : paludisme, méningite, rage, fièvre jaune en Afrique ; dengue, encéphalite, leptospirose, fièvre thyphoïde en Asie, etc. Par ailleurs, la pollution environnementale (air, eau...) fait également partie de ce type de risque ;
- le risque comportemental est inhérent au mode de vie d’une personne. Parmi ses facteurs, on peut

indiquer le tabagisme, l'alcoolisme et l'usage de drogues mais aussi le type d'alimentation. Ces facteurs sont dommageables pour la personne mais aussi parfois pour sa descendance ;

- le risque professionnel est associé au métier d'une personne et à son exposition à différents facteurs nuisibles : bruit, lumière, produits chimiques, radiation, animaux, ordinateur, posture debout ou assise, soulèvement de poids lourds, etc. ;
- les accidents sont des événements dommageables imprévus qui correspondent aux aléas de la vie. Indiquons par exemple des accidents domestiques (blessures avec les ustensiles de cuisine, empoisonnement alimentaire, chutes de personnes âgées...), sportifs (chutes de ski ou de vélo, élongations et ruptures de ligaments...), et loisir (noyades, accidents de voiture...);
- l'état clinico-biologique du patient, et surtout lorsque le patient montre des valeurs anormales d'examen, indiquent le risque potentiel à présenter ou développer une maladie. Par exemple, la haute tension artérielle, la taux de sucre, la surcharge pondérale peuvent être indicateurs de maladies qui se développent ;
- le risque lié aux traitements et aux soins médicaux peuvent être provoqués par le processus de soins ou le système de santé. Il s'agit typiquement de risques qui sont gérés par le code de la santé publique. On distingue ainsi des risques iatrogènes, car tout traitement est entaché d'une incertitude du résultat sur un plan individuel, ou encore des risques provenant de la prévention et de l'organisation du système de santé ;
- le risque diagnostic apparaît lorsque le diagnostic du patient n'a pas pu être fait ou bien a été mal fait. Dans les deux cas, un traitement approprié ne peut pas lui être proposé, ce qui peut mener à l'aggravation de son état de santé ;
- le risque lié à la recherche biomédicale est essentiellement causé par exemple par les traitements ou les examens nouveaux et pas totalement éprouvés auprès d'une large population ;
- l'information du patient et la communication entre le patient et le médecin occupent une place de plus en plus importante dans le processus de santé. Entre autres, la compréhension des informations de santé et l'éducation thérapeutique des patients apparaissent comme des facteurs importants menant à une meilleure communication entre les patients et les professionnels de santé et donc à un processus de santé plus réussi.

Il va de soi que ces différents types de risque peuvent se combiner et interagir, ce qui apporte un degré supplémentaire d'incertitude (Mayaux, 2011). Par ailleurs, le risque médical est une notion qui évolue en fonction (1) de la connaissance biomédicale sur les pathologies et les traitements et surtout (2) de l'état du patient.

Notons que le personnel de santé peut aussi encourir des risques, comme par exemple le risque nosocomial (dû au cotoyement des patients malades à l'hôpital), infectieux (contamination au travers des aiguilles ou du matériel utilisés), radiation, blessures, etc. Ces risques peuvent être corrélés aux risques professionnels du métier de santé.

### 3 Méthodes de TAL pour appréhender le risque médical

La requête MedLine "(*risk*) and (*nlp or natural language processing*)" appliquée aux mots-clés ou au texte brut a été lancée le 5 mars 2016. Elle a permis de collecter 197 références bibliographiques. Parmi ces références, 87 ne sont pas pertinentes car *NLP* peut y correspondre à différentes notions autres que *Natural Language Processing*, comme par exemple *no light perception*, *nodular lymphocyte predominance*, *normolipidemia*, *no late presentation*, *nanoliposome*, *nonlinear programming*. Dans d'autres références, si les méthodes de TAL sont impliquées, le travail ne concerne pas forcément le

traitement du risque médical.

Parmi les références pertinentes et par rapport à la figure 1, nous pouvons proposer cette analyse :

- une grande partie de références (n=66) correspond au risque clinique et plus spécifiquement au risque de développer une maladie étant donné une autre maladie, des signes cliniques, des facteurs de risque existants ou des résultats d'examens,
- une autre catégorie importante couvre les risques liés aux traitements (n=24). Il s'agit le plus souvent de risques médicamenteux, de radiologie ou des infections nosocomiales,
- le risque induit par le comportement (n=9) est aussi présent avec par exemple l'étude du tabagisme, de la consommation de drogues ou l'impact de la vie dans la rue,
- le risque génétique (n=7) est aussi étudié, grâce à l'étude des données génétiques et phénotypiques des patients et en relation avec des pathologies données,
- quelques références (n=4) sont dédiées à l'étude de l'information pour les patients et de leur communication avec les médecins.

