

## Research Article

# Fuzzy Intelligent System for Patients with Preeclampsia in Wearable Devices

Macarena Espinilla,<sup>1</sup> Javier Medina,<sup>1</sup> Ángel-Luis García-Fernández,<sup>1</sup>  
Sixto Campaña,<sup>2</sup> and Jorge Londoño<sup>3</sup>

<sup>1</sup>University of Jaén, Jaén, Spain

<sup>2</sup>Universidad Nacional Abierta y a Distancia, Bogotá, Colombia

<sup>3</sup>Universidad Pontificia Bolivariana, Medellín, Colombia

Correspondence should be addressed to Macarena Espinilla; [mestevez@ujaen.es](mailto:mestevez@ujaen.es)

Received 1 June 2017; Revised 4 August 2017; Accepted 10 September 2017; Published 12 October 2017

Academic Editor: Pino Caballero-Gil

Copyright © 2017 Macarena Espinilla et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preeclampsia affects from 5% to 14% of all pregnant women and is responsible for about 14% of maternal deaths per year in the world. This paper is focused on the use of a decision analysis tool for the early detection of preeclampsia in women at risk. This tool applies a fuzzy linguistic approach implemented in a wearable device. In order to develop this tool, a real dataset containing data of pregnant women with high risk of preeclampsia from a health center has been analyzed, and a fuzzy linguistic methodology with two main phases is used. Firstly, linguistic transformation is applied to the dataset to increase the interpretability and flexibility in the analysis of preeclampsia. Secondly, knowledge extraction is done by means of inferring rules using decision trees to classify the dataset. The obtained linguistic rules provide understandable monitoring of preeclampsia based on wearable applications and devices. Furthermore, this paper not only introduces the proposed methodology, but also presents a wearable application prototype which applies the rules inferred from the fuzzy decision tree to detect preeclampsia in women at risk. The proposed methodology and the developed wearable application can be easily adapted to other contexts such as diabetes or hypertension.

## 1. Introduction

We are currently living in the age of data; our everyday lives are surrounded by sensors that capture information associated with objects, humans, or environments such as vision sensors, motion sensors, light sensors, and medical sensors. The so-called Internet of Things (IoT) [1] is all around us, producing a huge amount of data that needs to be automatically organized and processed in order to produce easy to understand reports for the users, as trying to deal with this raw data directly is certainly far from the human capabilities.

Data analysis [2] is at the core of a relevant amount of recently published works, many of them focusing on data mining techniques to extract useful knowledge from the data and to shape it so that users understand what is happening and act accordingly. The research on methods to communicate this knowledge in a user-friendly way has led

to the concept of linguistic descriptions [3], which allow humans to abstract useful data into different levels and dimensions, therefore providing interpretability.

In this context, the computing with words (CWW) methodology emulates human cognitive processes to make reasoning processes and decisions in environments of uncertainty and imprecision [4]. The CWW methodology considers that inputs and outputs should be expressed in a linguistic domain in order to be close to human natural language, therefore providing interpretable and understandable results [5, 6].

The application of *soft computing techniques* for the development of methods for data analysis [7] has proved to be very successful. However, these methods produce results that in most cases still require some level of expertise to understand and to use them properly. This is the main motivation for the development of linguistic knowledge extraction techniques, which rely on the principles of fuzzy logic [7] in order

to communicate relevant information obtained from the analysis of the data.

This paper is focused on the health care area, given its relevance nowadays, although the presented techniques could be applied in other areas like wellness [8]. In general, modern health care systems make use of many technological advances to measure biomedical signals that produce large amounts of raw data.

In the context of the early detection of preeclampsia [9, 10], generating user-friendly knowledge to support the diagnosis and monitoring in scalable contexts is a problem that has not been completely solved yet because sometimes a rigid interpretation of the data can lead to misdiagnosis. For example, in [11] it is highlighted that factors like physiology, body size, and variations on the instrumentation can produce differences in the measured heart-rate values that might lead to misdiagnosis.

Therefore, we propose a methodology based on the linguistic approach to extract knowledge by applying a decision tree analysis [12] on a supervised preeclampsia dataset with the aim of early detection of preeclampsia. The proposed methodology is composed of two phases: (i) a linguistic transformation of the dataset to increase the interpretability and flexibility in the analysis of preeclampsia and (ii) the extraction of linguistic knowledge by means of inferring rules using decision trees to classify the dataset.

The use of white-box classification in the cases of decision trees and fuzzy logic [13] provides a wide range of benefits related to real deployments in Health Systems, among which we highlight the following:

- (i) *Increased human interpretability.* It is to identify risky conditions and extract interpretable linguistic knowledge for the medical staff, thanks to the intuitive representation of concepts instead of using numeric values [14].
- (ii) *Independence from health measuring instruments.* In our approach, the values are translated to degrees of membership in different fuzzy sets, and therefore it is more robust to variations in precision or granularity. We highlight that the flexibility achieved by the fuzzification of the attributes is key for translating the classification from the original dataset based on measurements with traditional health instruments to different contexts and advanced medical devices, such as wearable devices.
- (iii) *Inclusion of human observations in the system.* The need for diagnosing preeclampsia in developing countries, where some areas are remote or isolated from health centers, involves integrating human measurements described by itinerant health staff without advanced instruments [15]. In our study case, the age or weight, and even urine infection, can be described by human observation of the patient through linguistic transformation of the data in the decision tree classifier.

The rest of the paper is structured as follows. Section 2 describes the preeclampsia disease and provides some

significant data about its incidence worldwide. Section 3 reviews some notions related to mobile information systems and their application to health. Section 4 describes the proposed fuzzy linguistic methodology. Section 5 presents the results of the case study. Section 6 presents a prototype of wearable application which applies the rules inferred from the obtained fuzzy decision tree. Finally, Section 7 summarizes our conclusions and future work.

