Journal of Biomedical Informatics 44 (2011) 545–552

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Occupational exposome: A network-based approach for characterizing Occupational Health Problems

Laurie Faisandier [a,*], Vincent Bonneterre [a,b], Régis De Gaudemaris [a,b], Dominique J. Bicout [c,*]

[a] Laboratoire Environnement et Prédiction de la Santé des Populations, UMR CNRS 5525 Université Joseph Fourier, Domaine de la Merci, 38 706 La Tronche, France
[b] Service de Médecine et Santé au Travail, Centre Hospitalier Universitaire Grenoble, BP217, 38 043 Grenoble, France
[c] Biomathématiques et Epidémiologie, Laboratoire Environnement et Prédiction de la Santé des Populations-TIMC, UMR CNRS 5525 Université Joseph Fourier, VetAgro Sup, Campus Vétérinaire de Lyon, 1 avenue Bourgelat, 69280 Marcy l'Etoile, France

ARTICLE INFO

Article history:
Received 18 April 2010
Available online 27 February 2011

Keywords:
Occupational disease
Multi-exposure
Network

ABSTRACT

Surveillance of work-related diseases and associated exposures is a major issue of public health, in particular for identifying and preventing new threats for health. In the occupational health context, the French national occupational disease surveillance and prevention network (RNV3P) have constructed a growing database that records every year all Occupational Health Problems (OHPs) diagnosed by a network of physician specialists. The network aims to provide and develop an expertise on the disease–exposure relationships, and uses the RNV3P database for developing the surveillance of OHPs and for the detection of emerging associations between diseases and occupational exposures. In this paper, we have developed the theoretical framework of the occupational exposome, defined as a network of OHPs linked by similar occupational exposures, as a novel approach which allows to characterize and to analyze the disease–exposure associations reported in the RNV3P database in the form of a relational network. Next, the occupational exposome is structured in terms of occupational exposure groups which constitute informative sub-sets of hazards considered as the backbone tree spectrum of the occupational exposures potentially related to a disease.

To illustrate the wide possibilities of this method, the exposome approach is applied to the RNV3P database's sample of Non-Hodgkin Lymphomas (NHLs). As a result, we found that the NHL occupational exposome could be described in terms of 86 embedded exposure groups, defined as a set of OHPs sharing at least one component of the occupational multi-exposure. For example, "organic solvents and thinners" is the most represented hazards related to NHLs, but is also co-associated to "benzene", "ionizing radiations" or "agricultural products". From the knowledge stored in the database by physician experts, the occupational exposome constitutes a decisive step towards the evolving monitoring of multi-exposure associated to a given disease.

1. Introduction

Studies on health at work are designed primarily to identify occupational hazards, validate new atmospheric or biological indicators of exposure and effect, assess the impact of the occupational environment on health, gain information on the etiology of diseases and improve means of preventions. Surveillance of work-related diseases and exposures is a major issue of public health, in particular for identifying and preventing new threats for health. In such a context, surveillance conventionally involves in specific cohort and epidemiological follow-up of indicators (accidents and occupational diseases), often developed from surveillance networks which feed databases. An illustrative example of such a network is The Health and Occupational Reporting network (THOR) in the United Kingdom [1–3].

In France, the national occupational disease surveillance and prevention network (Réseau National de Vigilance et de Prévention des Pathologies Professionnelles, RNV3P) was set up in 2001 as a nation-wide network of experts from occupational health consultation centers in University Hospitals of metropolitan France that records in a structured and standardized RNV3P database all patient cases diagnosed with an Occupational Health Problem (OHP), i.e. patients with diseases potentially related to occupational exposures. The RNV3P network originated from the need of developing a common expertise on health problems by bringing together scientists and researchers working in multidisciplinary areas of health. Among others objectives, the network aims to provide and develop an expertise on the disease–exposure relationships, and uses the RNV3P database for developing the

* Corresponding authors. Fax: +33 4 76 76 89 10.
E-mail addresses: LFaisandier@chu-grenoble.fr (L. Faisandier), VBonneterre@chu-grenoble.fr (V. Bonneterre), RDegaudemaris@chu-grenoble.fr (R. De Gaudemaris), Dominique.Bicout@imag.fr, d.bicout@vetagro-sup.fr (D.J. Bicout).

1532-0464/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jbi.2011.02.010

surveillance of OHPs and for the detection of emerging associations between diseases and occupational exposures. To address these issues, a systematic data-mining approach based on statistical tests of disproportionate as used in pharmacovigilance [4,5] was applied to RNV3P database to generate pre-alerts of potentially emergent disease–exposure couples. Although very sensitive, this method appears to be less specific and fails to simultaneous handle all components of the exposure associated with a disease.

Following these first studies on the RNV3P database, we began to develop the occupational exposome approach [6,7], as an alternative and complementary approach to pharmacovigilance methods involving disease–single exposure couples, that incorporates all dimensions of the composite occupational exposures in analyses of the RNV3P database. Our aim is to develop an approach allowing to investigate characteristics or traits which gather or separate OHPs as many individual factors and complex occupational situations, where exposures of diverse origins and variable intensity over time, are combined in effects on health. In this framework, we conceptualized the occupational exposome as a network of OHPs sharing components of the set of occupational exposures. Such an occupational exposome will make sense only when dealing with diseases associated with multi-exposures, i.e., associated with more than one hazard.

