Computer-Aided Assessment of Drug-Induced Lung Disease Plausibility

Brigitte Séroussi[1], Jacques Bouaud[2], Hugette Lioté[1,3], and Charles Mayaud[3]

[1] Université Paris 6, UFR de Médecine, Paris, France; AP-HP, Hôpital Tenon, Département de Santé Publique, Paris, France
[2] AP-HP, DSI, STIM, Paris, France; INSERM, UMR_S 872, eq. 20, Paris, France
[3] Université Paris 6, UFR de Médecine, Paris, France; AP-HP, Hôpital Tenon, Service de Pneumologie, Paris, France
brigitte.seroussi@tnn.aphp.fr

Abstract. Drug-induced lung disease (DILD), often suspected in pneumology, is still a diagnostic challenge because of the ever increasing number of pneumotoxic drugs and the large diversity of observed clinical patterns. As a result, DILD can only be evoked as a plausible diagnosis after the exclusion of all other possible causes. PneumoDoc is a computer-based decision support that formalises the evaluation process of the drug-imputability of a lung disease. The knowledge base has been structured as a two-level decision tree. Patient-specific chronological and semiological criteria are first examined leading to the assessment of a qualitative intrinsic DILD plausibility score. Then literature-based data including the frequency of DILD with a given drug and the frequency of the observed clinical situation among the clinical patterns reported with the same drug are evaluated to compute a qualitative extrinsic DILD plausibility score. Based on a simple multimodal qualitative model, extrinsic and intrinsic scores are combined to yield an overall DILD plausibility score.

1 Introduction

Awareness of drug-induced lung disease (DILD) is increasing: a review published in 1972 identified only 19 drugs as having the potential to cause pulmonary disease. Now at least 400 agents are identified when querying the Pneumotox database [1] and the list continues to grow. Early diagnosis is important, because stopping the drug usually reverses toxicity, whereas unrecognized toxicity can be progressive and even fatal.

However, recognition of DILD remains a diagnostic challenge because there is no gold standard test, and clinical, radiologic, and histologic findings are non-specific. Thus, the diagnosis of drug-induced injury is currently only assessed as a plausible hypothesis. As a consequence, there is no reliable data on which DILD diagnostic probabilities could be estimated, forbidding numerical approaches to model the process. In this work, we propose a non-numerical empirical model of uncertainty to assess the plausibility of DILD as a qualitative drug-imputability score. This model accounts for the heuristic principles that are used in clinical routine to diagnose DILD [2]: qualitative plausibility scores are locally assessed from patient data, or *intrinsic* factors, and drug knowledge, or *extrinsic* factors. Qualitative scores are then combined through an intermediate quantitative step to yield the overall DILD plausibility.

R. Bellazzi, A. Abu-Hanna, and J. Hunter (Eds.): AIME 2007, LNAI 4594, pp. 359–363, 2007.

2 Reasoning Under Uncertainty

Plausibility is the state of being plausible, *i.e.* appearing worthy of belief, and is related to uncertainty. There has been considerable work to model uncertainty and belief, and how to combine uncertain information, or facts, to draw plausible inferences [3]. Uncertainty representations in AI fall into two basic categories: numerical (such as Bayes's, Dempster-Schafer's, and fuzzy theories) and non-numerical, or symbolic, approaches. Symbolic representations are mostly designed to handle the aspect of uncertainty derived from the incompleteness of information. Among these models, endorsement theory [4] relies on a heuristic approach of uncertainty. It was initially proposed as an alternative to the probabilistic handling of uncertainty: subjective degrees of belief are hardly quantified and generally do not behave as probabilities. The major advantage of this theory is that it makes *explicit* sources of uncertainty and the way they are combined so we may reason about them directly, instead of implicitly through some sort of numerical calculus. This qualitative reasoning about uncertainty is suited to model human expert-based knowledge, which can be context-dependent, when support data is missing. Some authors compared numerical theories and endorsement theory for a problem that combined data and heuristic knowledge [5].

3 Knowledge Model

The knowledge model (KM) is represented by a two-level tree-structured algorithm. The first level explores patient data to assess the intrinsic DILD plausibility. The second level explores bibliographical drug data to evaluate the potentiality of the suspected drug to induce pulmonary toxicity and assess the extrinsic DILD plausibility. This second phase has thus to be operated for each suspected drug.

