QUESTIONNAIRE SCREENING OF SLEEP APNEA CASES USING FUZZY KNOWLEDGE REPRESENTATION AND INTELLIGENT AGGREGATION TECHNIQUES

David Nettleton
Consulting Group – Business Intelligence
IBM Global Services, Spain
dnettlet@es.ibm.com

Lourdes Hernández
Respiratory Disease Institute,
Hospital Clinic
University of Barcelona, Spain
fburgos@medicina.ub.es

Summary

In this article, joint medical and data analysis expertise is brought to bear using fuzzy knowledge representation and 'intelligent' aggregation techniques to solve a difficult medical diagnosis problem, that of sleep apnea syndrome screening.
Key Words: fuzzy representation, sleep apnea diagnosis, questionnaire, aggregation

1. INTRODUCTION

Screening of Apnea cases is a difficult diagnosis problem, at present not satisfactorily resolved by standard statistical modelling techniques. We propose that part of the problem is due to the inherent fuzzy nature of a significant part of the data: questionnaire responses. In this article a fuzzy representation is proposed for the questionnaire responses, which permits 'hedging' on the part of the patient. Also, three contrasting data aggregation techniques 'fuse' the variables into a diagnosis for each case.

The article is structured as follows: in section 2, a clinical description of the sleep apnea syndrome is given; in section 3, the sleep studies questionnaire is described; in section 4, examples of processing with test cases for aggregation based diagnosis are detailed; in section 5, the fuzzy representation for the questionnaire responses is presented; in section 6, aggregation techniques are detailed; section 7 summarises and gives some conclusions on the present work and its future direction.

We address the *kind of knowledge used* in sections 2 and 3, as the domain knowledge is embodied in the questionnaire. The *need to exploit the available prior knowledge* inherent in the problem is covered in sections 3 and 4, this knowledge being represented in the form of reliability and relevance weights for each variable (question) and data value. The *way in which the available knowledge is represented* is detailed in sections 3, 4, 5 and 6. Finally, *the plan of how to use the derived knowledge* is covered in sections 4 and 7. We can consider as *derived knowledge* the membership grades of the patient responses, and the final aggregated value used for diagnosis of sleep apnea syndrome.

Figure 1a shows the traditional approach which is used in the literature for the processing of Apnea case questionnaire data. Questionnaire responses have a categorical or a numerical representation. Groups of variables are aggregated by indexes such as Epworth. This gives a crisp diagnosis value of 1 or 0. In Figure 1b we see the approach of this article. Questionnaire responses can have a fuzzy or categorical representation and three aggregation methods are tested. The final patient diagnosis is a grade of membership value.

Categorical and numeric representation

↓

Indexes to aggregate clinical responses to questions and data (Epworth, combinations of variables, …)

↓

Patient diagnosis [0,1]

Figure 1a: Standard (crisp) questionnaire screening

Fuzzy linguistic variables and numerical representation

↓

Aggregation functions (owa, wowa, principal components)

↓

Patient diagnosis(μ)

Figure 1b: Summary of proposed fuzzy method

2. THE SLEEP APNEA SYNDROME

The Obstructive Sleep Apnea Syndrome (OSAS) is a set of secondary clinical manifestations relating to the ceasing (apnea) or reduction (hypopnea) of air flow during sleep, caused by a partial or total collapse of the upper air way at the faringe level. The severity of the OSAS is defined by the *apnea hypopnea index* (AHI) or the number of apneas plus the number of hypopneas per hour during sleep. Generally an AHI ∃10-15 is considered pathological.

CLINICAL PRESENTATION: There are diverse symptoms associated with OSAS. They often become introduced insidiously during a certain period of time and are often overlooked in clinics and even by the patients themselves, due to their lack of specificity. The snore is one of the principal symptoms. The long snoring history which refer to patients with OSAS reflects the increase of resistance of the upper air tract during sleep. The presence of respiratory pauses witnessed by the room partner is another important data referenced in the literature, and tends to be a good symptom predictor.

PREVALENCE: The prevalence of OSAS oscillates between 1-9% according to studies. This difference in the percentages obtained reflects the diversity of methods and criteria used to diagnose OSAS and the possible differences in the populations that have been studied. The study of reference is that realised in the population of Wisconsin by Young et al[13], where the prevalence obtained reached 2% for females and 4% for males, showing minimum symptoms. When we extrapolate these results to the general population, 9% of women and 24% of men would present sleep related respiratory alterations. This elevated prevalence in adults is considered to be a significant problem for public health.

MORBIDITY AND MORTALITY: Daytime hypersomnolence has been related to a reduction of physical and mental effectiveness, in the daily activity of the individual, including the work environment, and the ability to drive automobiles (drive worse and have greater risk of suffering traffic accidents). As well as daytime hypersomnolence, a certain relation has been identified between OSAS and systemic arterial hypertension. The patient with OSAS tends to present an elevated sympatic activity, which can cause an increase in the daytime blood pressure.

