Expert Systems With Applications 219 (2023) 119621

Contents lists available at ScienceDirect

Expert Systems With Applications

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

# Flexible evaluation of electrotherapy treatments for learning purposes

Antonio J. Rueda [a], Carmen Martínez-Cruz [a,*], Ángeles Díaz-Fernández [b], María Catalina Osuna-Pérez [b]

[a] Center for Advanced Studies in Information and Communication Technologies (CEATIC), University of Jaén, Campus de las Lagunillas, s/n., 23071, Jaén, Spain
[b] Department of Health Sciences, University of Jaén, Campus de las Lagunillas, s/n., 23071, Jaén, Spain

ARTICLE INFO

*Keywords:*
Expert system
Fuzzy logic
OWA
e-learning
Serious game
Electrotherapy

ABSTRACT

Computers are extensively used for training and simulating tasks that are hazardous in a real-life situation or require expensive, delicate or difficult to access equipment. In the present work we describe a system for the assessment of clinical practices involving the use of electrotherapy. The existing variety of currents, electrical parameters and electrode configurations make this a challenging subject for undergraduate physical therapy students. Extensive training is required before they can treat real patients, since an incorrect practice may have no effect on the patient's condition, or even worse, cause severe pain or burns. Our system, implemented as a serious game, can facilitate this training, since the opportunity for supervised clinical practice with real equipment is limited.

Our goal was to mimic the assessment of an electrotherapy treatment by an expert, by evaluating each of its aspects in a flexible way, detecting and highlighting critical situations that can affect the patient's integrity and computing an overall score for the whole treatment. For this purpose, we have modeled the valid values of each electrical parameter (frequency, pulse width, intensity, etc.) and the treatment time using fuzzy sets. We also characterize the placement of the electrodes in the area of treatment through three geometric quantities (orientation, distance to center and spacing) on which three fuzzy terms are defined (well oriented, centered and well distributed). For each of these variables we get partial scores that are finally aggregated into an overall treatment score through a pessimistic exponential ordered weighted average operation. In order to test the reliability of the system we conducted an experiment with a group of physical therapy students that revealed a substantial agreement with a human expert.

## 1. Introduction

As early as the times of the Ancient Greece and Rome, the electricity generated by certain fishes was used for alleviating paralysis and pain. The invention of the battery in the 19th century started the golden era of electrotherapy, in which electricity was used for the treatment of many diseases, sometimes without a rigorous scientific basis. Due to the latter, it felt into some disgrace for some time, and it was not until the second half of the 20th century that there was a renaissance of electrotherapy, driven by new scientific developments that explained better it benefits. Today electrotherapy is commonly used in physical therapy by its proven anti-inflammatory and analgesic effects, and its utility to strengthen the muscle fibers for rehabilitation or improvement in sports performance.

An electrotherapy treatment consists of two steps: electrode placement and device configuration. The first step includes taking several decisions such as choosing the area of treatment, the number and size of the electrodes to use and their orientation in the treatment area (lengthwise or crosswise). The second step implies choosing the appropriate current and configuring its parameters, such as impulse time, frequency, type of modulation, intensity and time of treatment. These parameters and their value depend on the type of current applied, therapeutic objective (analgesic, anti-inflammatory, strengthener) and even the patient's physical characteristics. The example below illustrates a typical electrotherapy treatment.

*Case study.* *Patient with severe pain in the lower back after lifting heavy loads a few days ago. On examination he presents with bilateral contracture of the lower dorsal and lumbar paravertebral muscles with pain on palpation. He shows an antalgic position with high functional limitation.*

In view of this clinical picture, the physical therapist would apply an electrotherapy treatment in the lumbar region of the patient, placing two $6 \times 8$ cm electrodes crosswise. There are several currents that can

* Corresponding author.
*E-mail addresses:* ajrueda@ujaen.es (A.J. Rueda), cmcruz@ujaen.es (C. Martínez-Cruz), andiaz@ujaen.es (Á. Díaz-Fernández), mcosuna@ujaen.es (M.C. Osuna-Pérez).

https://doi.org/10.1016/j.eswa.2023.119621
Received 29 October 2021; Received in revised form 16 March 2022; Accepted 26 January 2023
Available online 30 January 2023