## 2. Preeclampsia Disease

In this section, the medical context of preeclampsia is reviewed and global quantitative data is provided in order to highlight the seriousness of the problem.

Preeclampsia (PE) [9, 10] is a multisystemic disorder that occurs during pregnancy, characterized by hypertension and proteinuria (excess of proteins in the urine). PE usually appears after 20 weeks of pregnancy. This condition is one of the most serious complications of pregnancy and also one of the most feared. In severe cases, it endangers the life of both the mother and the fetus [16]. Therefore, it is necessary to diagnose it and begin the treatment as soon as possible. In [17] we also find that severe preeclampsia is associated with an increment of maternal mortality (0.2%) and higher rate of maternal complications (5%) such as convulsions, pulmonary edema, acute renal or liver failures, liver hemorrhage, disseminated intravascular coagulopathies, and strokes. These complications are usually detected in women who develop preeclampsia before the 32nd week of gestation and in those with preexisting medical conditions [18]. PE may be mild or severe, according to the following clinical parameters [19]:

- (i) *Mild PE:* Blood pressure of at least 140/90 mmHg on two occasions six hours apart after the 20th week of pregnancy. Proteinuria greater than 300 mg in 24 hours. Moderate edema and urinary volume in 24 hours greater than 500 ml.
- (ii) *Severe PE:* Blood pressure greater than 160/90 mmHg on two occasions six hours apart after the 20th week of gestation. Systolic blood pressure greater than 60 mmHg over baseline and diastolic blood pressure greater than 30 mmHg over baseline. Proteinuria greater than 5 g in 24 hours, massive edema, and systemic symptoms such as pulmonary edema, headache, visual disturbances, pain in the right hypochondrium, high level of liver enzymes, or thrombocytopenia.

It is important to emphasize that “blood pressure” is one of the most important factors to be controlled in pregnant women at risk of PE. Therefore, it would be very useful to have a device for monitoring this factor, preferably in real time, and including the possibility of remote monitoring the reports generated by the device.

On the other hand, maternal mortality is a phenomenon that continues occurring despite the efforts of health institutions all over the world. In order to have a general view of the problem, we present here a brief description of the statistics

TABLE 1: Statistics related to PE in the world in the last 10 years.

Case study	Year	Results
The Woman with Severe Preeclampsia Who Died from Postpartum Complications [23].	2016/2017	PE is responsible for about 14% of maternal deaths per year in the world.
The World Bank. Trends in Maternal Mortality: 1990 to 2010 [24]	2015	The maternal mortality ratio (MMR), defined as the number of maternal deaths per 100000 live births, was estimated at 216 globally.
Competing Risks Model in Screening for Preeclampsia by Maternal Factors and Biomarkers at 30–34 Weeks' Gestation [25]	2013/2014	PE affects 2-3% of all pregnancies and is a major cause of maternal and perinatal morbidity and mortality.
Reducing Maternal Mortality from Preeclampsia and Eclampsia in Low-Resource Countries—What Should Work? [26]	2012	PE is thought to account for about 15% of maternal mortality worldwide, between 20% and 25% of fetal mortality, and with increased risk of asphyxia and preterm delivery for as much as 25% of neonatal mortality. Preeclampsia is one of the most serious pregnancy complications. The worldwide prevalence of PE ranges from 3% to 8% of pregnancies, affecting a total of 8.5 million women worldwide. PE is responsible for about 18% of maternal deaths and up to 40% of fetal mortality.
Review: Biochemical Markers to Predict Preeclampsia [27]	2010/2011	
World Health Statistics 2016: Monitoring Health for the SDGs [28].	2003/2009	The primary causes of maternal deaths are hemorrhage, preeclampsia, eclampsia, sepsis, or infections and indirect causes.

TABLE 2: PE statistics in Colombia.

Case study	Year	Results
Research Opportunities in Preeclampsia from the Perspective of Primary Prevention: A Reflection [29].	2006	23.8% of maternal mortality cases recorded in health centers were due to this pathology.
Preeclampsia: Facing a Complex Problem from its Physiology [30].	2010	The indicator of maternal mortality was 25.7% for countries such as Colombia and 16.1% for developed countries. Statistics also showed that the main cause of death due to maternal complications in Colombia in 2010 was PE.
Clinical Practice Guidelines for Approaching Pregnancy Associated Hypertensive Complications [31].	2013	35% of maternal mortality was associated with hypertensive disorders such as PE.
Protocol on Public Health Surveillance: Extreme Maternal Morbidity [32].	2014-2015	10,499 cases of maternal death in 2014, increasing by 35% the figures from 2013. By 2015, the number of cases increased to 15,067.

related to maternal mortality in relation to preeclampsia (PE), mainly referred to the last years (see Table 1).

This paper presents a methodology based on the linguistic approach to extract knowledge by applying a decision tree analysis [12] on a supervised preeclampsia dataset from Colombia with the aim of early detection of preeclampsia. For this reason, it is also important to refer to the PE statistics in this country. A statistical summary can be found in Table 2.

It has also been pointed out [20] that the management of PE has not changed significantly over time, possibly as a result of the poor progress in understanding this condition. All these factors motivate our work, which was developed taking as reference a dataset of patients who suffered PE during a five-year time frame, kindly provided by a hospital in Colombia.

In this paper, we analyze this information using data mining techniques and, as a result, we propose a prediction model and a pathology monitoring process that takes into account the values of relevant vital signs, using a fuzzy linguistic system to support the diagnosis of PE. This prediction model has been developed for a mobile device as a prototype of wearable application in order to detect preeclampsia in women at risk.

### 3. Mobile Applications in Health Care

In this section, we review some notions related to mobile information systems and some applications of mobile systems to health. A mobile application is software designed to run on a mobile device. Typically, mobile applications will access data, devices, and other applications from anywhere [21].