Quasiment toutes les références concernent la détection du risque. Il s'agit de tester des méthodes automatiques afin de prédire le risque. Les méthodes de TAL sont alors utilisées de différentes manières : le plus souvent pour extraire les informations pertinentes pour faire une prédiction, mais aussi pour catégoriser les données, ou encore pour construire des terminologies et ontologies dans l'objectif de traiter le risque. Une seule référence, liée à l'information et communication aux patients, propose d'aider la compréhension de documents de santé grâce à des services qui ajoutent des définitions aux documents cliniques et aux documents en ligne (Topac & Stoicu-Tivadar, 2013).

Il faut aussi ajouter que, si la requête utilisée couvre plusieurs références pertinentes, il existe plusieurs autres références qui ne répondent pas à cette requête mais qui traitent néanmoins du risque médical. Ces références peuvent comporter des termes plus précis et relatifs à un type de risque donné (*pharmacovigilance, nosocomial, interactions médicamenteuses*, etc.). L'explosion du concept *risk* du MESH (NLM, 2001) ne permettrait pas de retrouver ces références car il ne subsume par ce types de termes. Dans cette situation typiquement, nous pouvons observer que le risque médical n'est pas systématiquement indexés dans les références.

Nous discutons maintenant deux situations plus spécifiques du risque médical, qui sont fréquentes dans la pratique médicale et qui profitent déjà des méthodes de TAL : les risques liés aux traitements médicaux (section 3.1) et les facteurs de risque (section 3.2).

### 3.1 Risques liés aux traitements médicaux et au séjour à l'hôpital

Tout acte médical comporte une part de risque pouvant aboutir à une inefficacité ou à des effets indésirables. Le facteur de l'incertitude y joue un rôle important et dépend par exemple des connaissances médicales du moment, de la réaction métabolique et des facteurs de risque du patient et de bien d'autres choses potentiellement inconnues du prescripteur. Il serait difficile de mettre sur le médecin la responsabilité totale des facteurs qui dépassent sa compétence : le droit commun a ainsi mis à la charge du médecin une obligation de moyens, et non pas du résultat : le médecin est ainsi tenu de faire ce qui est en son pouvoir pour soigner au mieux le patient<sup>1</sup>. Si le risque est inhérent

---

1. Selon l'article L. 1142-1, I du code de la santé publique, en dehors du cas où la responsabilité du médecin serait encourue en raison du défaut d'un produit de santé, les professionnels de santé ne sont responsables des conséquences dommageables d'acte de prévention, de diagnostic ou de soin qu'en cas de faute. L'obligation de moyen montre qu'un patient ne peut attendre de son médecin un résultat déterminé, sauf avis contraire exprimé par celui-ci, compte tenu de l'aléa que comporte un acte médical

à la pratique médicale, il existe aussi des moyens pour leur prévention. En France, les risques dus aux soins sont gérés par le code de la santé publique : il existe des articles pour la prévention des affections nosocomiales et iatrogènes et pour les actions relatives à la politique du médicament et aux dispositifs médicaux stériles (articles R6111-1 à R6111-11). De même, le remboursement et la prise en charge de ces événements sont gérés par le code de la santé publique.

Les risques liés aux traitements médicaux ou aux séjours dans l'hôpital, communément associés aux erreurs médicales, occupent une place importante dans les travaux scientifiques (La Pietra *et al.*, 2005). Un modèle des erreurs est proposé et distingue l'erreur humaine (l'erreur qui provient du médecin, du patient ou de la source externe) et l'erreur due à la structure ou au processus (l'erreur provient des protocoles, de l'organisation ou des connaissances du moment). Il faut noter que les prescriptions de médicaments (*e.g.* analgésiques, sédatifs, chimiothérapie, anticoagulants, cardiovasculaires) sont une source fréquente du risque (Bates *et al.*, 1995; Ridley *et al.*, 2004; Rolland, 2004). Il a été également noté que certains risques et erreurs peuvent être prévenus et évités (Weingart *et al.*, 2000; Wilson *et al.*, 1995). Il existe ainsi plusieurs modèles classiques qui proposent des actions à entreprendre pour éviter ce type de risque, comme le modèle de Reason (Vincent *et al.*, 1998), HFMEA (Healthcare Failure Mode and Effects Analysis) (DeRosier *et al.*, 2002) ou des checklists (Heller *et al.*, 2016; Jones *et al.*, 2016).