**Table 1**
Most common currents in electrotherapy and their main parameters.

| Current | Electrical and treatment parameters |
|---|---|
| Galvanic direct/ interrupted | CC, PC, TT: 10–15 min, I: liminal (max 0.15 mA/cm$^2$) |
| Iontophoresis | CC, TT: 10-15 min, I: liminal (max 0.15 mA/cm$^2$) |
| Dyadinamic DF/ MF/LP/CP | CC, TT: 10-15 min, I: liminal |
| Dyadinamic RS | CC/CV, TT: 10-15 min, I: supraliminal |
| Trabert | CC/CV, PW: 2 ms, RT: 5 ms, F: 142.5 Hz, PC, TT: 15–20 min, I: max. tolerable |
| TENS with asymmetric biphasic waveform | CC, PW: 100–200 $\mu s$, F: 80–120 Hz, M: frequency/amplitude, TT: > 60 min, I: supraliminal |
| TENS with symmetric biphasic waveform | CC, PW: 200–300 $\mu s$, F: 50 Hz, M: pulse trains, TT: 10–25 min, I: supraliminal with visible contraction |
| Interferential bipolar | CC, F: carrier at 4000 Hz and AMF at 80–120 Hz, M: spectrum, TT: > 60 min, I: supraliminal |
| Interferential tetrapolar | CC, F: carrier at 4000 Hz and AMF at 80–120 Hz, M: vector scan, TT: > 60 min, I: supraliminal |
| Interferential (russian stimulation) | CV, F: carrier at 2500 Hz and AMF at 50 Hz, M: pulse trains, TT: 10-25 min, I: supraliminal with visible contraction |

CC: constant current, CV: constant voltage, PC: polarity change, PW: pulse width, RT: rest time, F: frequency, AMF: amplitude modulation frequency, M: modulation, TT: treatment time, I: intensity.

be effective; one of them is the *transcutaneous electrical nerve stimulation (TENS) with an asymmetrical biphasic waveform*. The electrical parameters that have to be adjusted are the *pulse width* (100 to 200 μs) and the *pulse frequency* (80 to 120 Hz). The *intensity* of the current must be set to the subjective *supraliminal* level of the patient, usually between 10 and 30 mA. Finally the *treatment time* is set between 60 and 240 min according to the number of sessions received so far and the availability of the specialist.

The above example illustrates the complexity of the process. An error in any of the steps or parameters can lead to a treatment without any positive effect or in the worst case, cause severe pain in the patient. Certainly, modern electrotherapy equipment has several predefined current setups for most common treatments but still the physical therapist should have a deep knowledge of the different currents and its effects, and be able to fine-tune their physical parameters to personalize the treatment. Table 1 illustrates the variety of currents and parameters that have to be mastered by the specialist.

Because of the explained above, undergraduates in physical therapy receive a solid instruction in the matter. To support the learning process, a group of associated professors at the University of Jaén from the physical therapy and computer science fields developed an application for tablets to help the students to learn the concepts of electrotherapy in an interactive way (Díaz-Fernández et al., 2016). This application shows the most common areas of treatment, currents and their electrical parameters, and the time of treatment depending on the different treatment goals (analgesic, anti-inflammatory or muscle stimulation).

Looking for a way to improve learning and engage students more in this sometimes dry matter, we resolved to apply *gamification* techniques (Kapp, 2012) by adding a *serious game* mode to the application. This game shares the same screens but goes beyond by challenging students to figure out the correct treatment for a clinical case, without providing any guidance or clue, as in the final exam of the course. In the game mode, a first screen shows the description of a clinical case (Fig. 1(b)). Then, the student chooses the type and number of electrodes, placing them in the appropriate area of the body (Fig. 1(c)). Next, a simulated electrotherapy device is shown for the selection of the current family, current and configuration of its electrical parameters (Figs. 1(d)–1(f)). In the treatment screen the student has to choose the duration of the treatment and apply the proper intensity following the feedback of the patient (Fig. 1(g)). Finally, the game finishes showing a table with the detailed evaluation of each of the aspects of the treatment (Fig. 1(h)).

The last step is key for the success of the game and ideally, it should make a similar assessment to that of the professors of the course. This assessment is a complex task since it involves many different variables, not all of the same importance, and many of them inherently fuzzy (electrode position in the zone of treatment, chosen intensity, time of treatment, etc.), since sometime there is not a clear distinction between *correct* and *incorrect*. Considering all the valid values for the variables as crisp would lead to an unrealistic and too strict assessment, which could result in frustration for the students.

In addition to allowing students to perform simulated clinical practices and to evaluate their knowledge of the subject, a major benefit of the application is to allow students to practice with a realistic electrotherapy device at home, since unfortunately, the time available to practice with real electrotherapy equipment in the laboratory is limited.