One of the features offered by mobile platforms is real-time connectivity, which has facilitated the creation of various tools that efficiently guide the daily activities of their users. Mobile applications have become the preferred way for users to connect from their devices [22] and have different application areas, such as health, industry, commerce, marketing, entertainment, and sports, among others. For the case that we analyze, the specific domain is the one that refers to the application of mobile information systems to health care. To provide a global idea about some applications developed in this context, Table 3 shows some related works.

We can divide the mobile applications for health purposes into two main groups: applications that are designed for specific purposes (treatment of a single pathology) and general purpose applications that are intended to prevent or attend to emergency situations. The common feature of all

TABLE 3: Applications of mobile systems to health.

Application	Year	Service
Fetal heart rate monitor using a mobile phone [33]	2009	A mobile application that analyses the fetal heartbeat and calculates the heart rate using a beat-to-beat accuracy algorithm. This data is sent and stored in a server, and then a midwife can examine it through a web browser [34].
Mobile health system for the treatment of pediatric HIV [35]	2009	A Web based EHR system that uses an embedded comprehensive pediatric HIV knowledge base and a clinical decision system. This system allows physicians to integrate clinical information to manage pediatric HIV.
Mobile phone image transmission for diagnosis [36]	2011	A mobile phone with camera functionalities is used as image transmission unit. This system allows communicating with medical image experts in remote locations. The images are sent using the Multimedia Message Service (MMS) via mobile phones.
Mobile phone pulse oximeter [37]	2012	A mobile phone is used to analyze and display the information received from an oximeter placed on a finger. It can aid physicians in detecting clinical events and making decisions [34].
Portable telemedicine unit [38]	2008	A portable telemedicine unit that combines a mobile telemedicine system with a computer server. Communication between these two elements is set via GSM, CDMA, Internet, or satellite. This system is intended for healthcare services in rural or remote regions and can be used for several health services, such as recording, reporting, and teleconsultation [34].

these applications is that all of them use mobile technologies for their operation, including the monitoring through remote transmissions of the pathology or purpose for which they were created.

In this paper, we present a mobile information system in the context of health, especially, a prototype of wearable application to detect preeclampsia in women at risk. Moreover, the most important variables to be monitored are identified; the values of these variables can be collected from health sensors and/or from human observation straightforwardly.

#### 4. Methodology

In this section, a methodology based on a linguistic approach is presented to extract knowledge by applying a decision tree analysis on a supervised preeclampsia dataset. The aim of this process is to allow early detection of preeclampsia.

The dataset contains quantitative data from patients with pregnancy disorders and risk of preeclampsia in Colombia, as well as a human expert diagnosis label for each case.

Firstly, a series of attributes is identified as key factors for diagnosing preeclampsia, based on the experience of health experts and other academic studies. This process led us to the following selection of attributes:

- (i) *Age*: age is between 13 and 46 years old, as stated in [39].
- (ii) *Body mass index (ICM)*: body mass index is between 22 and 38 Kg/m<sup>2</sup>, as explained in [40].
- (iii) *Trimester of pregnancy*: pregnancy is to be in the second or third trimester.
- (iv) *Blood pressure*: diastolic (DBP) or systolic (SBP) between 80 and 200 mmHg is highly related to preeclampsia, according to [41].
- (v) *Family history*: if the mother of the patient suffered from preeclampsia, it is labeled as first degree; if only one of the patient's grandmothers suffered from

preeclampsia, it is labeled as second degree; otherwise, no label is set. This is pointed out as a risk factor in [42].

- (vi) *Socioeconomic stratum*: it is related to preeclampsia because of its consequences on the supplementation of multivitamins and folic acid [43]. It is described with discrete values in the range [1, 4]. The value of 1 is related to the highest socioeconomic stratum and 4 the lowest, which is the most critical for preeclampsia.
- (vii) *Race/ethnicity*: according to [44], African-American women are more prone to suffer from severe preeclampsia. This attribute is defined with discrete values (*indigenous*, *African-American*, and *mestizo*).
- (viii) *New mother*: this attribute reveals whether it is the first time the patient is pregnant.
- (ix) *Proteinuria*: the presence of excess proteins in the urine. It can be measured by a simple dipstick test that returns a value in a scale ranging from 0 to 8 mg/dl. It is also related to preeclampsia diagnosis, as explained in [45].
- (x) *Preeclampsia label*: this is the target attribute to classify.

The last attribute, *preeclampsia label*, is defined as a discrete attribute with three possible values: *nonpreeclampsia*, *moderate preeclampsia*, and *severe preeclampsia*.

We highlight the fact that the cases included in the dataset were collected because of their complexity and relevance to preeclampsia detection. No trivial episodes have been included in the dataset, and the diagnosis of preeclampsia is divided into moderate and severe. The complexity of the cases in the dataset should be taken into account before presenting the results.

In the next subsections, the two phases of the proposed methodology are presented.

*4.1. Extracting Knowledge by Means of Decision Trees.* The first phase of our methodology is focused on extracting

TABLE 4: Data for trapezoidal shaped representation.