DIAGNOSIS: The predictive value of the clinical data in OSAS diagnosis is low. Hoffstein [5] published results that indicated that clinical data explains 36% of the variability of the AHI (apnea hypopnea) and Katz [6] reported a figure of 39%, other authors report lower figures (Table1). The subjective clinical evaluation of the interviewer has also been evaluated and tends to have a low sensibility and specificity, in the order of 55%-65% respectively, for correctly classifying the sick. On the other hand, The predictive models for AHI based in clinical data have a higher sensibility of up to 90%. Their specificity, in the best of cases, does not reach 70% .

The reference method for OSAS diagnosis is the polysomnogram. It consists of the simultaneous recording of a number of sleep parameters, which allow us to identify its different phases and the correlation of these with cardiorespiratory events such as apneas, desaturation of oxyhaemoglobine and changes in cardiac rhythm. For sleep measurement, including body position changes, respiratory effort and efficiency in ventilation, there exist multiple methods and each clinic tends to use its own variables which are obtained with the resources available in each centre.

At present, it is not appropriate to define rigid diagnostic criteria in this rapidly developing area. Neither is it possible to identify the ideal equipment for sleep studies.

3. DESCRIPTION OF THE SLEEP STUDIES QUESTIONNAIRE

The purpose of the questionnaire is to provide an information profile of the patient which allows a pre-diagnosis of his/her condition. This acts as a ‘screening’ which avoids patients entering into the sleep centre for expensive and time consuming testing, when they have a low probability of suffering from Apnea Syndrome, or have some other pathology.

The data which is captured by the questionnaire represents all the information we have about a patient in order to diagnosis Sleep Apnea.

Table 1. Multiple linear regression models

Study	n	Diagnostic criterion	Predictive variables	r^2
Stradling (1991)	1001	ID4%>5	Neck circumference, alcohol consumption, age, obesity	0.14
Davies (1992)	150	ID4%	Sleep when inactive Neck circumference	0.13 0.35
Hoffstein (1993)	594	AHI>10	BMI,age, sex, snoring, exploration of ORL	0.36
Flemons (1994)	180	AHI>10	Neck circumference, HTA, snoring, observed apneas	0.34
Deegan (1994)	250	AHI👓15	BMI, age, alcohol consumption	0.19

ID4%: index de desaturation with fall of 4%. AHI: apnea-hypopnea index. r^2: regression coefficient. BMI: body mass index. ORL: otorrinolaringologic exploration. HTA: arterial hypertension.

The *prior knowledge* which exists in the questionnaire is implicit on several levels: on a primary level, knowing which questions to ask; on a secondary level, knowing how to phrase and represent the questions. *A Priori* knowledge is also embodied in the relevance and reliability weights later described in sections 4 and 6. Prior medical knowledge is fundamental to diagnosis and thus the representation and aggregation methods are designed to embody and exploit it.

The questionnaire consists of two main sections: the first records clinical data, with 15 key clinical variables: age, sex, presence of a partner, profession, work hours, education level, weight, height, neck circumference, BMI (body mass index), blood pressure, alcohol intake, cigarette intake, auto-evaluation of most important symptom, other illnesses; the second section consists of 41 questions to which the patient responds on a five point scale {never, rarely, sometimes, frequently, always}. The questions are divided in 3 subsections: 15 general sleep questions, 16 respiratory related questions and 9 somnolence related questions. Based on this information, the doctor then gives a clinical evaluation: healthy; simple snorer; doubtful; typical apnea; other illness. We interpret this as: typical apnea; no apnea, with the corresponding grade of membership.

Example of a general sleep question:

G11. Do you know or have you been told that you move your legs a lot when you are sleeping?

never	rarely	sometimes	frequently	always

Question G11 relies on the presence of a partner as witness.

Example of a respiratory related question:

R7. Have you been told that you “stop breathing” while you are sleeping?

never	rarely	sometimes	frequently	always

Question R7 also relies on the presence of a partner as witness.

Example of a somnolence related question:

S10. Do you fall asleep in your workplace while you realise your normal work activies?

never   rarely   sometimes   frequently   always

Question S10 may not be answered truthfully by the patient due to possible employment repercussions.

The crisp response would be stored as a single value corresponding to the one and only one linguistic label chosen as the response for that question. For example the response to S10 (above) could be stored as: {1}. This indicates that only the linguistic label 'rarely' is considered as the response. The linguistic labels are considered as being in ascending order, from left to right. Thus we can meaningfully assign ordered ascending integer values thus: {0,1,2,3,4}

We have chosen a subset of variables from the questionnaire which in the literature (see Table 1) have been identified as the most discriminatory variables with respect to apnea diagnosis. These are: age, sex, weight, body mass index, neck circumference, alcohol intake, blood pressure, snoring and daytime sleepiness. The first 7 variables are clinical data and are crisp. For snoring information, we have used the responses to 4 respiratory related questions: R1, R2, R11 and R13 (see Table 2).