3.1 Intrinsic Factors

Intrinsic factors include chronological and semiological criteria. The assessment of chronological criteria involved checking the suspected drug intake is before pulmonary manifestations occurred (otherwise the chronology is *incompatible*), searching for previous pulmonary episode with the same drug or hypersensitivity reaction following the suspected drug intake. The assessment of semiological criteria is only developed when the chronology is not *incompatible*. The first step consists in the exclusion of non drug-induced diagnoses for which pathognomonic characterizations exist. The 5 differential diagnoses modeled (pulmonary œdema, pulmonary embolism, infective pneumonia, and malignancy or systemic disease) are thus incrementally explored. As soon as one is proven, the DILD plausibility associated to semiology is *incompatible*. When not *incompatible*, the second step based on the evaluation of clinical, radiologic, and histologic (BAL) findings is performed. Opacities observed on X-rays are first characterized as localized or disseminated. When disseminated, the delay between the first manifestation of clinical and radiologic symptoms and the hospitalization date is assessed

making the difference between overacute, acute, and chronic pneumonia. BAL findings as well as other clinical parameters (fever, acute respiratory distress syndrome, etc.) are integrated and the DILD plausibility score from semiological criteria is attached to the different patterns.

3.2 Extrinsic Factors

As opposed to intrinsic factors assessed from patient-specific data, extrinsic factors represent drug-based information. This information is evaluated using the Pneumotox database (*www.pneumotox.com*). The website offers a comprehensive catalog of drugs known to be responsible of DILD and gives for each drug a rough estimate of adverse effects frequency (as reported in the literature) scored with stars: '*', isolated case reports (1 to 5) which await confirmation; '**', about 10 available cases; '***', in the range of 20 to 100 cases; '****', more than 100 reported cases. No star means suspicious drug but no data published yet. For any suspected drug, absolute frequencies of pulmonary toxicity given by Pneumotox have been weighted to integrate frequencies of the observed clinical situation among the reported clinical patterns. The assessment of extrinsic DILD plausibility is summarized in figure 1.

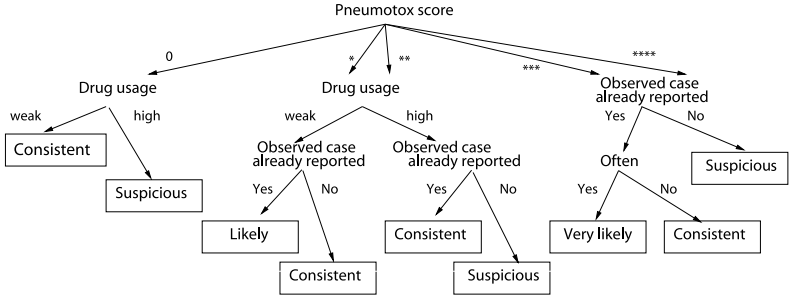

Fig. 1. Assessment of extrinsic DILD plausibility score

4 Plausibility Model

The plausibility representation of a statement related to drug imputability is based on an ordered set of qualitative values: *incompatible*, *suspicious*, *consistent*, *likely*, *very likely*. More specifically, *certain* is not considered in our application domain, since DILD certainty can never be established in the diagnostic process.

To manage the combination of plausible statements in a probability-like manner using the cross product, numerical values and intervals have been assigned to the qualitative plausibility values. In this context, *incompatible* means that the DILD hypothesis should be rejected whatever the plausibility of other statements: *incompatible* has thus to be the null element. *Consistent* has to be the neutral element: there is no specific argument to doubt or believe in DILD; when a DILD *consistent* statement is combined

Table 1. Numerical spaces assigned to symbolic plausibility values

Incompatible	*Suspicious*	*Consistent*	*Likely*	*Very likely*
0	$1/2$	1	2	> 2

with another statement, the resulting plausibility is that of this other statement. Table 1 reports such an assignment.

Let P_{chrono} denotes the chronology-related DILD plausibility, P_{semio} the semiology-related DILD plausibility, and $P_{intrinsic}$ the intrinsic DILD plausibility. Since $intrinsic = chrono \wedge semio$, and as chronology parameters and semiology data are considered independent, then $P_{intrinsic} = P_{chrono \wedge semio} = P_{chrono} \times P_{semio}$. As for extrinsic DILD plausibility denoted $P_{extrinsic}$, it is derived by considering literature citations and drug use as modeled in the KM (see figure 1). Thus, the overall DILD plausibility score $P_{overall}$ is obtained by the combination of intrinsic and extrinsic DILD plausibility scores: $P_{overall} = P_{intrinsic} \times P_{extrinsic}$.

However, whatever the strict value of the overall DILD plausibility score, the qualitative nature of the combination allows us to distinguish 5 different situations (Table 2) which lead to different interpretations and actions. There are 3 situations where intrinsic and extrinsic plausibilities are in the same range: (i) DILD is *suspicious* and the drug toxicity hypothesis will be rejected, (ii) DILD is *likely* or *very likely* and the DILD hypothesis will highly be considered with the immediate stop of the treatment, (iii) DILD is *consistent*, which is the most difficult case, there is neither evidence to support the hypothesis nor to reject it. In this case, any other administered drug that would better explain the clinical situation should be considered. By default, the suspected drug should be stopped to assess actual DILD.