L'informatisation des hôpitaux et la création de méthodes automatiques fournissent d'autres moyens pour réduire le risque. Il est estimé que jusqu'à 50 % d'erreurs de prescription peuvent alors être réduites avec un système analysant automatiquement les ordonnances (Koren, 2002). Le contrôle du processus hospitalier peut intervenir à différents moments grâce aux systèmes de prescription assistés par ordinateur et aux alertes automatiques (Tamblyn *et al.*, 2003; Horsky *et al.*, 2004; Boussadi *et al.*, 2008), au contrôle du code-barre des médicaments (Wimmer *et al.*, 2015; Bowers *et al.*, 2015), à la détection de valeurs anormales de laboratoire (Deleger *et al.*, 2014; Wang *et al.*, 2015; Hoogendoorn *et al.*, 2016), à la pharmacovigilance (Meyboom *et al.*, 2002; Bousquet *et al.*, 2005; Trifirò *et al.*, 2009; Dupuch & Grabar, 2015), à l'étude de l'interaction médicamenteuse (Duda *et al.*, 2005; Ayvaz *et al.*, 2015; Liu *et al.*, 2016) ou entre les médicaments et les aliments (Linnoila, 1973; Greenblatt & Derendorf, 2013; Duke Med Health News, 2013). Les outils automatiques permettent en effet de prévenir certains risques et d'améliorer la sécurité des patients. Tout type de documents cliniques peuvent contribuer à la détection de ce types de risque (prescriptions, comptes-rendus, résumés...), mais ils peuvent aussi générer d'autres types de risques, tels que l'e-iatrogénie (Battles & Keyes, 2002).

### 3.2 Facteurs de risque

Comme l'indique la figure 1, les facteurs de risque désignent un ensemble de facteurs orthogonaux comme le comportement, environnement, pathologies existantes ou le patrimoine génétique, considérés de manière large, qui augmentent la probabilité pour une personne de manifester une maladie donnée. La découverte de facteurs de risque permet de mettre en place des stratégies de prévention. Il s'agit d'un défi biomédical important. Plusieurs travaux de recherche existants sont effectués avec des données provenant de différents domaines (biologie, épidémiologie, santé publique) et exploitent différentes méthodes, dont les méthodes de TAL. Cependant, malgré les recherches existantes, les facteurs de risque traditionnels pour les maladies coronaires par exemple ne couvrent que 50 % de risques existants (Allen, 2000). La majorité de travaux en médecine exploitent des méthodes statistiques sur de grandes populations pour détecter les facteurs de risque de plusieurs pathologies et

conditions : maladie cardiovasculaires (Gouni-Berthold *et al.*, 2009; Perret-Guillaume *et al.*, 2009), hypertension (Meneton *et al.*, 2008), athérosclérose (Dentali *et al.*, 2009; Portugal *et al.*, 2009), cancers (Turnbull & Hodgson, 2005; Fleshner & Lawrentschuk, 2009), mortalité de la population âgée (Glynn *et al.*, 1995; Bennett, 2002; Ahmad & Bath, 2005).

Il existe aussi quelques travaux en fouille de textes et de données. Une expérience en fouille de données a été effectuée avec les codes de la CIM-9 pour prédire les groupes de personnes avec des facteurs de risque similaires (Cerrito, 2004), ou avec les données des compagnies d'assurance pour estimer le coûts des remboursements (Kolyshkina & van Rooyen, 2006). Notons que le défi KDD<sup>2</sup>, qui a eu lieu en 2004, était dédié à la fouille de données biomédicales associées aux groupes sociaux de patients. L'objectif principal était alors d'identifier les facteurs de risque d'athérosclérose et de contrôler l'évolution de ces facteurs de risque et leur impact (Brisson *et al.*, 2004).

Il existe aussi quelques travaux de TAL sur des documents non structurés, essentiellement la littérature scientifique : (1) combinaison de méthodes manuelles et automatiques pour l'extraction de facteurs de risque liés au cancer du sein (Blake, 2004). Les résultats sont constants avec les travaux publiés, car ils montrent une association positive avec la consommation d'alcool et une association négative avec l'ancien tabagisme ; (2) méthodes automatiques à base de patrons et d'association de mots-clés appliquées à tout type de risque et évaluées sur le risque cardiovasculaire (Hamon *et al.*, 2010). Là aussi, les résultats sont cohérents avec l'état de connaissance actuel et indiquent plusieurs facteurs de risque connus, comme *hypertension*, *tabagisme*, *diabète*, *âge*, *obésité*, *hypercholestérolémie*, *hyperlipidémie* ou *histoire familiale*.