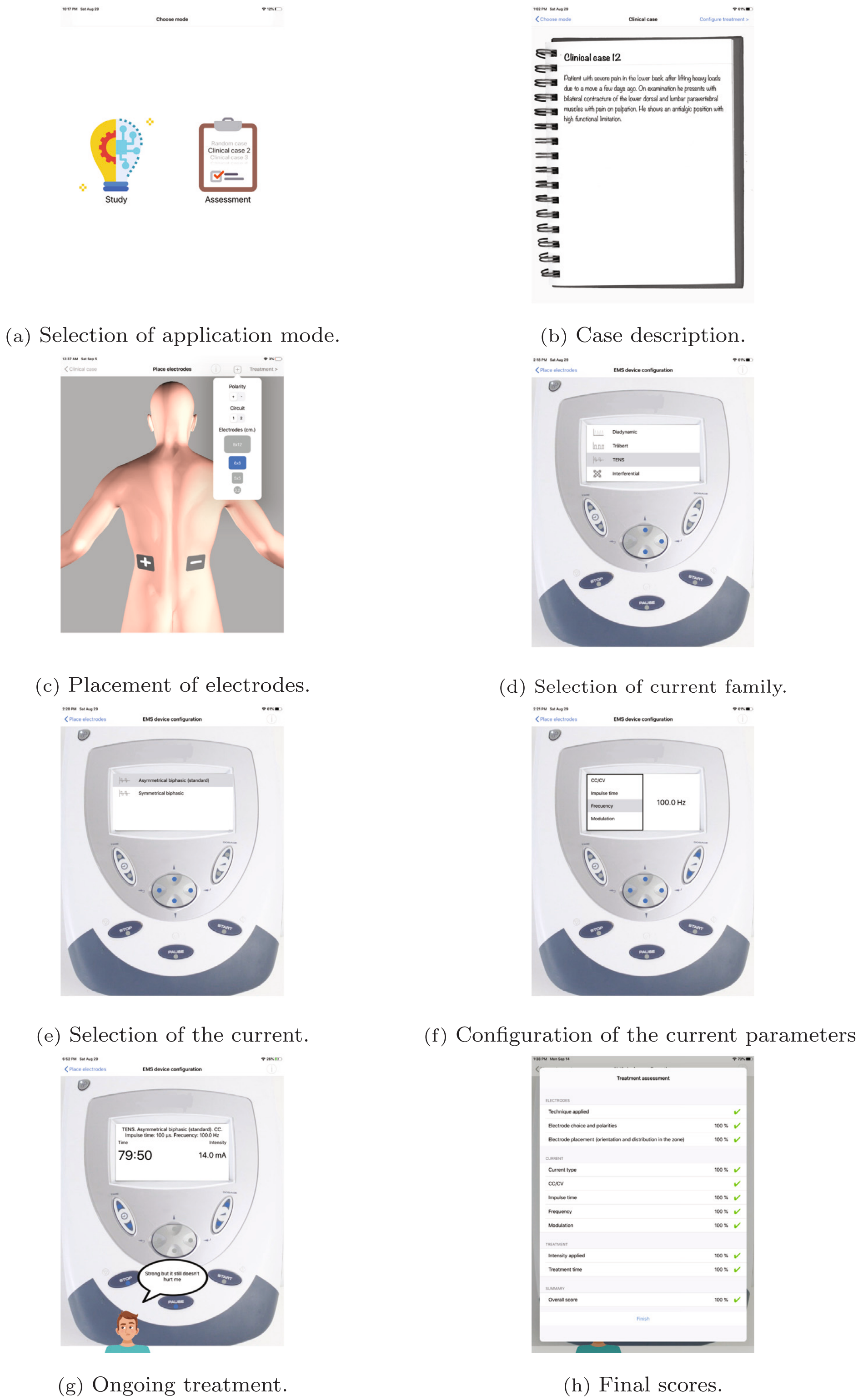
This work describes a flexible assessment system for electrotherapy treatments in the context of a serious game that resembles that of a human expert. It includes fuzzy geometric features together to more traditional fuzzy representations for the right values of the electrical parameters and the modeling of the different subjective sensitive levels of a patient. In the next sections we describe every aspect of the assessment, including the calculation of the final overall score.

## 2. Background

The evaluation process in any learning system involves the examination of an arbitrary number of variables. Variables whose values are limited and well-defined make this process simple. However, there are parameters whose evaluation have an implicit degree of uncertainty or subjectivity (Pangaro, 2000), due to factors such as the number of evaluators involved or their previous background (Rasmani & Shen, 2006). One solution to handle the inherent imprecision in this evaluation process, widely studied in the literature, is to use the fuzzy logic theory, defined by Zadeh (1965). This brings the automatic evaluation process closer to that performed by a human. It helps instructors to establish grades in a more natural way through linguistic terms, and in the case of students, it facilitates the interpretability of their results. For example, Annabestani et al. (2020) described an evaluation system for students which uses a set of fuzzy-modeled indicators and previous assessments to overcome the limitations of a descriptive evaluation. Zhou et al. (2001) proposed a method that uses fuzzy logic to assess students projects with evaluations performed with different formats. Chai et al. (2015) also used fuzzy logic to deal with the use of words in the assessment process applied to cooperative learning. Ma and Zhou (2000) implemented a method to define the evaluation criteria and weights using fuzzy logic in order to engage the students in their learning process. Capuano et al. (2017) proposed a method to perform a peer assessment based on the principles of fuzzy group decision making.