Attribute	Term	Trapezoidal membership function
Systolic pressure	Normal	T (120 mmHg, 120 mmHg, 135 mmHg, 140 mmHg)
Systolic pressure	High	T (135 mmHg, 140 mmHg, 155 mmHg, 160 mmHg)
Systolic pressure	Severe	T (140 mmHg, 150 mmHg, 160 mmHg, 170 mmHg)
Systolic pressure	Very severe	T (160 mmHg, 170 mmHg, 200 mmHg, 200 mmHg)
Diastolic pressure	Normal	T (85 mmHg, 85 mmHg, 85 mmHg, 90 mmHg)
Diastolic pressure	High	T (85 mmHg, 90 mmHg, 105 mmHg, 110 mmHg)
Diastolic pressure	Severe	T (100 mmHg, 105 mmHg, 115 mmHg, 120 mmHg)
Diastolic pressure	Very severe	T (115 mmHg, 120 mmHg, 200 mmHg, 200 mmHg)
Proteinuria	Normal	T (0 g/dl, 100 m/dl, 900 g/dl, 1000 g/dl)
Proteinuria	High	T (500 g/dl, 1000 g/dl, 4500 g/dl, 5000 g/dl)
Proteinuria	Severe	T (2000 g/dl, 4500 g/dl, 5000 g/dl, 5000 g/dl)
Age	Girl	T (13 y, 13 y, 15 y, 18 y)
Age	Young	T (15 y, 20 y, 30 y, 35 y)
Age	Elder	T (30 y, 38 y, 38 y, 38 y)
Age	Risk	T (35 y, 38 y, 40 y, 40 y)
Weight (ICM)	Normal	T (20 Kg/m <sup>2</sup> , 20 Kg/m <sup>2</sup> , 20 Kg/m <sup>2</sup> , 25 Kg/m <sup>2</sup> )
Weight (ICM)	Overweight	T (23 Kg/m <sup>2</sup> , 25 Kg/m <sup>2</sup> , 30 Kg/m <sup>2</sup> , 30 Kg/m <sup>2</sup> )
Weight (ICM)	Obesity	T (25 Kg/m <sup>2</sup> , 30 Kg/m <sup>2</sup> , 35 Kg/m <sup>2</sup> , 35 Kg/m <sup>2</sup> )
Weight (ICM)	Severe obesity	T (30 Kg/m <sup>2</sup> , 35 Kg/m <sup>2</sup> , 35 Kg/m <sup>2</sup> , 35 Kg/m <sup>2</sup> )

knowledge from the dataset on pregnant women with high risk of preeclampsia using decision trees. In this stage, decision trees are used to infer rules to classify the dataset.

Among all the standard decision tree strategies, we propose using the C4.5 statistical classifier [46] to analyze the dataset. This classifier is shaped as a tree-like graph for continuous and discrete attributes [47] in which the most relevant attributes are located in upper layers. The hierarchy is built taking into account the entropy of each attribute [48], which is a measurement of the homogeneity and the relevance of the information of an attribute with respect to the target attribute to be classified.

The result of applying a decision tree is a white-box model [49], in which the tree-like graph can be easily understood and modified by humans. Moreover, it can be translated into an inference rule system in which each path, from the root to each end node, represents an induction rule [50].

When analyzing real-world datasets, the model tends to increase the tree size to obtain the best accuracy in the classification. However, this may result in the overfitting of the tree, as well as making it more difficult to interpret; therefore, pruning techniques are applied to reduce the tree complexity [51]. Pruning decreases the size of the tree-like graph by removing those nodes whose instances present low entropy. In Section 5, the effects of applying two different pruning techniques on real data are discussed.

**4.2. Linguistic Fuzzy Transformation.** The fuzzy linguistic approach is included in the second phase of the classification of the dataset in order to increase the interpretability of the resulting decision tree. Fuzzy logic has been proposed for this stage because of the increased expressiveness of learning methods based on fuzzy logic predicates [52].

We have integrated a fuzzification step in which we relate each numeric attribute with an attribute described using linguistic terms as presented in [13]. In this step, each attribute with continuous values has been translated to a linguistic variable with different terms. The fuzzification consists of describing each attribute by one variable whose terms are related to a membership degree between [0, 1]. For example, *systolic pressure: 138 mmHg* would be described as *systolic pressure is normal (0.8)* and *systolic pressure is high (0.2)*. Next, the continuous value of each attribute is replaced by the linguistic term with the highest membership degree.

The membership functions of the linguistic terms have been provided by health experts, which defined the values and ranges for each attribute based on their experience and the attribute relevance as a risk factor. The attributes, linguistic terms, and membership functions proposed are detailed in Table 4 using a trapezoidal shaped representation, as shown in Figure 1.

## 5. Evaluation of Results

In this section, a real evaluation of the methodology proposed in this paper is carried out in order to show the efficiency of our proposal.

We have used a dataset from patients with a diagnosis of possible preeclampsia collected in the Departmental Hospital of Nariño (Colombia); this meant an arduous task of compilation with private and public funding. According to medical experts, the dataset contains a representative sample of preeclampsia cases in the last 5 years in this area and consists of 729 records with the attributes and values described in Section 4. We present the results of analyzing