## 4 Conclusion et Perspectives

Nous avons présenté une analyse du risque médical. D'une part, nous avons proposé une modélisation du risque médical, en distinguant les types de risques selon leurs sources. D'autre part, nous avons analysé la littérature scientifique, en mettant l'accent sur deux types de risque : les risques dus aux traitements médicaux et au séjour à l'hôpital et les facteurs de risque. Dans les deux cas, il existe en effet plusieurs travaux de TAL qui visent essentiellement la détection du risque, mais aussi sa prévention grâce aux systèmes de prescription automatique, par exemple. Nous avons aussi observé que l'indexation du risque médical n'est pas effectuée de manière systématique dans la base bibliographique biomédicale MedLine.

La détection, mais surtout la prévention et la gestion du risque médical, est une problématique de recherche très importante. En effet, tous les patients sont potentiellement vulnérables face au risque médical. Comme le risque représente un certain coût humain, économique et social, cela donne la motivation principale pour les travaux de recherche. La réduction du risque médical concerne en premier lieu la sécurité des patients et la qualité des soins médicaux qui leurs sont dispensés.

Le travail proposé ici peut être enrichi en suivant plusieurs directions. Une étude plus systématique de la littérature scientifique et de l'activité des sociétés savantes et d'intérêt public qui oeuvrent pour la sécurité des patients et des soins médicaux pourra être effectuée. Comme nous l'avons vu, il est aussi nécessaire de cibler les types de risque précisément afin d'accéder aux travaux pertinents. Cela représente aussi une orientation future de notre travail. Finalement, nous avons vu que, dans la variété de risques médicaux existants, plusieurs types de risque sont peu ou pas traités avec les méthodes de

TAL. Cela ouvre donc plusieurs autres pistes d'investigation.

## Références

- AHMAD R. & BATH P. A. (2005). Identification of risk factors for 15-year mortality among community-dwelling older people using Cox regression and a genetic algorithm. *Journal of Gerontology*, **60**(8), 1052–8.
- ALLEN J. (2000). Genetics and cardiovascular disease. *Nurs Clin North Am*, **35**(3), 653–62.
- AYVAZ S., HORN J., HASSANZADEH O., ZHU Q., STAN J., TATONETTI N., VILAR S., BROCHHAUSEN M., SAMWALD M., RASTEGAR-MOJARAD M., DUMONTIER M. & BOYCE R. (2015). Toward a complete dataset of drug-drug interaction information from publicly available sources. *J Biomed Inform*, **55**, 206–17.
- BATES D., CULLEN D., LAIRD N., PETERSEN L., SMALL S., SERVI D., LAFFEL G., SWEITZER B., SHEA B. & ET AL R. H. R. (1995). Incidence of adverse drug events and potential adverse drug events. *JAMA*, **274**, 29–34.
- BATTLES J. & KEYES M. (2002). Technology and patient safety: a two-edged sword. *Biomed Instrum Technol*, **36**, 84–8.
- BENNETT K. (2002). Low-level social engagement as a precursor of mortality among people in later life. *Age Ageing*, **31**, 165–8.
- BLAKE C. (2004). A text mining approach to enable detection of candidate risk factors. In *Medinfo*, p. 1528–1528.
- BOUSQUET C., HENEGAR C., LILLO-LE LOUËT A., DEGOULET P. & JAULENT M. (2005). Implementation of automated signal generation in pharmacovigilance using a knowledge-based approach. *Int J Med Inform*, **74**(7-8), 563–71.
- BOUSSADI A., BOUSQUET C., SABATIER B., COLOMBET I. & DEGOULET P. (2008). Specification of business rules for the development of hospital alarm system: application to the pharmaceutical validation. In *Stud Health Technol Inform*, p. 145–50.
- BOWERS A., GODA K., BENE V., SIBILA K., PICCIN R., GOLLA S., DANI F. & ZELL K. (2015). Impact of bar-code medication administration on medication administration best practices. *Comput Inform Nurs*, **33**(11), 502–8.
- BRISSON L., PASQUIER N., HEBERT C. & COLLARD M. (2004). Hasar: mining sequential association rules for atherosclerosis risk factor analysis. In *PKDD 2004*, p. 14–25, Pisa Italy.
- CERRITO P. (2004). Inside text mining. *Health management technology*, **25**(3), 28–31.
- CORNU G. (2007). *Vocabulaire juridique*. Paris: PUF, Association Henri Capitant.
- DELEGER L., BRODZINSKI H., ZHAI H., LI Q., LINGREN T., KIRKENDALL E., ALESSANDRINI E. & SOLTI I. (2014). Developing and evaluating an automated appendicitis risk stratification algorithm for pediatric patients in the emergency department. *J Am Med Inform Assoc*, **20**(e2), 212–20.
- DENTALI F., SQUIZZATO A. & AGENO W. (2009). The metabolic syndrome as a risk factor for venous and arterial thrombosis. *Semin Thromb Hemost*, **35**(5), 451–7.
- DEROSIER J., STALHANDSKE E., BAGIAN J. & NUDELL T. (2002). Using health care failure mode and effect analysis: the va national center for patient safety's prospective risk analysis system. *Jt Comm J Qual Improv*, **5**, 248–67.