Table 2. Discriminant variables: example minimum set with weighting factors for aggregation

variable	description	relia-bility*	rele-vance*
age	age in years	E	0.50
sex	sex 1 or 2	E	0.70
weight	weight in Kg	M	0.70
IMC	body mass index in Kg/m^2	M	0.70
Neck circum-ference	Neck circumference in cm.	E	1.00
alcohol	Alcohol intake	M	0.50
HTA	Arterial hypertension mmHg	E	0.70
R_1	Do you snore when sleeping or have you been told that you do?	H	0.90
R_2	Does your snoring wake your partner or can it be heard from another room?	E	0.90
R_{11}	Do you have head-ache when you wake up in the morning?	E	0.80
R_{13}	Do you feel as if you haven't rested when you get up in the mornings?	H	0.70
S_3	Do you fall asleep when at the cinema, theatre, or other spectacle?	M	0.70
S_4	Do you sleep in meetings or in public places?	M	0.80
S_5	Do you fall asleep while driving on the motorway?	L	0.85
S_6	Do you fall asleep against your will during the daytime?	M	0.85

*the values of these columns are then converted proportionately to normalised values so that $\Sigma\omega = 1$ and $\Sigma\rho = 1$, as in Tables 3 and 4.

For daytime sleepiness information, we have used the responses to 4 somnolence related questions: S3, S4, S5, S6 (see Table 2). They were chosen as the key discriminatory questions with the highest statistical correlations with the output flag (apnea, yes or no).We have test data to demonstrate the techniques used in a simplified manner, as summarised in Tables 3 and 4, and in section 4 which follows.

4. EXAMPLES OF USE OF THE REPRESENTATION AND AGGREGATION METHODS

With reference to Tables 3, 4 and 5, we follow the processing with example test cases and compare the resulting diagnostic outcomes. All of the weights (ω,ρ) and the membership functions are defined by the medical expert. The data analysis expert then checks for any inconsistencies from a statistical point of view.

4.1 APPLYING THE AGGREGATION TECHNIQUES TO THE DATA.

We have aggregated a preselected subset of 15 variables with high discriminant value, which include 7 clinical data variables (age, neck circumference, etc. ..) and 8 question responses (from a total of 41). These variables are listed in Table 2. We have used expert medical knowledge, results from the literature (Table 1) and statistical analysis techniques to select the key variables. In aggregating we have considered all data as numeric. The question responses have orderable numeric values from 0 to 4, where 0 represents the linguistic label *never* and 4 represents the linguistic label *always*.

Table 3. ρ vector: each variable has a ρ weight which indicates its relevance. Σρ = 1

	Question Response Variable							
	R_1	**R_2**	**R_{11}**	**R_{13}**	**S_3**	**S_4**	**S_5**	**S_6**
ρ vector	0.14	0.14	0.12	0.11	0.11	0.12	0.13	0.13

Table 4. ω vector: each variable has an vector which weights the ordered data responses for that variable, in terms of their reliability. Σω = 1

	ω vector					
Variable	ω_1	ω_2	ω_3	ω_4	ω_5	Weighting Bias on:
R_1	0.10	0.10	0.20	0.30	0.30	High values
R_2	0.20	0.20	0.20	0.20	0.20	Even
R_{11}	0.20	0.20	0.20	0.20	0.20	Even
R_{13}	0.10	0.10	0.20	0.30	0.30	High values
S_3	0.10	0.25	0.30	0.25	0.10	Middle values
S_4	0.10	0.25	0.30	0.25	0.10	Middle values
S_5	0.30	0.30	0.20	0.10	0.10	Low values
S_6	0.10	0.25	0.30	0.25	0.10	Middle values

Aggregation of the questionnaire responses and clinical data

We assign one ρ weight to each question, which in this context indicates the significance of the question to the global outcome; and we assign a vector of ω weights which indicate the reliability of the response to each question. For example, if we though that a response of 'always' {4} is not reliable or to be expected for question S_5, we can give it a relatively small weight which, in the case of {4} being the response, will diminish its impact on the global outcome. We can also use the ω weight vector to eliminate outliers, for example, giving more weight to the values closest to the mean.

With reference to Table 4, the row for variable S_5 has the following vector ω: {0.3, 0.3, 0.2, 0.1, 0.1} which gives a low values bias. This pattern, once interpolated by the WOWA aggregation operator to give a continuous curve, will act on an ordered column of numeric data (in this case corresponding to the variable S_5) and will diminish responses of *frequently* and *always*, while strengthening responses such as *never* and *rarely*. With reference to Table 3, we see one ρ weight for each variable. Variable R_1 is given a relatively strong weight of 0.14 while variable S_3 has 0.11, which indicates that R_1 is more relevant to the outcome than S_3 in a ratio of approx 14:11.

