Abstract
Computerized Clinical Practice Guidelines (CPGs) improve quality of care by assisting physicians in their decision making. A number of problems emerges since patients with close characteristics are given contradictory recommendations. In this article, we propose to use fuzzy logic to modelize uncertainty due to the use of thresholds in CPGs. A fuzzy classification procedure has been developed that provides for each message of the CPG, a strength of recommendation that rates the appropriateness of the recommendation for the patient under consideration. This work is done in the context of a CPG for the diagnosis and the management of hypertension, published in 1997 by the French agency ANAES. A population of 82 patients with mild to moderate hypertension was selected and the results of the classification system were compared to whose given by a classical decision tree. Observed agreement is 86.6% and the variability of recommendations for patients with close characteristics is reduced.

INTRODUCTION
Clinical Practice Guidelines (CPGs) aim at improving quality of care, reducing variation of medical practice among physicians and controlling costs. Guidelines can be implemented in Decision Support Systems to suggest treatment or diagnostic procedures to physicians through messages [1]. Moreover, it has been proved that computerized CPGs allow better accessibility to information and that output messages are more respected [2]. A message contains the following information: 1) the content of the message itself, linked to clinical data; 2) the grade of recommendation. defined according to the level of evidence thanks to the methodology of Evidence-based Medicine [3].

A CPG for the diagnosis and the management of hypertension has been published in 1997 by the French Agency ANAES (Agence Nationale d'Accréditation et d'Évaluation en Santé) [4]. This CPG has been computerized as a decision tree within the context of the EsPeR project (Individualized Estimation of Risks). This project provides physicians with appropriate risk estimates and corresponding CPGs, individualized according to patients’ characteristics [5]. The CPG contains a set of messages such as "For a 60 to 80 years old patient usually having a Systolic Blood Pressure above 160 mmHg, the benefit of treatment is proved (Grade A)". A number of problems have emerged in the application of the guideline. First, the level of evidence, associated to the message, expresses the overall validity of the related literature but is not tailored to an individual patient. Second, messages contain precise decision thresholds [6]. For instance, in the previous example, the message includes three thresholds: 1) age 60, 2) age 80, 3) SBP 160. Such precise thresholds lead to unconsistency concerning the appropriateness of the message for a given patient. Consequence of this uncertainty is shown in figure 1 where three close patients are given contradictory recommendations using the decision tree shown figure 1.

Figure 1: Example of use of decision tree from 1997 Hypertension guideline for three patients with close characteristics.

In this article, we propose to use fuzzy logic to modelize uncertainty due to the use of thresholds in messages.

Such approach allows to associate to each couple (patient, message) a strength of recommendation that expresses to which extend the conclusive part of the message (the recommendation) should be followed for this particular patient. The results show that the outcome variability for patients with close characteristics is then reduced.
BACKGROUND

Uncertainty in Guidelines
Uncertainty in CPGs occurs at three levels. The first one is the uncertainty due to the use of precise thresholds. Indeed, such thresholds have to be understood as concepts rather than true values. For example, for a condition such as: "age over 80", the value 80 has a higher degree of confidence for a very old person than for someone just over 80. In the latter case, the physician interpretation for the condition will depend on the physiological state of the patient. The second level is the uncertainty inherent to the messages conclusions. Indeed, such conclusions are natural language expressions that are naturally imprecise like for instance "treatment is recommended" (figure 1). In the studied CPG published by ANAES, there are two basic conclusions: "begin drug treatment", "do not treat". The third level of uncertainty concerns the levels of evidence and is supported by the evidence-based medicine approach.

The present work deals with uncertainty due to precise thresholds.

Guidelines and fuzzy logic
L. Ohno-Machado et al. [7] reported recently how much it is necessary to take into account uncertainty in computerized CPGs. They quoted three sources of fuzziness, corresponding to a lack of data: 1) a missing biologic exam result, 2) an unknown family history and 3) an unavailable result. They suggested the fuzzy decision tree approach to perform the classification procedure that assign a patient to a recommendation.

Furthermore, Liu et al. [8] developed a fuzzy model from a CPG published by the American Academy of Pediatrics on the management of first simple febrile seizures. The guideline 1) describes the children who are eligible for its recommendation, 2) states that routine diagnostic use is not generally indicated, and 3) lists factors involved in the decision to perform lumbar puncture (LP) as part of the diagnostic process. In their context, the criteria for deciding whether to perform LP, and the strength of recommendation are not clear-cut. They describe the executability of the recommendation of an LP in terms of four classes: an LP can be recommended, strongly considered, considered, or not routinely warranted. The appropriateness of LP is modelized by a fuzzy set for each three decision variables. This work assumes that any of these three variables alone would be sufficient to trigger the LP decision in a child and assumes also a good knowledge of how a variable and a recommendation are linked. In the context of the CPG for the diagnosis and the management of hypertension, a variable alone cannot trigger the recommendation. In the present work, the strength of recommendation (appropriateness) is a continuous variable defined for each message.