- DUDA S., ALIFERIS C., MILLER R., SLATNIKOV A. & JOHNSON K. (2005). Extracting drug-drug interaction articles from Medline to improve the content of drug databases. In *AMIA Symp*, p. 216–20.
- DUKE MED HEALTH NEWS (2013). Grapefruit: enemy of many medications. in some patients, the interaction of fruit and drug may put their life and health at risk. *Duke Med Health News*, **19**(2), 1–2.
- DUPUCH M. & GRABAR N. (2015). Semantic distance-based creation of clusters of pharmacovigilance terms and their evaluation. *J Biomed Inform*, **54**, 174–185.
- FLESHNER N. & LAWRENTSCHUK N. (2009). Risk of developing prostate cancer in the future: overview of prognostic biomarkers. *Urology*, **73**(5), 21–7.
- GLYNN R., FIELD T., ROSNER B., HEBERT P., TAYLOR J. & HENNEKENS C. (1995). Evidence for a positive linear relation between blood pressure and mortality in elderly people. *Lancet*, **345**(8953), 825–9.
- GOUNI-BERTHOLD I., KRONE W. & BERTHOLD H. (2009). Vitamin d and cardiovascular disease. *Curr Vasc Pharmacol*, **7**(3), 414–22.
- GREENBLATT D. & DERENDORF H. (2013). Grapefruit-medication interactions. *CMAJ*, **185**(6), 507.
- HAMON T., GRAÑA M., RAGGIO V., GRABAR N. & NAYA H. (2010). Identification of relations between risk factors and their pathologies or health conditions by mining scientific literature. In *Medinfo*, p. 964–8.
- HELLER S., REZAPOOR M. & PARVIZI J. (2016). Minimising the risk of infection: a peri-operative checklist. *Bone Joint J*, **98-B**(1 Suppl A), 18–22.
- HOOGENDOORN M., SZOLOVITS P., MOONS L. & NUMANS M. (2016). Utilizing uncoded consultation notes from electronic medical records for predictive modeling of colorectal cancer. *Artif Intell Med*.
- HORSKY J., KAUFMAN D. & PATEL V. (2004). Computer-based drug ordering: evaluation of interaction with a decision-support system. *Stud Health Technol Inform*, **107**(2), 1063–7.
- JONES E., LEES N., MARTIN G. & DIXON-WOODS M. (2016). How well is quality improvement described in the perioperative care literature? a systematic review. *Jt Comm J Qual Patient Saf*, **42**(5), 196–216.
- KOLYSHKINA I. & VAN ROOYEN M. (2006). *Text mining for insurance claim cost prediction*, In *Date Mining LNAI 3755*, p. 192–202. Springer-Verlag.
- KOREN G. (2002). Trends of medication errors in hospitalised children. *J Clin Pharmacol*, **42**, 707–10.
- LA PIETRA L., CALLIGARIS L., MOLENDINI L., QUATTRIN R. & BRUSAFERRO S. (2005). Medical errors and clinical risk management: state of the art. *Acta Otorhinolaryngol Ital.*, **25**(6), 339–346.
- LINNOILA M. (1973). Drug interaction on psychomotor skills related to driving: hypnotics and alcohol. *Ann Med Exp Biol Fenn*, **51**(3), 118–24.
- LIU S., TANG B., CHEN Q. & WANG X. (2016). Drug-drug interaction extraction via convolutional neural networks. *Comput Math Methods Med*.
- MAYAUX L. (2011). *Les grandes questions du droit des assurances*. Paris: Lextenso.