In general terms, we can classify existing methods into two groups: (i) those that use fuzzy logic and aggregation methods, such as the OWA operator defined by Yager (1988), to allow the combination of heterogeneous or purely fuzzy variables (Carlsson & Fullér, 1997) and, (ii) the proposals that implement a Fuzzy Inference System (FIS) with a set of fuzzy rules, such as those defined by Mamdani and Assilian (1975) or Takagi and Sugeno (1985), to compute an overall evaluation.

Belonging to the first group, where the present work is included, there are many applications such as that of Kwok et al. (2007) who present a fuzzy multi-criteria decision making (MCDM) model to solve assessment problems in group projects, related to the use of different criteria and scores in different formats. Also, it uses as aggregation method the OWA operator together with the definition of a fuzzy




(a) Selection of application mode. (b) Case description. (c) Placement of electrodes. (d) Selection of current family. (e) Selection of the current. (f) Configuration of the current parameters. (g) Ongoing treatment. (h) Final scores.

**Fig. 1.** Screenshots of the electrotherapy serious game.

quantifier to create the weight vector. This approach requires a big effort to configure the system by the professor when any format is changed. Andayani et al. (2017) propose a method to assess student achievements using a dataset of student grades, given in numerical and linguistic format, which is transformed into a 2-tuple linguistic approach. The aggregation is performed with an arithmetic mean. This method presents two problems: it needs to transform every single variable into a 2-tuple format because it cannot deal with quantitative and qualitative information at the same time, and the results are very dependent of outliers. Nguyen et al. (2017) perform the evaluation of personal selection of student-internship by using fuzzy linguistic terms and a fuzzy OWA operator. This system represents user/position characteristics into fuzzy values, modeled by 2-tuple linguistic model, which are compared to make a suitability raking. This proposal requires a complex definition process to include new offers and merits. In the field of e-learning and peer assessment, there are many proposals where fuzzy logic and fuzzy aggregation operators are applied. However, these systems are mainly focused in the representation of the relationship among the different actors and items valuated. This is the case of Chang and Chen (2009) that develop a fuzzy peer assessment system (FPAS) to satisfy the requirements of cooperative e-learning environments. Despite using fuzzy logic to represent the preference relationships and final evaluation, the assessment of the variables uses only crisp measures. In Capuano et al. (2017), it is implemented a peer assessment system in massive open online courses. A classic OWA operator is used to aggregate student evaluations. Because their goal is peer assessment, the definition of the preferences of some users with respect to others using fuzzy relationships is needed, thus adding complexity to the model. In contrast to all these proposals, our system is much simpler since only one student is evaluated at a time and the modeled variables are of two types, crisp and fuzzy. Fuzzy variables are modeled with fuzzy member functions defined by our experts. Crisp values are discriminating, and consequently, aggregation is performed on the fuzzy variables using the pessimistic exponential OWA operator (Filev & Yager, 1998) to qualify the evaluation in a more restrictive way, similar to the method used by the professor. Also, our proposal makes the transformation into fuzzy data using a real simulation (serious game) where the user performs the definition of the fuzzy values transparently and intuitively.

Belonging to the second group, there are many proposals that solve the student assessments problem by defining a FIS, which returns a semantic and understandable output, such as that of Echauz and Vachtsevanos (1995) who propose a fuzzy-based grading system to generate a fair mark distribution by using the student scores and instructor performance that modify a set of fuzzy grades. This proposal represents scores with a membership function similar to ours, but evaluation is done with a FIS and clustering operations. The definition of the model and professor evaluation is the most challenging part of this proposal. Samarakou et al. (2017) develop a system to assist the assessment of students with respect to their learning process. Similar to ours, the system models the inputs in a fuzzy way but it uses a FIS to evaluate several variables that describe the type of student that is being evaluated. However, results are not transformed into any final score. Annabestani et al. (2020) use the Mamdani inference model to classify each student's assessment. Here, it performs a continuous evaluation of students by defining a FIS with two inputs, one of which is the output of the previous assessment. This approach only works if the weight of the two inputs is always the same. Aziz et al. (2019), Chuang et al. (2015) and Jamsandekar and Mudholkar (2013) built alternative FIS to calculate the performance of the students. In the first one, different aspects of the academic performance are considered, which include class attendance and marks obtained in class tests. The second one uses a game-based system to assess their creativity, and the third one defines a FIS to monitor the progress and provide timely guidance to students in order to achieve a better performance score. Previous approaches are interesting because of their data modeling strategies and the solutions provided. However, in the evaluation process we used an aggregation operator instead of a FIS, since we have a hybrid system with many variables of two different types that have to be summarized in a numerical rating, and a defuzzification process cannot provide it to the extent we need.