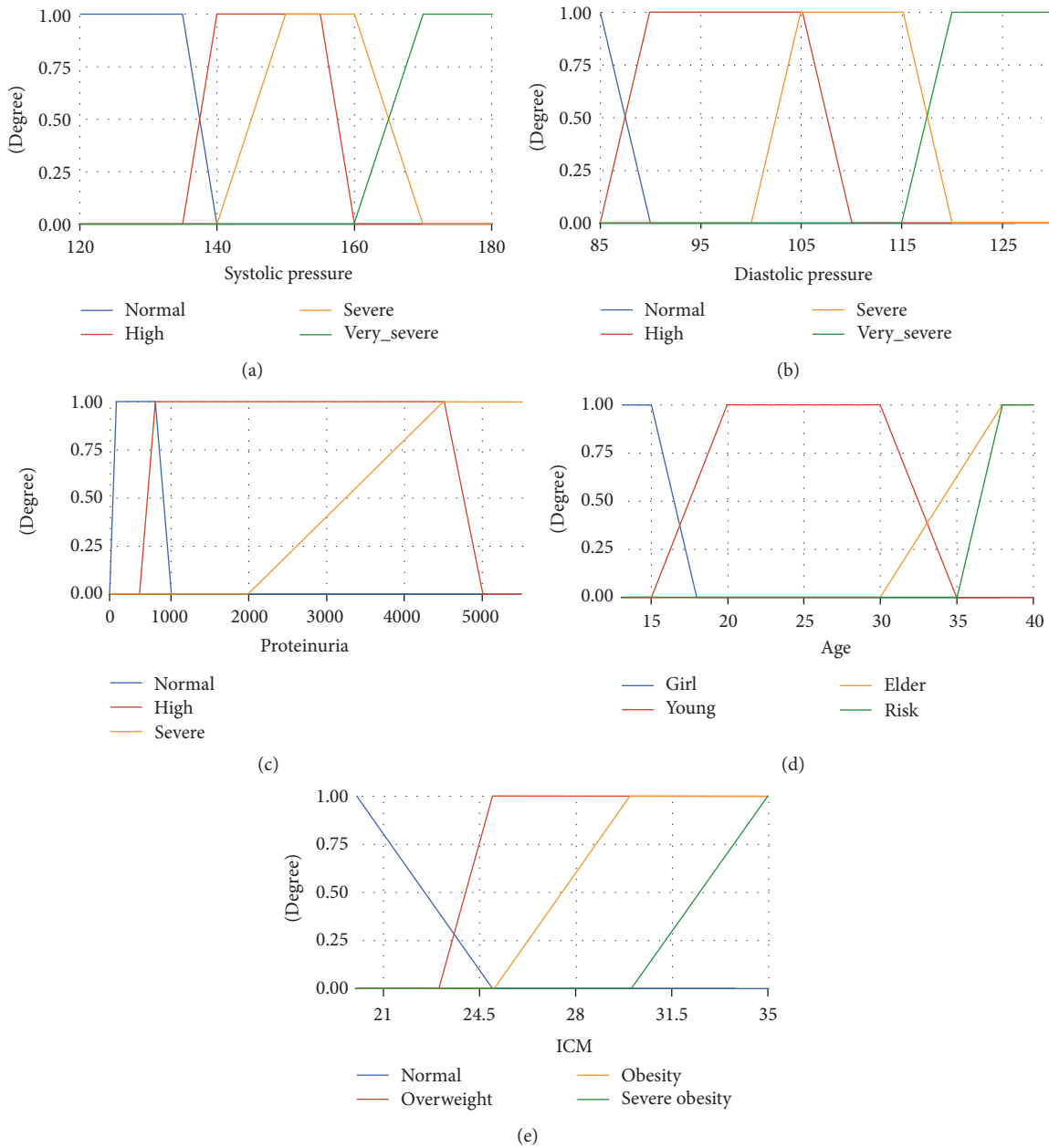


FIGURE 1: (a) Linguistic terms and membership functions related to systolic pressure. (b) Linguistic terms and membership functions related to diastolic pressure. (c) Linguistic terms and membership functions related to proteinuria. (d) Linguistic terms and membership functions related to the age of the patient. (e) Linguistic terms and membership functions related to ICM.

the dataset using the open source implementation of C4.5 available in Weka [53], called J48 [54].

The initial approach was to classify the overall dataset using simple C4.5 without pruning techniques and without a limit in the number of required instances for creating a tree node. The accuracy obtained was 99.45% but at the cost of a tree size of 598 nodes. As we stated before, the problems of overfitting and interpretability can be reduced at the expense of accuracy using pruning. For this reason, in the second and third attempt, the accuracy ratio is computed on the test set by means of a leave-one-out cross-validation, which is a

particular case of  $f$ -cross-validation when  $f$  is equal to 1. The main advantage of this validation is that all the activities in the dataset are used for training and testing, therefore avoiding the problem of considering how the dataset is divided.

In the second attempt, we applied two pruning techniques (one was applied on the finished tree and the other one was applied during the tree construction); these were the following: (i) reducing the generated nodes with a confidence factor of entropy equal to 0.25 and (ii) disabling nodes with just one instance in construction. This approach produced a notable accuracy of 82.16%, with a tree size of 236 nodes.

```

Age ≤ 20
  Etnia = mestizo
    Trimester of pregnancy = 3T
      Body Mass Index ≤ 28 THEN Risk of preeclampsia is severe
      Body Mass Index > 28 THEN Risk of preeclampsia is moderate
    Trimester of pregnancy = 2T
      Age ≤ 18 THEN Risk of preeclampsia is moderate
      Age > 18 THEN Risk of preeclampsia is severe
    
```

Box 1: Example of rules generated from original data.

```

patient is a girl
  patient with severe obesity
    Proteinuria is not normal
      patient without antecedents
        Etnia = african-american THEN Risk of preeclampsia is severe
        Etnia = indigenous THEN Risk of preeclampsia is moderate
        Etnia = mestizo THEN Risk of preeclampsia is moderate
      patiente with antecedets THEN Risk of preeclampsia is moderate
    
```

Box 2: Example of interpretable rules generated by linguistic approach.

TABLE 5: Results of the methods proposed for classifying preeclampsia patients.

C4.5 features	Accuracy	Relative size	Advantages	Weak points
(i) No pruning (ii) No limit in number of nodes	0.9945	100%	(i) Good accuracy	(i) Low interpretability because of the tree size (ii) Overfitting (iii) Difficult extrapolation out of the dataset
(i) Postpruning (confidence factor = 0.25) (ii) Online pruning (number of instances/node ≥ 2)	0.8216	39%	(i) Encouraging accuracy (ii) Increased interpretability due to the reduction in the tree size	(i) Difficult extrapolation out of the dataset (ii) Dependence on measuring instruments (iii) Not good accuracy
(i) Fuzzy linguistic representation (ii) Postpruning (confidence factor = 0.25) (iii) Online pruning (number of instances/node ≥ 2)	0.7503	32%	(i) Increased interpretability due to the linguistic approach (ii) Flexibility on measuring instruments (iii) Human observation of attributes	(i) Loss of precision

For the third attempt, we evaluated the fuzzy approach of tree classification, integrating the linguistic terms and membership functions proposed by health experts as described in Section 4.2. It resulted in a tree with 197 nodes and an improvement of interpretability at the expense of a lower accuracy of 75.03%. Obviously, the generation of more scalable classification system based on knowledge involves a loss of accuracy, which is related in [55, 56] as *a reasonable small loss of accuracy for the sake of interpretability*. In exchange, it can be adapted to other contexts using human observation or different health devices, such as a wearable.