- MENETON P., GALAN P., BERTRAIS S., HEUDES D., HERCBERG S. & MÉNARD J. (2008). High plasma aldosterone and low renin predict blood pressure increase and hypertension in middle-aged caucasian populations. *J Hum Hypertens*, **22**(8), 550–8.
- MEYBOOM R., LINDQUIST M., EGBERTS A. & EDWARDS I. (2002). Signal selection and follow-up in pharmacovigilance. *Drug Saf*, **25**(6), 459–465.
- NLM (2001). *Medical Subject Headings*. National Library of Medicine, Bethesda, Maryland. [www.nlm.nih.gov/mesh/meshhome.html](http://www.nlm.nih.gov/mesh/meshhome.html).
- NLM (2015). *Medline: medical literature on-line*. National Library of Medicine, Bethesda, Maryland. [www.ncbi.nlm.nih.gov/sites/entrez](http://www.ncbi.nlm.nih.gov/sites/entrez).
- PERRET-GUILLAUME C., JOLY L. & BENETOS A. (2009). Heart rate as a risk factor for cardiovascular disease. *Prog Cardiovasc Dis*, **52**(1), 6–10.
- PORTUGAL L., FERNANDES L. & ALVAREZ-LEITE J. (2009). Host cholesterol and inflammation as common key regulators of toxoplasmosis and atherosclerosis development. *Expert Rev Anti Infect Ther*, **7**(7), 807–19.
- RIDLEY S., BOOTH S. & ET AL. C. T. (2004). Prescription errors in uk critical care units. *Anaesthesia*, **59**, 1193–200.
- ROLLAND P. (2004). Occurrence of dispensing errors and efforts to reduce medication errors at the central arkansas veteran's healthcare system. *Drug Saf*, **27**, 271–82.
- TAMBLYN R., HUANG A., PERREAULT R., JACQUES A., ROY D., HANLEY J., MCLEOD P. & LAPRISE R. (2003). The medical office of the 21st century (MOXXI): effectiveness of computerized decision-making support in reducing inappropriate prescribing in primary care. *CMAJ*, **169**(6), 549–56.
- TOPAC V. & STOICU-TIVADAR V. (2013). Patient empowerment by increasing the understanding of medical language for lay users. *Methods Inf Med*, **52**(5), 454–62.
- TRIFIRÒ G., PARIENTE A., COLOMA P., KORS J., POLIMENI G., MIREMONT-SALAMÉ G., CATANIA M., SALVO F., DAVID A., MOORE N., CAPUTI A., STURKENBOOM M., MOLOKHIA M., HIPPISELY-COX J., ACEDO C., VAN DER LEI J. & FOURRIER-REGLAT A. (2009). Eu-adr group. data mining on electronic health record databases for signal detection in pharmacovigilance: which events to monitor? *Pharmacoepidemiol Drug Saf*, **18**(12), 1176–84.
- TURNBULL C. & HODGSON S. (2005). Genetic predisposition to cancer. *Clin Med*, **5**(5), 491–8.
- VINCENT C., TAYLOR-ADAMS S. & STANHOPE N. (1998). Framework for analysing risk and safety in clinical medicine. *BMJ*, **316**(7138), 1154–7.
- WANG Y., LUO J., HAO S., XU H., SHIN A., JIN B., LIU R., DENG X., WANG L., ZHENG L., ZHAO Y., ZHU C., HU Z., FU C., HAO Y., ZHAO Y., JIANG Y., DAI D., CULVER D., ALFREDS S., TODD R., STEARNS F., SYLVESTER K., WIDEN E. & LING X. (2015). Nlp based congestive heart failure case finding: A prospective analysis on statewide electronic medical records. *Int J Med Inform*, **84**(12), 1039–47.
- WEINGART S., WILSON R., GIBBERD R. & HARRISON B. (2000). Epidemiology of medical errors. *BMJ*, **320**, 774–7.
- WILSON R., RUNCIMAN W., GIBBERD R., HARRISON B., NEWBY L. & HAMILTON J. (1995). The quality in australian health care study. *Med J Aust*, **163**, 6–75.
- WIMMER S., NEUBERT A. & RASCHER W. (2015). The safety of drug therapy in children. *Dtsch Arztebl Int*, **112**(45), 781–7.