Table 2, summarizes the different proposals analyzed here. This comparison highlights the main contributions of the present paper with respect to previous ones.

Throughout the present work, variables that represent flexible concepts are defined by the following fuzzy sets (Dubois, 2000; Zadeh, 1965):

T1: Trapezoidal shape

$$T1(x; l_1, l_2, l_3, l_4) = \begin{cases} 0 & x \le 0 \\ (x - l_1)/(l_2 - l_1) & l_1 \le x \le l_2 \\ 1 & l_2 \le x \le l_3 \\ (l_4 - x)/(l_4 - l_3) & l_3 \le x \le l_4 \\ 0 & l_4 \le x \end{cases}$$

T2: Triangular shape

$$T2(x; l_1, l_2, l_3) = \begin{cases} 0 & x \le 0 \\ (x - l_1)/(l_2 - l_1) & l_1 \le x \le l_2 \\ (l_3 - x)/(l_3 - l_2) & l_2 \le x \le l_3 \\ 0 & l_3 \le x \end{cases}$$

TR: Right shoulder shape

$$TR(x; l_1, l_2) = \begin{cases} 1 & x \le l_1 \\ (l_2 - x)/(l_2 - l_1) & l_1 \le x \le l_2 \\ 0 & l_2 \le x \end{cases}$$

TL: Left shoulder shape

$$TL(x; l_1, l_2) = \begin{cases} 0 & x \le l_1 \\ (x - l_1)/(l_2 - l_1) & l_1 \le x \le l_2 \\ 1 & l_2 \le x \end{cases}$$

D: Discrete value

$$D(x; l_1) = \begin{cases} 1 & x = l_1 \\ 0 & x \ne l_1 \end{cases}$$

where $x$ is the input value to be represented by the fuzzy set and $l_n$ are the vertices of the shape.

### 2.1. Motivation

As discussed in the previous section, fuzzy logic has been mainly used in the context of education to overcome the subjectivity of the evaluator or group of evaluators and to communicate the results of the evaluation in a natural way through linguistic terms. Although this may also be relevant to our problem, or primary motivation was to enable a flexible evaluation of naturally fuzzy concepts such as the vertical/horizontal orientation of the electrodes, their centered placement or correct distribution in a zone of the body. In the same way, certain current-related variables such as the intensity or frequency admit a range of valid values with gray areas in the extremes where values are not the ideal but definitively not wrong either. A crisp evaluation of these variables would lead to artificially bad results in many cases, causing frustration in the students and loss of interest in the application.

To sum up, the contributions of the present work are outlined next:

- A flexible evaluation of variables of a heterogeneous nature: discrete and continuous, scalar and spatial. As far as we know, this is the first use of fuzzy logic to evaluate spatial variables for learning purposes.
- The use of the pessimistic exponential OWA (POWA) (Filev & Yager, 1998) to simulate the evaluation of an expert in the overall assessment, that tends to penalize partial variables with particularly low scores. Again, as we discussed in previous section,
- The integration of the intelligent system in an e-learning application not as a simple evaluation/auto-evaluation tool as in previous approaches but as a serious game that mimics a real clinical practice.

A comparison of the main characteristics of this proposal with respect to others is included in Table 2. Here, it is analyzed the following issues: (i) the use of fuzzy variables, (ii) the use of crisp variables, (iii) the modeling of spatial variables using fuzzy logic, (iv) the use of OWA aggregators, (v) the use for e-learning purposes, (vi) integration into serious game environment. In this table, it can be observed how the model presented here, despite being framed in a student evaluation model, has some very specific particularities, which make its comparison with others very complicated. Specifically, there is no student assessment proposal using serious game that models spatial variables with e-learning objectives. There are many hybrid models or models that use aggregation using OWA and fuzzy logic, although not the method implemented here.