In the 2 other situations, there is a mismatch between what is observed for the patient and what is currently known about the drug, this corresponds to "dissociation patterns". First, when extrinsic plausibility is *likely*, or *very likely*, while intrinsic plausibility is *suspicious*, extrinsic data should take precedence: although patient data should be carefully re-assessed, DILD should be considered and the treatment should be stopped. Second, when extrinsic plausibility is *suspicious* and intrinsic plausibility is *likely*, or *very likely*, if there is no other suspected drug, the actual clinical case could be a candidate for a new clinical pattern. The administered drug should be stopped for DILD assessment. In both situations, if DILD is confirmed, the clinical case is a new occurrence and should be ideally "published" to actualize the Pneumotox database.

Table 2. Overall DILD plausibility from the combination of intrinsic and extrinsic plausibilities

<table>
<tr><td rowspan="2">Intrinsic</td><td colspan="4">Extrinsic</td></tr>
<tr><td>Suspicious</td><td>Consistent</td><td>Likely</td><td>Very likely</td></tr>
<tr><td>Suspicious</td><td>Suspicious</td><td>Suspicious</td><td colspan="2">Dissociation</td></tr>
<tr><td>Consistent</td><td>Suspicious</td><td>Consistent</td><td>Likely</td><td>Very likely</td></tr>
<tr><td>Likely</td><td rowspan="2">Dissociation</td><td>Likely</td><td>Very likely</td><td>Very likely</td></tr>
<tr><td>Very likely</td><td>Very likely</td><td>Very likely</td><td>Very likely</td></tr>
</table>

5 Conclusion

PneumoDoc is a computer-based decision support system designed to help physicians assessing the overall plausibility of DILD. Developed in the documentary paradigm of decision making initially proposed with OncoDoc [6], the system relies on a knowledge base (KB) which is interactively browsed by the user physician. Structured as a two-level decision tree, the KB implements the DILD diagnosis strategy proposed by the KM described. Based on the heuristic principles used in clinical routine, the system makes the most of patient data (intrinsic factors) and literature-based drug information (extrinsic factors). The proposed, specific, plausibility model is based on qualitative ordered values which are combined according to heuristic, domain-dependent, knowledge. Akin in spirit to the endorsement framework, the advantage is the explicit specification of how uncertainty is propagated, which allows for the contextual interpretation of the system's result in the case of an actual patient.

As DILD diagnosis relies on the successive exclusion of all other possible causes, beyond structuring the reasoning process, the sequence and the choice of the investigations proposed by PneumoDoc to eliminate the 5 main differential diagnoses stand for the appropriate etiologic search strategy. In addition, PneumoDoc may help the detection of new toxicity cases (new pneumotoxic drug or new clinical pattern of a known pneumotoxic drug) with the identification of dissociation patterns.

The system has been tested on 20 actual medical records of DILD and lead to 100% of *very likely* DILD plausibility score. A retrospective evaluation is currently under process on 50 randomly selected pneumological records (including known DILD and non DILD). A multicentric survey is planned to be carried out to measure the impact of PneumoDoc on medical practices evaluated in terms of medico-economical parameters (length of hospitalization, number and type of laboratory tests used in the etiologic search, etc.).

References

1. Camus, P., Fanton, A., Bonniaud, P., Camus, C., Foucher, P.: Interstitial lung disease induced by drugs and radiation. Respiration 71(4), 301–326 (2004)
2. Mayaud, C., Fartoukh, M., Parrot, A., Cadranel, J., Milleron, B., Akoun, G.: Les pneumopathies infiltrantes diffuses d'origine médicamenteuse : un probléme avant tout diagnostique. Rev Pneumol Clin 61(3), 179–185 (2005)
3. Bonissone, P.: Reasoning, plausible. In: Shapiro, S. (ed.) Enc Artif Intell, pp. 854–863. John Wiley and Sons, Inc., Chichester (1987)
4. Cohen, P.R.: Heuristic reasoning about uncertainty: an artificial intelligence approach. Pitman, Boston (1985)
5. Comber, A.J., Law, A.N.R., Lishman, J.R.: A comparison of Bayes', Dempster-Shafer and endorsement theories for managing knowledge uncertainty in the context of land cover monitoring. Computers, Environment and Urban Systems 28, 311–327 (2004)
6. Séroussi, B., Bouaud, J.: Using oncodoc as a computer-based eligibility screening system to improve accrual onto breast cancer clinical trials. Artif Intell Med 29(1–2), 153–167 (2003)