The possible values for the ω vector are defined as follows: even bias (E) is assigned as {0.2, 0.2, 0.2, 0.2, 0.2}; low values bias (L) is assigned as{0.3,0.3,0.2,0.1,0.1}; high values bias (H) is assigned as{0.1,0.1,0.2,0.3,0.3}; high&low values bias (O) is assigned as{0.3,0.15,0.1,0.15,0.3}; middle values bias (M) is assigned as{0.1,0.25,0.3,0.25,0.1}.(See Table 2 reliability column, and Table 4).

Incorporating the membership grade information for each question

We have three weights: relevance of each variable ρ, reliability of the data values of each variable ω, and the membership grade of each data value of each variable μ. The WOWA operator allows two weight vectors as input, ρ and ω. To incorporate μ, we propose that u acts as a secondary weight on ω, thus incorporating additional information conditioned by the data capture, that is, the subjective choice of the patient in answering the question.

For example, given a ω vector for questionnaire variable S_5 of {0.3, 0.3, 0.2, 0.1, 0.1}, the patient responds to question S_5 on a continuous scale giving the following membership grades to each linguistic label: {1:0.0, 2:0.3, 3:0.7, 4:0.0, 5:0.0}. Thus the patient must have indicated his/her response between 'rarely' and 'sometimes', and closest to ‘sometimes’. We then alter the values of the ω vector thus: {0.3×0.0, 0.3×0.3, 0.2×0.7, 0.1×0.0, 0.1×0.0}. This incorporates the patients own confidence in his/her response, modifying ω using only the nonzero membership grades. A μ of 1.0 would indicate no modification to the ω weight. In this case, the ω vector, weighted by μ, will be {0.0, 0.09, 0.14, 0.0, 0.0}. To maintain the requirement that $\Sigma\omega = 1$, we then recalculate, giving {0.0, 0.39, 0.61, 0.0, 0.0}. Thus responses with a high ω reliability but a low μ reliability, can have their outcomes significantly altered.

4.2 INTERPRETING THE AGGREGATION RESULTS

In interpreting the aggregation results for all aggregation techniques, we need to define a threshold which indicates where 'do not admit' ends and 'admit' starts. We establish this by running known cases through and noting the values generated as output. We need a spectrum of cases, from a strongly positive case, to a strongly negative case, and a spectrum of intermediate cases ordered by degree of evidence of the apnea syndrome. This is measured clinically in terms of < 10 apneas / hour and >= 10 apneas / hour, so it is possible to assign a numeric quotient to the grade of incidence of apnea.

In this problem, we can consider the *derived knowledge* as the responses the patient gives, which is then interpreted by the membership functions and used as input to the aggregation functions. Also we can consider as *derived knowledge* the aggregated value produced for each patient. This will be used for Apnea screening, and provides another index to be considered along with the conventional (Epworth, …) indexes which the doctor uses in screening the patients for admission to the sleep centre.

Table 5. Input responses for 8 questions (see Table 2) with corresponding outcomes from aggregation methods.

	Input								Outcomes		
	Normalised question responses (0.00=never, 0.25=rarely, 0.50=sometimes, 0.75=frequently, 1.00=always)										
	R_1	**R_2**	**R_{11}**	**R_{13}**	**S_3**	**S_4**	**S_5**	**S_6**	Wowa	Owa	Principal Components
Data vector for Patient P_1	0.75*	0.50	0.75	0.75	0.50	0.75	0.25	0.75	0.52	0.84	1.15291
Membership-Grade μ	0.70*	0.85	0.80	0.90	0.75	0.80	0.65	0.85	admit	admit	admit
Data vector for Patient P_2	0.75	0.25	0.50	0.25	0.50	0.50	0.00	0.25	0.49	0.84	1.15324
Membership-Grade μ	0.85	0.65	0.70	0.80	0.80	0.65	1.00	0.80	do not admit	admit	admit
Data vector for Patient P_3	0.75	0.75	0.75	0.75	0.75	0.75	0.00	0.50	0.56	0.89	1.15466
Membership-Grade μ	0.65	0.75	0.75	0.80	0.80	0.65	0.70	0.75	admit	admit	admit
Data vector for Patient P_4	0.75	0.25	0.25	0.25	0.25	0.75	0.00	0.25	0.46	0.84	1.15353
Membership-Grade μ	0.75	1.00	0.80	0.90	0.80	0.90	1.00	1.00	do not admit	admit	admit

*for clarity, only the response with the highest membership grade is shown.