**Table 2**
Comparison of proposals in the literature.

| References | Fuzzy var. | Crisp var. | Spatial var. | Aggreg. OWA | e-learning | Serious game |
|---|---|---|---|---|---|---|
| Martinez-Cruz et. al (this one) | X | X | X | PE OWA | X | X |
| Kwok et al. (2007) | X | – | – | OWA | – | – |
| Andayani et al. (2017) | X | X | – | X | – | – |
| Nguyen et al. (2017) | X | X | – | FOWA | – | – |
| Chang and Chen (2009) | X | X | – | – | X | – |
| Capuano et al. (2017) | X | – | – | OWA | X | – |
| Echauz and Vachtsevanos (1995) | X | – | – | – | – | – |
| Samarakou et al. (2017) | X | – | – | – | – | – |
| Annabestani et al. (2020) | X | – | – | – | – | – |
| Aziz et al. (2019) | X | – | – | – | – | – |
| Chuang et al. (2015) | X | – | – | – | – | X |
| Jamsandekar and Mudholkar (2013) | X | – | – | – | – | – |

## 3. Description of the model

Two of the authors of the present paper are experts in electrotherapy, and have been teaching this matter to the students in the Bachelor's Degree in Physical Therapy at the University of Jaén for more than 15 years. This teaching relies extensively on the resolution of simulated clinical cases by the students, that are evaluated according to several aspects that have been translated as close as possible to 16 different variables in the application.

These variables are organized into three sections: (i) *Electrodes*, (ii) *Current*, and (iii) *Treatment*, as shown in Table 3. From these, 6 are categorical variables evaluated as simple yes/no or correct/incorrect and 10 are continuous variables with a 0%–100% score, modeled as fuzzy variables. The membership functions of these variables have been proposed by our experts, according to the range of valid values and the degree of flexibility allowed in its evaluation.

Several of the variables are critical, in the sense that a wrong value implies that none of the rest of variables of the section are subsequently evaluated, since it would not make sense. This indeed represents a major error, and as a result, the overall score would be zero or very penalized. For example, if a wrong electrode technique is selected, the electrode choice or their placement are not evaluated. In the same way, if a non-suitable current is chosen, the rest of variables related to the electrical parameters are not evaluated. Even worse, if the student exceeds the safety intensity level for the selected current or if the patient suffers burns due to an excessive intensity, the overall score is directly set to zero, ignoring the evaluation of the partial variables. The electrode placement variable is the most subjective and complex to interpret, and therefore required a careful study that resulted in a model that is described in detail in the next section.

At the implementation level, the descriptions of the clinical cases and their correct treatment, including the electrode placement details and the currents that are suitable, are coded in a JSON configuration file with a simple format that can be examined and edited by the electrotherapy experts.

In the next three sections we describe in detail the evaluation of each of the variables.

**Table 3**
Variables of the treatment evaluated.

| Section | Variable | Evaluation |
|---|---|---|
| Electrodes | Technique[a] | Correct/incorrect |
| | Choice | Correct/incorrect |
| | Placement | 0%–100% |
| Current[b] | Type[a] | 0%–100% |
| | CC/CV | Correct/incorrect |
| | Pulse width | 0%–100% |
| | Rest time | 0%–100% |
| | Frequency | 0%–100% |
| | Carrier Frequency | 0%–100% |
| | AMF | 0%–100% |
| | Polarity change | Correct/incorrect |
| | Modulation | 0%–100% |
| Treatment | Safety limits exceeded[c] | Yes/no |
| | Burns[c] | Yes/no |
| | Intensity | 0%–100% |
| | Time | 0%–100% |

[a]Critical variable: a *wrong* value implies that the rest of variables in the section are not evaluated.
[b]Variables in this section are dependent on the chosen current type.
[c]Major error: a *yes* value implies that the intensity and time variables are not evaluated and the overall score is set to zero.

### 3.1. Electrode selection and placement

This is the most involved part of the evaluation. The *technique* and *electrode choice* are relatively simple variables evaluated as *correct/incorrect* but the *placement* implies several geometric measurements that are evaluated in a flexible way.

*Electrode technique.* The electrode technique refers to the number of electrodes used during the treatment: bipolar (two electrodes), tetrapolar (four electrodes) and monopolar (two electrodes with the positive electrode twice the size of the negative one). The *technique* variable checks that the technique chosen by the student is one of the allowed for the clinical case, but also ensures that the size restriction for the monopolar technique is met and that the electrodes are equally-sized for the bipolar and tetrapolar techniques. Professional electrotherapy devices usually have more than one electrical circuit. The evaluation