MATERIAL

Decision variables
The decision variables involved to prescribe antihypertensive medication are the five following ones [5]:
1. Systolic Blood Pressure (SBP). The universe of discourse for this variable is the interval [80, 300] in mmHg unit. The value of SBP is confirmed by several measures.
2. Diastolic Blood Pressure (DBP). The universe of discourse for this variable is the interval [30, 160] in mmHg unit. The value of DBP is confirmed by several measures.
3. Age. Values are from 20 to 80 years old.
4. History of Cardiovascular disease (CVH). This binary variable is defined by the previous occurrence of a myocardial infarction or a stroke.
5. Global Risk (GR). The global cardiovascular risk of a patient is computed from the Framingham equations and is rated from 0 to 100 % [9].

In this study, we collected the data from a database designed for the medical management and follow-up of hypertensive patients. We use the population of a study designed to evaluate the preventive effects of anti-hypertensive drugs [10]. In this study, data were collected in 1993 (82 patients) and most patients had mild to moderate hypertension.

Table 1: Descriptive clinical characteristics of the database (82 patients)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average</th>
<th>Standard Deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBP</td>
<td>162.75</td>
<td>25.35</td>
<td>112</td>
<td>260</td>
</tr>
<tr>
<td>DBP</td>
<td>99.61</td>
<td>12.51</td>
<td>68</td>
<td>134</td>
</tr>
<tr>
<td>Age</td>
<td>48.37</td>
<td>12.24</td>
<td>30</td>
<td>80</td>
</tr>
<tr>
<td>CVH</td>
<td>4%</td>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>GR</td>
<td>11.13</td>
<td>1</td>
<td>1</td>
<td>46</td>
</tr>
</tbody>
</table>

Clinical Practice Guideline
The CPG published by the ANAES for the diagnosis and the management of essential hypertension in adults from 20 to 80 years old is used in this work [4].

The two basic recommendations ("begin drug treatment", "do not treat") are distributed inside seven messages that are listed in Table 2. They comprise a condition, a recommendation and a grade.

Table 2: the seven messages for the studied CPG.
METHODS

In this section, we first present the modeling of the conditions in the context of the fuzzy set theory and then describe how the strength of recommendation is computed for each message using a fuzzy classification system.

For a given message, we introduce the following notations:

Let $X_i$ for $i = \{1, 2, 3, 4, 5\}$ be the decision variables involved to prescribe anti-hypertensive medication. For each variable, a finite universe of discourse $[A_i, B_i]$ is defined as well as a given number $m_i$ of thresholds specific to this variable. The cut-out of the universe of discourse with $m_i$ thresholds produces $m_i+1$ sub-sets that are noted $D_j^i$ for $j = 0, ..., m_i$ (see an example figure 4).

Fuzzy representation of thresholds

Briefly, a fuzzy set $S$, defined on a referential $X$, is a set whose membership function $\mu_S$ takes its value in the interval $[0, 1]$. $\forall x \in X$, $\mu_S(x)$ expresses to which extent the value $x$ belongs to $S$. The value 0 corresponds to the absolute non-membership and the value 1 corresponds to the absolute membership. In the context of quantifiable variables, the referential $X$ is the set of reals $\mathbb{R}$ and we adopt the classical trapezoidal representation of a fuzzy set (see figure 2) [11].

As it was pointed out earlier, thresholds values come from literature and are naturally uncertain. They are represented by fuzzy sets with a unique modal value (fuzzy numbers). The figure 2a shows the definition of the membership function for a fuzzy threshold $S$ and its graphical representation. The values $a$ and $b$ that appear in the membership function may be provided directly by an expert of the domain or thanks to statistical methods. In the present work, the values $a$ and $b$ are determined arbitrarily.

Let "<S" and ">S" be the sets of values "smaller than S" and "greater than S". They are defined by their membership functions given in figure 2b.

The fuzzy representation of the thresholds leads to the fuzzy representation of the subsets $D_j^i$. Indeed, for a given threshold $S_j$, we have

$$D_j^i = \{ x \in [A_i, B_i] / S_j \leq x < S_j \}$$

Consequently, the membership function associated to the subset $D_j^i$ is given by :

$$\forall x \in [A_i, B_i], \mu_{D_j^i}(x) = \min(\max(\mu_{<S_j}(x), \mu_{S_j}(x)), \mu_{>S_j}(x)).$$

Computing the strength of recommendation

The fuzzy classification consists in computing for each message a membership degree that express to which extent the condition in the message is satisfied by the patient characteristics. The conditions involve several variables and two kinds of operators are necessary to combine these variables.