We include a fragment of the rules generated from the original data and the linguistic approach, respectively, in Boxes 1 and 2. Note the interpretability of the linguistic approach which can deal with human perception of age and symptoms.

We have summarized the results of analyzing the preeclampsia dataset in Table 5.

## 6. Development Using Wearable Devices

In this section, we present a prototype of wearable application which contains the rules from the fuzzy decision tree detailed in Sections 4 and 5. It has been developed using the Android Wear platform. As we have described in this paper, the aim of the methodology is to generate medical knowledge that can be applied with the aid of wearable devices in situations in which access to health centers is restricted.

Firstly, the wearable application collects the input values for the attributes to monitor from the patient. This information can be collected in different ways:



FIGURE 2: (a) Introducing input information with linguistic terms, (b) getting the measurement value from a wireless health device, including the fuzzification and representation of the value with a pie circle chart (green for *normal* and yellow for *severe*), and (c) representation of the degree of preeclampsia with a pie circle chart (green for *nonpreeclampsia*, yellow for *moderate preeclampsia*, and red for *severe preeclampsia*).

- (i) Measurements from wireless health devices, such as wireless blood pressure monitors connected through Bluetooth Low Energy (BLE). In this case, the continuous value of the device is replaced by a linguistic term under the fuzzification described in Section 4.2.
- (ii) Linguistic terms related to human interpretation, for example, overweight or obesity.
- (iii) Linguistic terms related to human observation of colored strips from portable urine tests. It is successful, for example, for visually evaluating the excess of proteinuria [57] in a cheap and portable way.

Secondly, the linguistic terms are evaluated in the decision tree integrated in the wearable device, providing a matching with the inferred rules.

Finally, the degree of risk of preeclampsia in the unit interval is provided as the degree of membership to each target class (*nonpreeclampsia*, *moderate preeclampsia*, and *severe preeclampsia*). Figure 2 shows the wearable application.

Two features have prevailed in the software development for the wearable device: on the one hand, experimentability, that allows the tool to be used to evaluate and experiment with different linguistic rules generated by different datasets and on the other hand, maintainability, due to the use of frameworks and standard, well-documented programming languages. Due to these features, the software can be easily modified and fixed for further development.

## 7. Conclusions and Future Work

Early detection of preeclampsia is an important worldwide problem that should be addressed. This paper has presented a solution to support the diagnosis and monitoring of this disease, overcoming the limitation of a rigid interpretation of the data that can usually lead to misdiagnosis. In this paper, biomedical signals are used to identify risk conditions and

extract linguistic knowledge for the medical staff. A methodology based on a fuzzy linguistic approach that provides interpretable results is proposed. The proposed methodology is composed of two phases that include a knowledge extraction by means of decision trees and a linguistic transformation of the data. Therefore, this methodology allows a linguistic monitoring in real time. A real evaluation of the proposed methodology has been carried out, providing good results and offering interpretable rules for monitoring. Furthermore, a prototype of a wearable application, which applies the rules from the fuzzy decision tree derived from the analysis, has been presented. Finally, our future work is focused on trying to obtain a preeclampsia dataset from another health center, preferably from another country, and applying the proposed methodology in order to compare and analyze the inferred linguistic rules for the new dataset.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement no. 734355, the Spanish government with the TIN2015-66524-P Project, and the Asociación Universitaria Iberoamericana de Postgrado.

## References

- [1] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): a vision, architectural elements, and future directions," *Future Generation Computer Systems*, vol. 29, no. 7, pp. 1645–1660, 2013.
- [2] H. Tanaka, "Fuzzy data analysis by possibilistic linear models," *Fuzzy Sets and Systems. An International Journal in Information Science and Engineering*, vol. 24, no. 3, pp. 363–375, 1987.