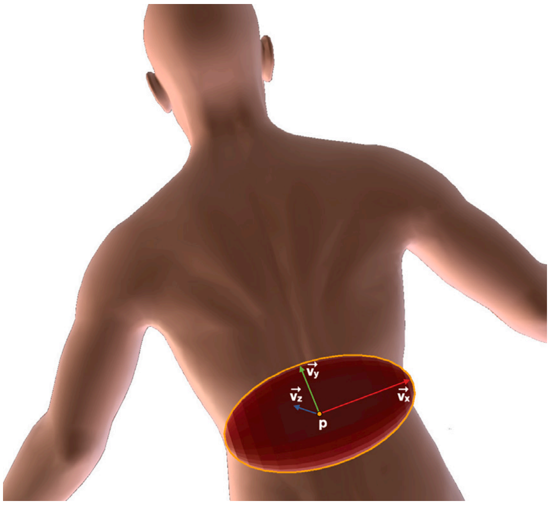



**Fig. 2.** Upper lumbar area defined by an enclosing ellipsoid.

of this variable also ensures that the electrodes are connected to the same circuit in the monopolar and bipolar techniques. In the tetrapolar technique there should be two circuits with two electrodes each. If any of the previous conditions is not met the technique is considered to have been applied incorrectly and the *technique* variable is set to *wrong*. As a result neither the electrode choice nor its placement in the treatment area are evaluated.

*Electrode choice.* The application supports the most common electrode sizes: 12 × 8 cm and 6 × 8 cm rectangular electrodes, 5 × 5 cm square electrode and 3.2 cm diameter round electrode. The size of the electrodes depend on the extend of the area of the body to be treated and the current used. The *electrode choice* variable checks that the size of the electrodes is one of the allowed in the configuration file for the given clinical case, and also checks that the polarities are correct, that is, there is an equal number of positive and negative electrodes. In the particular case of the tetrapolar method, the previous condition must be met for the electrodes in each circuit.

*Electrode placement: orientation, centering and distribution.* The evaluation of the electrode placement is much more involved. First we had to find a way to define each area of the body in an approximate way, since in general the boundary of these areas is blurry, and neighboring areas usually overlap. For instance, the elbow area may overlap with the forearm and arm areas. Then we had to agree with our physical therapy experts on what it means that an electrode is well placed in the treatment area. We reduced this to three conditions that we can model geometrically: correctly oriented, correctly centered and correctly distributed in the area. The term *correct* in these three conditions is inherently fuzzy, and therefore is modeled by a fuzzy set.

Our physical therapy experts have identified 52 areas in the human body that are relevant in electrotherapy. We have defined each of these areas through an oriented ellipsoid, as depicted in Fig. 2. This surface, although simple, provides us with enough flexibility to represent each area with a reasonable accuracy. An oriented ellipsoid $E$ is characterized by its center $p$ and three orthogonal vectors: $\vec{v_x}$, $\vec{v_y}$, and $\vec{v_z}$ that define its shape and orientation. Any point $q$ inside $E$ verifies:

$$|\overrightarrow{(q-p)} \cdot \vec{v_x}| \leq 1; |\overrightarrow{(q-p)} \cdot \vec{v_y}| \leq 1; |\overrightarrow{(q-p)} \cdot \vec{v_z}| \leq 1 \tag{1}$$

Given two electrodes centered at points $e_1$ and $e_2$, its orientation is defined by the angle defined by vector $v_y$ and the vector between $e_1$ and $e_2$ (Fig. 3):

$$o(E, e_1, e_2) = \frac{|\overrightarrow{(e_1 - e_2)} \cdot \vec{v_y}|}{\|\overrightarrow{e_1 - e_2}\| \cdot \|\vec{v_y}\|} \tag{2}$$

This gives a value between 0 and 1, where 0 means perfect crosswise orientation and 1 perfect lengthwise orientation. In order to evaluate the orientation in a flexible way, we defined the crosswise and lengthwise geometric terms by using two fuzzy sets whose membership function are $TL(o; 0.2, 0.35)$ and $TR(o; 0.65, 0.8)$ respectively, as depicted in Fig. 3.

Two electrodes are centered if the midpoint of the vector connecting them is close to the center of the ellipsoid (Fig. 4). This distance is computed as follows:

$$c(E, e_1, e_2) = \frac{\|\frac{1}{2}\overrightarrow{(e_1 + e_2)} - \vec{p}\|}{\max\{\|\vec{v_x}\|, \|\vec{v_y}\|, \|\vec{v_z}\|\}} \tag{3}$$

The denominator normalizes it to a $[0, 1]$ value. We define the linguistic term *centered* as a fuzzy set with the membership function $TL(c; 0.1, 0.7)$ over this normalized distance.