Similar conceptual approaches can be found in literature including the wording exposome which had been already introduced by Wild [8]. In contrast to the occupational exposome, the Wild's exposome represents the collection and succession of individual and environmental exposures encountered during an individual's lifetime. The author proposes reconstructing for a given individual, an exposure network in order to better understanding the role of each exposure and thus generate research hypotheses for the disease etiology. Although this requires a certain amount of information on traceability of exposures, this exposome view lays the groundwork for a coherent debate geared towards the development of relational networks for monitoring of exposures of various origins. Beyond the idea of collecting multi-exposures, Barabasi [9] has constructed a "diseasome" to illustrate, under a network, the environmental and social factors which might have a potential role in the origins of obesity; and Goh et al. [10] have explored a network of human diseases that implicates similar genetic mutations. In a similar way, Christakis et al. [11] have used the concept of relational networks that link individuals sharing social ties to monitor their weight over time. Assuming that the weight of an individual may be influenced by his or her surroundings, this approach has highlighted the possibility that social networking could be a factor in the spread of obesity. What all these analyses and various views have in common is the search of similar characteristics for understanding underlying mechanisms and identifying key factors in the onset and development of diseases.

Our main goal is to develop a framework allowing analyzing the RNV3P database in terms of an evolving and growing complex network. For this purpose, objectives of this paper are twofold: first, we outline the framework of the occupational exposome approach by providing definitions and showing how to construct exposomes, and second we show how to characterized and structured the topology of the constructed exposome by means of occupational exposure groups (sets of nodes or OHPs in a network sharing similar characteristics) which represent motifs carrying information on occupational exposures potentially related to the disease. To illustrate how this approach can be implemented, we considered the sample of Non-Hodgkin Lymphomas from the RNV3P database and, subsequently, we show the potential of this approach for monitoring and studying associations and relationships between diseases and occupational multi-exposures.

2. Material

Each diagnosed patient is recorded in the RNV3P database as an Occupational Health Problem (OHP) represented by a four codes item (Fig. 1) that includes a disease or pathology code (ICD-10 classification) plus a three-item exposure code holding information on the kind of industrial products, chemicals, psychological or organizational constraints (characterized by hazard codes) to which the patient was exposed during his or her occupational activity (characterized by occupation and activity sector codes). The hazards are coded using a hierarchical code owned by the national social security organism (Caisse Nationale d'Assurance Maladie, CNAM), the patient's professional activity is coded according to the international classification of occupational type (Classification Internationale Type des Professions, CITP-88), the French equivalent of the International Standard Classification of Occupations (ISCO), and the activity sector is coded according to the French occupational classification (Nomenclature des Activités Professionnelles, NAP-03). Two OHPs may differ either by their pathology and/or by at least one component of the exposure, i.e., hazards, occupation and activity sector. Likewise, several patients can present an identical OHP as well. From 2002 to 2007, the RNV3P database was full of 90,335 OHPs of which 75% were reported associated with one hazard, 17% with two hazards, 5% with three hazards, 2% with four hazards and less than 1% with five hazards.

3. Theory

3.1. Construction of the occupational exposome

To study the complexity of the relationships between diseases and composite occupational exposures and to seek similarities in associated exposures that potentially lead to identical diseases, we developed the concept of the occupational exposome. To process the RNV3P data in the form of a network of relationships, each OHP in the RNV3P database is represented by a node (vertex) $v = (p,e)^{T}$, which is a unique combination of a disease (pathology) "p" and its associated three dimensional composite occupational exposure $e = (h, o, s)^{T}$, characterized by a set of hazards "h", an occupation "o" and an activity sector "s" (Fig. 1). By convention, the hazard vector $h = (h_1, h_2, h_3, h_4, h_5)^{T}$ may comprise from 1 up to 5 distinct hazards that were present in the patient's occupational environment as characterized by "o" and "s", and are suspected or confirmed to be related or cause the disease. Each node (i.e., OHP) is weighted by the total number "w" of identical OHP copies in the database (see Supplementary information Fig. S3). In this way, the initial RNV3P database of OHPs is mapped into

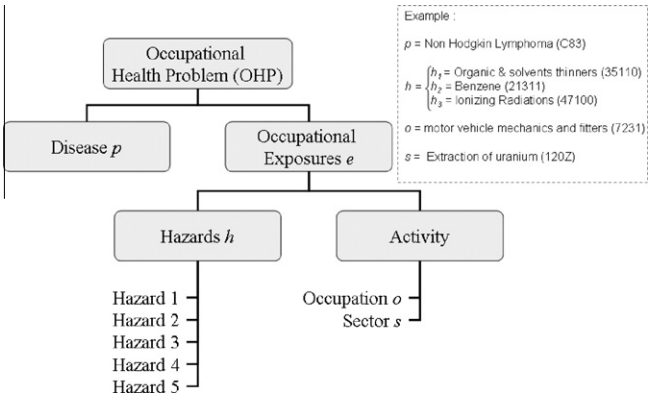

Fig. 1. Structure of an Occupational Health Problem (OHP). Inset: example of an OHP with item codes given between parentheses.