In Table 5 we see the results of executing the method using four theoretical test cases and 3 aggregation methods. Rows 1 to 3 are positive cases (admit), case 2 being *borderline*: from which we derive the % success rate of correct diagnosis of patients who have apnea syndrome; and row 4 is a strongly negative case (do not admit): from which we derive the % success rate of correct diagnosis of patients who do not have apnea syndrome.

The case data is not only weighted by the ρ and ω vectors, but also by the membership grade associated with the linguistic label of each question response.

We see that WOWA agrees with OWA and principal components for cases 1 and 3, and does not agree for the borderline case (2) and the strongly negative case (4). Principal components and OWA give positive outcomes for all four cases, thus having a high precision for positive diagnosis and low precision for negative diagnosis (high sensibility and low specificity as commented in Section 2.4) which is a typical result for standard statistical techniques used in the literature [11].

5. REPRESENTING THE LINGUISTIC LABELS – QUESTIONNAIRE RESPONSES

In this section we consider how to represent linguistic labels as fuzzy sets, first as simple trapeziods and then as curves which give a smoother transition between one label and the next.

5.1 PARMENIDEAN PAIRS

In general, the basic representation for parmenidean pairs is based on the use of fuzzy partitions with a trapezoidal membership function. In [2] a method is presented which automatically constructs a system of 5 linguistic labels which represent the ordered values of a variable derived from what we call a *'parmenidean pair'*. This responds to the basic opposing values that the variable may assume. This method is very appropriate for variables which represent responses to questions like 'Do you snore while you sleep or have you been told that you do?' for which we can define the fuzzy values NEVER, RARELY, SOMETIMES, FREQUENTLY, ALWAYS derived from the basic opposing values of NEVER, ALWAYS.

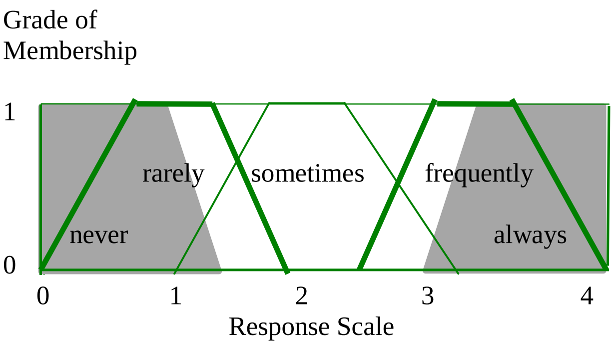


Figure 2. Representation of Ordinal Variables

Figure (2) shows a simple fuzzy representation for a typical questionnaire 'response' using a FLV (Fuzzy Linguistic Variable). From a semantic point of view, a FLV can be identified by 3 parameters: its relative *position* with respect to the other ones, its degree of *imprecision*, and its degree of *uncertainty*, these last can be merged into a single concept of *softness*, as opposed to *crispness*.

5.2 LINEAR TO NON-LINEAR MEMBERSHIP FUNCTIONS

Trapezoids formed by straight lines are really approximations of real membership functions. Thus we can get closer to having a natural representation (best fit) for the linguistic labels by generating a curve in place of a straight line for the ascending and descending gradients. To achieve this, we can use an interpolation technique such as that described in [1] to construct the curve. In some cases we may wish to strengthen a transition with hedges like "very" or "extremely" or weaken it with, for example, "slightly". We can perform strengthening by, for example, a sigmoid-like function.

We can use Zadeh's S-Function:

$$S(x;\alpha,\beta,\gamma) = \begin{cases} 0 & x \le \alpha \\ \left(\dfrac{x-\alpha}{\beta-\alpha}\right)^2 & \alpha < x \le \beta \\ \left(\dfrac{x-\beta}{\gamma-\beta}\right)^2 & \beta < x \le \gamma \\ 1 & \gamma < x \end{cases}$$

Now

$$f(x) = \begin{cases} \dfrac{1}{2} + \dfrac{\sqrt[3]{(x-1/2)}}{2\times\sqrt[2]{(1/2)}} & x > 1/2 \\[2ex] \dfrac{1}{2} - \dfrac{\sqrt[3]{(1/2-x)}}{2\times\sqrt[2]{(1/2)}} & x \le 1/2 \end{cases}$$

The use of $f(S(x;\alpha,\beta,\gamma))$ increases all the membership values above 0.5, and decreases all the others. This is the definition for "very"; for "extremely" we can replace in formula 2 the 3rd root by the nth root (for a suitable $n > 3$, n odd).

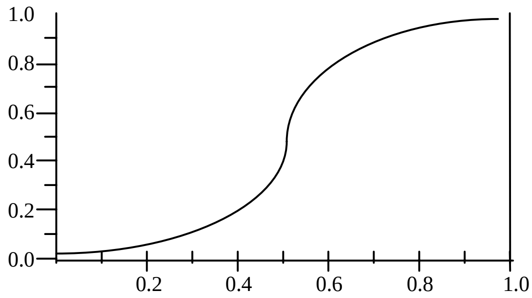


Figure 3. Zadeh's s-function can be used to customise membership transition

For linear and non-linear membership gradients, we assume a symmetrical relation between the descending membership value for the preceding fuzzy set and the ascending membership grade for the following fuzzy set (which sum to 1), as can be seen in Figure 3.