A first operator $op$ is necessary to combine, for a single variable, the different $D_j^i$ appearing in the condition. For example, in the message n°5 (table 2), for $X_i = " AGE "$ the condition is "AGE < 60" or "AGE > 80". In this example, $op = "or"$.

$$\forall x \in X_i, \mu_S(x) = 1 - \sup_{y \neq x} \mu_S(y)$$

The "Minimum" expresses the simultaneous satisfaction of elementary conditions and corresponds to the "and" operator:

$$\forall (x, y) \in [0, 1]^2, OM(x, y) = \min(x, y)$$

The "Maximum" expresses the fact that the satisfaction of only one elementary condition is sufficient and corresponds to the "or" operator:

$$\forall (x, y) \in [0, 1]^2, OM(x, y) = \max(x, y)$$

Using the fuzzy representation of $D_j^i$ and the operators, a compatibility degree is computed between the patient data and a condition in a message. This degree is the strength of recommendation for the message under consideration. The figure 3 gives a graphical representation of the fuzzy classification procedure. The recommendation with the greatest degree is then proposed to the physician.
RESULTS

Among the five variables, one variable is binary (CVH), four are continuous for which nine thresholds are defined. The figure 4 shows the fuzzy sets associated to $D_{DBP}$ for the DBP variable.

A fuzzy classifier has been developed in Java. The database of 82 patients described earlier was used to test it. For each patient, the recommendation with the strongest degree was compared to the message delivered by the decision tree implemented within the context of the EsPeR project. Results are displayed in figure 5.

Observed agreement is 86.6 % that is 13.4 % of the patients classified differently according to the method, the fuzzy classifier or EsPeR decision tree (figure 5a). The results show that the fuzzy system tends to overestimate the seriousness of the patients’ state (possibly because the database contains mainly patients with mild hypertension). We observe also that the message $n=5$ is never selected. The figure 5b shows the results for the three patients of figure 1 (these three pedagogic cases are not part of the database). The variability between the conclusive messages is reduced but the differences still appear in the values of the strength of recommendation.

DISCUSSION AND CONCLUSION

The modeling of thresholds by fuzzy numbers allows a flexible cut-out of the universe of discourse of a decision variable. From these fuzzy thresholds, we developed a fuzzy classification procedure that provides for each recommendation message of the CPG, strength of recommendation that rates the appropriateness of the recommendation for the patient under consideration. It is important to note that this strength of recommendation does not come from a lack of data (missing data, imprecision of data) as described in [7] but from the uncertainty of the message itself.

However, the strength of recommendation is complementary to the level of evidence. The level of evidence refers to the validity of a message and the strength of recommendation refers to the appropriateness of this message for patients close to thresholds. These two complementary quantities appear somehow in many ‘machine learning’ approaches where a distinction is made between the errors made in the construction of a model (internal validity) and the error made while generalizing the model to an individual (external validity).

The main limits of this approach are the ones generally pointed out while using fuzzy sets. The definition of the thresholds membership functions has to come from experts. In our work, we prototyped these functions since we don’t think that a small change in the $a$ and $b$ values (figure 2) will have any great impact on the final result. However, obtaining these values from statistical approaches will allow to verify this hypothesis [12]. The definition of the two kinds of operators is also preliminary in our work. Our application of the conjunction/disjunction operators assumes that decision variables are of equal importance in messages’ conditions. In order to take into account variables of different importance, weighted operators...
may be used. These weights can be constant and relative to the variable like in [12] or they can vary according to the characteristics of the patient under consideration [13].

Besides this, the output of the system can be discussed. Indeed, it could be interesting to provide the physician with more than one recommendation, let’s say, for instance, the two recommendations with the best values. It will lead to a more precise comparison with non-fuzzy approaches. Another point concerns the couple (strength of recommendation, level of advice). Is it preferable to provide a strong recommendation with grade D rather than a less strong recommendation but with grade A? The Pareto approach could be used to order the output and answer this question [13].

Within the framework of this project, we compared results obtained with the fuzzy classification and with the original decision tree. The concordance between the two approaches is perfect for patients whose characteristics are not close to thresholds. For the others, we observed that the fuzzy classification tends to make the situation worse than the decision tree due to the fact that the database contains essentially patients with mild hypertension. A more representative database should be used to get a more balanced difference between the two approaches. It will also be interesting to define a gold standard for the comparison (e.g. a panel of experts). However, the fuzzy classification seems already interesting for patients with close characteristics.

The procedure has been developed in the context of a CPG for the diagnosis and the management of hypertension, published in 1997 by the French agency ANAES. The methodology is not specific to this singular CPG and could be generalized to other CPGs.

Finally, a perspective of this work should be to investigate similar approaches to take into account the two other levels of uncertainty in CPGs.

References