- [3] A. Ramos-Soto, A. Bugarin, S. Barro, and J. Taboada, "Automatic Generation of Textual Short-Term Weather Forecasts on Real Prediction Data," *Fuzzy Systems, IEEE Transactions*, vol. 23, no. 1, pp. 44–57, 2013.
- [4] L. A. Zadeh, "From computing with numbers to computing with words—from manipulation of measurements to manipulation of perceptions," *IEEE Transactions on Circuits and Systems. I. Fundamental Theory and Applications*, vol. 46, no. 1, pp. 105–119, 1999.
- [5] L. Martínez, D. Ruan, and F. Herrera, "Computing with words in decision support systems: An overview on models and applications," *International Journal of Computational Intelligence Systems*, vol. 3, no. 4, pp. 382–395, 2010.
- [6] J. M. Mendel, L. A. Zadeh, E. Trillas et al., "What computing with words means to me: Discussion forum," *IEEE Computational Intelligence Magazine*, vol. 5, no. 1, pp. 20–26, 2010.
- [7] L. A. Zadeh, "Fuzzy logic, Neural networks, and soft computing," *Communications of the ACM*, vol. 37, no. 3, pp. 77–84, 1994.
- [8] T. M. Uribe, "El autocuidado y su papel en la promoción de la salud," *Investigación y Educación en Enfermera*, vol. 17, no. 2, 2013.
- [9] C. W. G. Redman, G. P. Sacks, and I. L. Sargent, "Preeclampsia: an excessive maternal inflammatory response to pregnancy," *American Journal of Obstetrics & Gynecology*, vol. 180, no. 2, pp. 499–506, 1999.
- [10] L. M. Yamaleyeva and S. H. Lindsey, "Potential for miRNAs as Biomarkers and Therapeutic Targets in Preeclampsia," *Hypertension*, vol. 69, no. 4, pp. 580–581, 2017.
- [11] J. M. Goldman, "The challenge of acquiring accurate, complete, near-patient clinical data for data science analysis," in *Data Science Symposium*, p. 43, 2014.
- [12] J. R. Quinlan, "Simplifying decision trees," *International Journal of Man-Machine Studies*, vol. 27, no. 3, pp. 221–234, 1987.
- [13] M. E. Cintra, M. C. Monard, and H. A. Camargo, "A fuzzy decision tree algorithm based on c4. 5," *Mathware & Soft Computing*, vol. 20, pp. 56–62, 2013.
- [14] Y. Zheng, L. Peng, L. Lei, and Y. Junjie, "R-C4.5 decision tree model and its applications to health care dataset," in *Proceedings of the 2005 International Conference on Services Systems and Services Management, ICSSSM'05*, vol. 2, pp. 1099–1103, Chongqing, China, June 2005.
- [15] P. Singh, "Employing UNICEF Open Source Software Tools in mHealth Projects in Nicaragua," in *Ubiquitous Computing and Ambient Intelligence*, vol. 10069 of *Lecture Notes in Computer Science*, pp. 286–293, Springer International Publishing, 2016.
- [16] C. I. Arroyo Vasquez, "Factores de riesgo independientes para la presencia de preeclampsia," 2014.
- [17] B. M. Sibai, "Diagnosis and management of gestational hypertension and preeclampsia," *Obstetrics & Gynecology*, vol. 102, no. 1, pp. 181–192, 2003.
- [18] B. M. Sibai, M. A. Villar, and B. C. Mabie, "Acute renal failure in hypertensive disorders of pregnancy. Pregnancy outcome and remote prognosis in thirty-one consecutive cases," *American Journal of Obstetrics and Gynecology*, vol. 162, no. 3, pp. 777–783, 1990.
- [19] L. A. V. Egan and S. P. C. Peña, "Conceptos actuales sobre la preeclampsia/eclampsia," *Revista de la Facultad de Medicina UNAM*, vol. 50, no. 2, 2007.
- [20] F. A. English, L. C. Kenny, and F. P. McCarthy, "Risk factors and effective management of preeclampsia," *Integrated Blood Pressure Control*, vol. 8, pp. 7–12, 2015.
- [21] J. G. Enriquez and S. I. Casas, "Usabilidad en aplicaciones móviles," *Informes Científicos-Técnicos UNPA*, vol. 5, no. 2, pp. 25–47, 2014.
- [22] L. Ramirez, E. Guillen, and Y. Cifuentes, "Validation strategy to mobile health applications Estrategia de validación para aplicaciones móviles de salud," *Actas de Ingeniería*, vol. 2, pp. 325–333, 2016.
- [23] A. M. Gashi, "The woman with severe preeclampsia who died from postpartum complications," *Revista Cubana de Obstetricia y Ginecología*, vol. 42, no. 4, pp. 519–523, 2016.
- [24] WHO, UNICEF, and UNFPA, *The World Bank. Trends in Maternal Mortality: 1990 to 2010*, WHO, UNICEF, UNFPA, and The World Bank Estimates, 2012.
- [25] A. Tsiakkas, Y. Saiid, A. Wright, D. Wright, and K. H. Nicolaidis, "Competing risks model in screening for preeclampsia by maternal factors and biomarkers at 30–34 weeks' gestation," *American Journal of Obstetrics and Gynecology*, vol. 215, no. 1, pp. 87.e1–87.e17, 2016.
- [26] R. L. Goldenberg, B. Jones, J. B. Griffin et al., "Reducing maternal mortality from preeclampsia and eclampsia in low-resource countries - What should work?" *Acta Obstetrica et Gynecologica Scandinavica*, vol. 94, no. 2, pp. 148–155, 2015.
- [27] U. D. Anderson, M. G. Olsson, K. H. Kristensen, B. Åkerström, and S. R. Hansson, "Review: biochemical markers to predict preeclampsia," *Placenta*, vol. 33, pp. S42–S47, 2012.
- [28] World Health Organization, *World Health Statistics 2016: Monitoring Health for the SDGs Sustainable Development Goals*, World Health Organization, 2016.
- [29] L. A. Díaz Martínez and N. C. Serrano Díaz, "Oportunidades de investigación en preeclampsia, desde la perspectiva de prevención primaria. Un artículo de reflexión. A paper aimed at inviting reflection on," *Revista Colombiana de Obstetricia y Ginecología*, vol. 59, no. 3, pp. 206–215, 2008.
- [30] C. Piedrahita Ochoa and B. Agudelo Jaramillo, "Preeclampsia: un problema complejo para enfrentar desde su fisiología," *Medicina & Laboratorio*, vol. 16, no. 11/12, 2010.
- [31] A. Cinets, "Guía práctica clínica para el abordaje de las complicaciones hipertensivas asociadas al embarazo. Clinical practice guidelines for approaching pregnancy-associated hypertensive complications," *Revista Colombiana de Obstetricia y Ginecología*, vol. 64, no. 3, pp. 289–326, 2013.
- [32] M. L. Ospina, M. E. Martínez, O. E. Pacheco, and H. Q. Bonilla, "Protocolo de Vigilancia en Salud Pública," *Morbilidad Materna Extrema*, 2016.
- [33] C. S. Lee, M. Masek, C. P. Lam, and K. T. Tan, "Advances in fetal heart rate monitoring using smart phones," in *Proceedings of the 9th International Symposium on Communications and Information Technology (ISCIT '09)*, pp. 735–740, IEEE Press, Icheon, Republic of Korea, September 2009.
- [34] B. M. C. Silva, J. J. P. C. Rodrigues, I. de la Torre Díez, M. López-Coronado, and K. Saleem, "Mobile-health: a review of current state in 2015," *Journal of Biomedical Informatics*, vol. 56, pp. 265–272, 2015.
- [35] S. Paul, S. Bhattacharya, A. Sudar et al., "A web-based electronic health care system for the treatment of pediatric HIV," in *Proceedings of the 11th International Conference on e-Health Networking, Applications and Services (Healthcom '09)*, pp. 175–180, Sydney, Australia, December 2009.
- [36] L. Bellina and E. Missoni, "M-learning: Mobile phones' appropriateness and potential for the training of laboratory technicians in limited-resource settings," *Health and Technology*, vol. 1, no. 2–4, pp. 93–97, 2011.