5.3 EXAMPLE OF FUZZY REPRESENTATION OF A QUESTIONNAIRE RESPONSE

For each question we design a membership function which can be overlaid on each scale to read off the grade of membership to each linguistic label.

S5. Do you fall asleep while driving on the motorway?

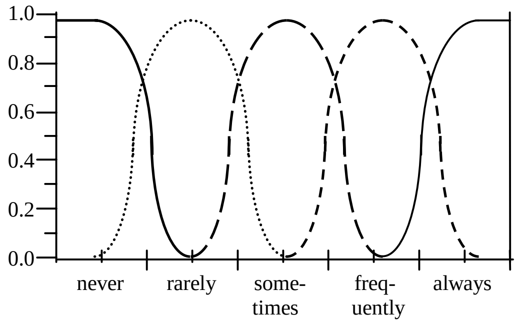
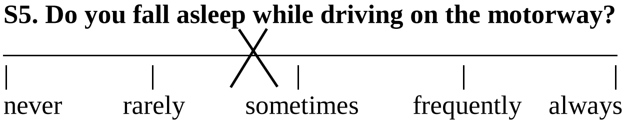


Figure 4. Example of representation for a critical question.

In Figure 4 we see that the curves are formed by Zadeh's s-function. We can manipulate this type of curve as detailed in section 5.2, in order to strengthen or weaken a linguistic label. The patient draws a cross on the continuous scale (eg. S5) to indicate his/her response to the question. In the questionnaire, this question would appear as:

S5. Do you fall asleep while driving on the motorway?



The fuzzy response would be stored as a quintuple, with a membership grade for each linguistic label. For example the response to S5 (above) could be stored as: {0:0.0, 1:0.3; 2:0.7; 3:0.0; 4:0.0}. This indicates that only linguistic labels 'rarely' and 'sometimes' have non-zero membership values, being 0.3 and 0.7 respectively. We can simply take the linguistic label with the highest membership grade, which in this case is 'sometimes'. Note that we can convert to categorical if we so desire, and that way we have both crisp and fuzzy data capture.

The membership grades of the responses can be read or by writing a computer programme which finds the corresponding point on the y-axis, for the point indicated by the response on the x-axis. Other wise, we can overlay a transparent sheet on each response line and read off the membership grade on the y-axis. Each sheet would have been drawn or created by a statistical package. We have chosen at present the latter method, which avoids dedicating time to programming and enables us to tailor one sheet of membership functions for each question.

6. AGGREGATION TECHNIQUES

In order to use the questionnaire data to make an initial diagnosis which serves for screening, some form of aggregation, or totalling scheme must be used. For example, an index could be created which could be devised as the sum of all the responses. The higher the index reading the higher the probability that the patient suffers from Apnea Syndrome and therefore should be admitted to the Sleep Centre for extensive tests.

6.1 FUNCTION DESCRIPTIONS

In this paper, we consider data aggregation methods based on the use of weighting vectors to bias the data, with respect to relevance and reliability. Three methods are proposed: principal components (PC), ordered weighted average (OWA), and weighted ordered weighted average(WOWA). These methods contrast different weighting factors; PC correlates the input variables in order to reduce the dimensionality in one or more factors, OWA weights the variables, and WOWA weights both the variables and the data values. Weighted mean (WM) is an aggregation technique incorporated in WOWA which has as input a data vector and a weight vector. The weight vector contains one degree of reliability value between 0 and 1, for each corresponding variable.

6.1.1 PC – PRINCIPAL COMPONENTS

PC correlates the input variables in order to reduce the dimensionality in one or more factors.

6.1.2 OWA - ORDERED WEIGHTED AVERAGE

OWA is an 'intelligent' data aggregation method which was originally defined by Yager in [12]. Ordered Weighted Average has two vectors as input: a data vector and a weight vector. The weight vector contains two or more degree of relevance values between 0 and 1. These points are subsequently interpolated to form a continuous function which can be used to interpret all the possible data values OWA permits an AND/OR effect on the data inputs, controlled by the relevance weights.

Definition: A mapping F from

$$I^n \rightarrow I \text{ (where } I = [0, 1])$$

is called an OWA operator of dimension n if associated with F, is a weighting vector ω,

$$\omega = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \ldots.. \\ \omega_n \end{bmatrix}$$

such that

1) $\omega_i \in (0,1)$
2) $\Sigma_i \omega_i = 1$

and where

$$F(a_1, a_2, ..., a_n) = \omega_1 b_1 + \omega_2 b_2 + ... + \omega_n b_n,$$

where b_i is the ith largest element in the collection $a_1, a_2, ..., a_n$. **B** is called an ordered argument vector if each element $b_i \in [0,1]$ and $b_i \geq b_j$ if $j > i$. Given an OWA operator F with weight vector ω and an argument tuple $(a_1, a_2, ..., a_n)$ we can associate with this tuple an ordered input vector **B** such that **B** is the vector consisting of the arguments of F put in descending order. It is important to note that weights are associated with a particular ordered position rather than a particular element.

6.1.3 THE WOWA OPERATOR

Torra in [9] described the Weighted OWA operator (WOWA), which combines advantages of the weighted mean and the OWA operator, thus solving some of the shortcomings of the latter two operators. It considers two weight vectors: ρ corresponding to the relevance of the sources (as in weighted mean), and ω corresponding to the relevance (which we interpret as ‘reliability’) of the values as in OWA. One of the difficulties in using aggregation operators is the initial fixing of the associated parameters, for example the relevance weights ρ of each information source. In [8], Nettleton et al. explain contrasting data analysis methods for determining the weights of the aggregation function.

6.2 USE OF THE RELIABILITY VECTOR ω

In Figures 5a to 5e we can see each of the characteristic curves defined by the E, L, H, O and M vectors of Table 4. These vectors represent the ω vector which can be used to strengthen some responses while diminishing others, as we see in Figures 5a to 5e. For example, in Figure 5b, a response of ‘never’ will be strengthened to affect the (aggregated) outcome more than a response of ‘always’, which will have its contribution to the (aggregated) outcome diminished.

Note that we distinguish this weighting effect for the response reliability from the membership grade of the responses as detailed in Section 5. We can say that the membership grade is reflecting the qualitative information provided by the patient, whereas the ω weighting of the responses reflects the medical experts knowledge of what responses are most expected for each question.

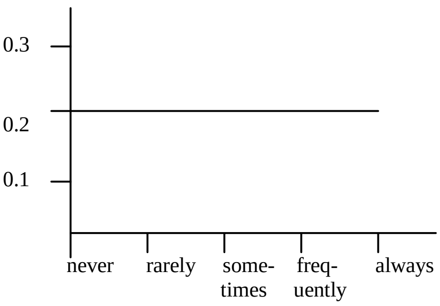


Figure 5a. Even bias vector (E)

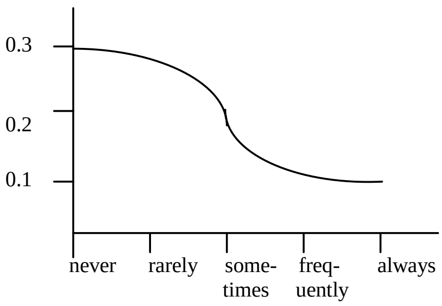


Figure 5b. Low bias vector (L)

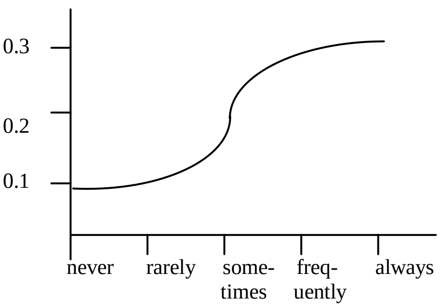


Figure 5c. High bias vector (H)

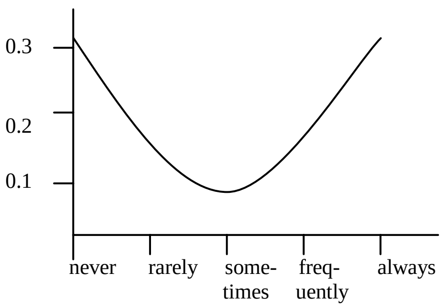


Figure 5d. High & Low bias vector (O)

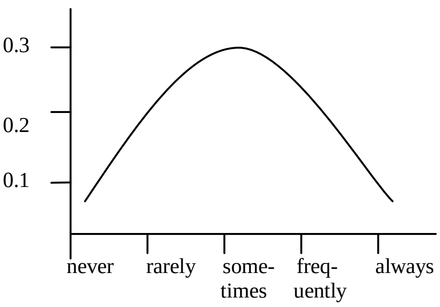


Figure 5e. Middle bias vector (M)

With reference to Table 4, we can see five values defined for each variable. From these value points, WOWA uses the interpolation method of Chen and Otto [1], to create a continuous function curve which can be used to weight all the values of each variable.

To diagnose a patient, WOWA is called thus:

$$\sum_{j=1}^{n} (A_i, V\rho, V\omega_j),$$

where A_i is the data vector for patient **i** (see Table 5), $V\rho$ is the variable weight vector for all variables (see Table 3), and $V\omega_j$ is the data value vector for variable **j** (see Table 4).