- [37] D. Dunsmuir, C. Petersen, W. Karlen, J. Lim, G. A. Dumont, and J. M. Ansermino, "The phone oximeter for mobile spot-check," *Anesthesia and Analgesia*, vol. 115, no. 2S, p. 21, 2012.
- [38] E. Sutjiredjeki, S. Soegijoko, T. L. R. Mengko, and S. Tjondronegoro, "Development of mobile telemedicine system with multi communication links to reduce Maternal Mortality Rate," in *Proceedings of the 6th IASTED International Conference on Biomedical Engineering, BioMED 2008*, pp. 402–407, ACTA Press, Anaheim, Calif, USA, February 2008.
- [39] J. S. Moldenhauer, J. Stanek, C. Warshak, J. Khoury, and B. Sibai, "The frequency and severity of placental findings in women with preeclampsia are gestational age dependent," *The American Journal of Obstetrics and Gynecology*, vol. 189, no. 4, pp. 1173–1177, 2003.
- [40] T. E. O'Brien, J. G. Ray, and W.-S. Chan, "Maternal body mass index and the risk of preeclampsia: A systematic overview," *Epidemiology*, vol. 14, no. 3, pp. 368–374, 2003.
- [41] M. A. Brown, M. D. Lindheimer, M. de Swiet, A. van Assche, and J.M. Moutquin, "The classification and diagnosis of the hypertensive disorders of pregnancy: statement from the International Society for the Study of Hypertension in Pregnancy (ISSHP)," *Hypertension in Pregnancy*, vol. 20, no. 1, pp. 9–14, 2001.
- [42] R. Arngrimsson, S. Björnsson, R. T. Geirsson, H. Björnsson, J. J. Walker, and G. Snaedal, "Genetic and familial predisposition to eclampsia and pre-eclampsia in a defined population," *BJOG: An International Journal of Obstetrics & Gynaecology*, vol. 97, no. 9, pp. 762–769, 1990.
- [43] S. W. Wen, X. K. Chen, M. Rodger et al., "Folic acid supplementation in early second trimester and the risk of preeclampsia," *American Journal of Obstetrics and Gynecology*, vol. 198, no. 1, pp. 45–e1, 2008.
- [44] A. A. Goodwin and B. M. Mercer, "Does maternal race or ethnicity affect the expression of severe preeclampsia?" *American Journal of Obstetrics and Gynecology*, vol. 193, no. 3, pp. 973–978, 2005.
- [45] E. Schiff, S. A. Friedman, L. Kao, and B. M. Sibai, "The importance of urinary protein excretion during conservative management of severe preeclampsia," *American Journal of Obstetrics and Gynecology*, vol. 175, no. 5, pp. 1313–1316, 1996.
- [46] J. R. Quinlan, *C4. 5: Programs for Machine Learning*, Elsevier, 2014.
- [47] J. R. Quinlan, "Improved use of continuous attributes in C4.5," *Journal of Artificial Intelligence Research*, vol. 4, pp. 77–90, 1996.
- [48] C. E. Brodley and M. A. Friedl, "Decision tree classification of land cover from remotely sensed data," *Remote Sensing of Environment*, vol. 61, no. 3, pp. 399–409, 1997.
- [49] B. Delibasic, M. Vukicevic, and M. I. L. O. Jovanovic, "White-box decision tree algorithms: A pilot study on perceived usefulness, perceived ease of use, and perceived understanding," *International Journal of Engineering Education*, vol. 29, no. 3, pp. 674–687, 2013.
- [50] J. R. Quinlan, "Generating Production Rules from Decision Trees," in *IJCAI*, vol. 87, pp. 304–307, 1987.
- [51] J. Mingers, "An Empirical Comparison of Pruning Methods for Decision Tree Induction," *Machine Learning*, vol. 4, no. 2, pp. 227–243, 1989.
- [52] O. Cordon, "A historical review of evolutionary learning methods for Mamdani-type fuzzy rule-based systems: Designing interpretable genetic fuzzy systems," *International Journal of Approximate Reasoning*, vol. 52, no. 6, pp. 894–913, 2011.
- [53] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: an update," *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 1, pp. 10–18, 2009.
- [54] N. Bhargava, G. Sharma, R. Bhargava, and M. Mathuria, "Decision tree analysis on j48 algorithm for data mining," *Proceedings of International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 6, 2013.
- [55] S. Guillaume and B. Charnomordic, "Generating an interpretable family of fuzzy partitions from data," *IEEE Transactions on Fuzzy Systems*, vol. 12, no. 3, pp. 324–335, 2004.
- [56] J. M. Alonso and L. Magdalena, "Special Issue on Interpretable Fuzzy Systems," *Information Sciences*, vol. 181, no. 20, pp. 4331–4339, 2011.
- [57] D. A. De Silva, A. C. Halstead, A.-M. Côté, Y. Sabr, P. von Dadelzen, and L. A. Magee, "Urinary Dipstick Proteinuria Testing: Does Automated Strip Analysis Offer an Advantage Over Visual Testing?" *Journal of Obstetrics and Gynaecology Canada*, vol. 36, no. 7, pp. 605–612, 2014.