7. CONCLUSIONS

This work is being jointly developed with medical and data analysis expertise. We have chosen an area in which there is real room for improvement, due to the lack of precision of existing screening methods (especially for negative case prediction), and the high cost and resource requirements for sleep centre testing. We have considered two fundamental aspects from a data analysis point of view: representation of the data and aggregation.

We have defined smooth curve forms as membership functions for the fuzzy representations of the questionnaire responses. We can compare different s-curves in order to choose the most effective for each question variable. With respect to aggregation, we have compared three contrasting methods for aggregating the data values: Principal Components, OWA and WOWA. The proposed approach is previously untried in the literature of apnea diagnosis, which has tended to focus on multiple linear regression and logistic regression models (Table 1).

At present we have tested the aggregation techniques on a real crisp apnea case data set in collaboration with the Hospital Clinic of Barcelona. This work is summarised in [8] and has demonstrated a good precision for both positive and negative cases. In [8] the data was captured by crisp (categorical) question responses and processed by one standard aggregation algorithm (PC) and two fuzzy aggregation algorithms (OWA, WOWA).

To achieve the next step, that is, to evaluate the aggregation of questionnaire responses captured on a fuzzy (continuous scale), we are currently collaborating with a Sleep Clinic in Salamanca, Spain. Questionnaires are being filled in by patients, and this shall continue until we have approx. 100 cases which will then be processed with the methods described in this paper. The patients fill in two questionnaires, one in the fuzzy/continuous form and the same one but in crisp/categorical form. This will enable us to compare the diagnosis using crisp and fuzzy representation methods.

The work gives a novel approach for questionnaire data capture and processing where linguistic labels and subjective / uncertain inputs play an important role, and enables expert knowledge and statistical knowledge to be incorporated into data processing. The authors are grateful to the Medical Faculty of the University of Barcelona, and to Vicenc Torra of the Institute for Investigation in Artificial Intelligence, Bellaterra, Spain, for their collaboration

References

[1] Chen, Joseph E., Otto, Kevin N. *"Constructing membership functions using interpolation and measurement theory"*. Fuzzy Sets and Systems 73 (1995) 313-327.

[2] Delgado, M., Verdegay, J.L., Vila, M.A. *"On Aggregation Operations of Linguistic Labels"*. International Journal of Intelligent Systems, Vol. 8, 351-370 , John Wiley & Sons, 1993.

[3] Gerstorfer, E., Hellendoorn, H. *"On the Role of Fuzzy Logic in Production Planning"*. Proceedings of the Sixth IEEE International Conference on Fuzzy Systems, Volume III, pp. 1271-1275. (1997).

[4] Guilleminault C., Stoohs R., Clerk A et al. *"From obstructive sleep apnea syndrome to upper airway resistance syndrome: consistency of daytime sleepiness"*. Sleep. 1992; 15: 513-516

[5] Hoffstein, V., Szalai J.P. *"Predictive value of clinical features in diagnosing obstructive sleep apnea"*. Sleep 1993; 16: 118-122.

[6] Katz I., Stradling J., Slutsky A.S., et al. *"Do patients with sleep apnea have thick necks?"* American Review of Respiratory Diseases, 1990; 141: 1228-1231.

[7] Nettleton, D.F., *"Attribute fusion using a heterogeneous representation of crisp and fuzzy medical data"*. Eighth International Fuzzy Systems Association World Congress. Taipei, Taiwan, 1999.

[8] Nettleton, D.F. , Hernández, L., *"Evaluating reliability and relevance for WOWA aggregation of sleep apnea case data"*. Congress of the European Society of Fuzzy Logic and Technology EUSFLAT-ESTYLF-99, Palma de Mallorca, Spain. (1999)

[9] Torra, V., *"The Weighted OWA Operator"*. International Journal of Intelligent Systems, Vol. 12, 153-166. John Wiley & Sons (1997)

[10] Torra, V., *"On the learning of weights in some aggregation operators"*. VIII Congreso Español sobre Tecnologías y Lógica Fuzzy. Universitat Pública de Navarra, Pamplona (1998)

[11] Ward Flemons, W, McNichols, Walter T. *"Clinical prediction of the sleep apnea syndrome"*. Sleep Medicine Reviews, Vol. 1, Nº 1, pp 19-32, 1997.

[12] Yager, Ronald R. *"Families of OWA operators"*. Fuzzy Sets and Systems 59 (1993) pp125-148.

[13] Young T, Palta M, Dempsey J., et al. *"The occurrence of sleep-disordered breathing among middle-aged adults"*. N Engl J Med 1994; 328: 1230-1235.

[14] Zadeh, L.A. *"Similarity Relations and Fuzzy Orderings"*. Information Science, Vol. 3, pp-177-200. Elsevier Science Publishing Company, Inc. (1971).