In order to assess the distribution of the electrodes in the treatment area we have to measure the distance between them respect to the longest possible spacing (Fig. 5):

$$s(E, e_1, e_2) = \frac{\|\overrightarrow{e_1 - e_2}\|}{2\|o(E, e_1, e_2)\vec{v_y} + \sqrt{1 - o(E, e_1, e_2)^2}\vec{v_x}\|} \tag{4}$$

The expression in the denominator is the length of the longest possible axis of the ellipsoid in the orientation of the electrodes, depicted as a dotted line in Fig. 5. A well distributed pair of electrodes in the area should not be neither too close nor too far apart. This is represented as a fuzzy set with membership function $T1(s; 0.25, 0.5, 0.7, 0.95)$, as shown in the same figure.

These three geometric measures are also computed for the monopolar technique. The tetrapolar method has some specific aspects that, for the sake of simplicity, we only sketch here. There is a first geometric condition that must be met: the segments connecting the electrodes of the two circuits must intersect, as shown in Fig. 6. A simple intersection test for two 2D line segments can be used for this purpose (Vince, 2017), after projecting the four electrodes in the local 2D coordinate system defined by the axis $\vec{v_x}$ and $\vec{v_y}$. The orientation is computed as the average of the vectors connecting the electrodes of the two circuits $e_1 - e_2$ and $e_3 - e_4$. Likewise, the distance to the center of the ellipsoid (Eq. (3)) is computed from the midpoint of the four electrodes. Finally, the four electrodes are well distributed if the pair of electrodes in each of the two circuits are well distributed.

### 3.2. Current selection and configuration of electrical parameters

Similarly to the electrode technique, the *current type* is a critical parameter that enables the evaluation of the rest of variables in the section. A particular clinical case can accept several currents, although their score could not necessarily be 100%, since some of them can be more suitable than others. An unsuitable current (score 0%) would promptly terminate the evaluation of this section.

As shown in Table 1 the electrical parameters and their proper values are dependent on the current type. The *CC/CV* and *polarity change* parameters take simple discrete values (Yes/No) that can be evaluated as *correct/incorrect* but the rest of parameters are values in an interval that are better evaluated using fuzzy logic. For instance, the pulse frequency in a TENS current with an asymmetrical biphasic waveform should be set between 80 to 120 Hz, but slightly lower (60 to 80 Hz) or higher values (120 to 140 Hz) could be acceptable, although with a small penalty. These parameters and the representation of their correct values as fuzzy sets are detailed in Table 4. The *treatment time* parameter (TT), also shown in this table, is evaluated together with the intensity in the section below.

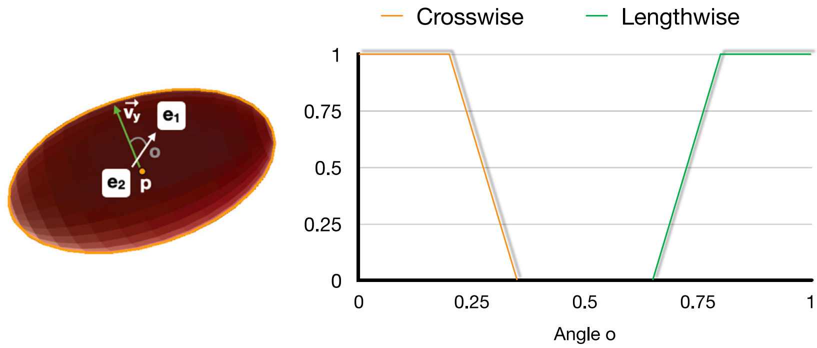



**Fig. 3.** Visual representation of the angle between the electrodes and the $\overline{v_y}$ axis of the ellipsoid ($o(E, e_1, e_2)$), and definition of the *crosswise* and *lengthwise* terms.

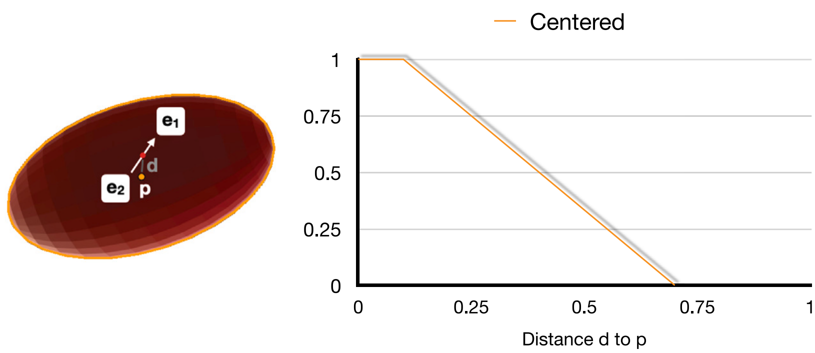



**Fig. 4.** Visual representation of the distance between the midpoint of the electrodes and the center of the ellipsoid ($c(E, e_1, e_2)$), and definition of the *centered* term.

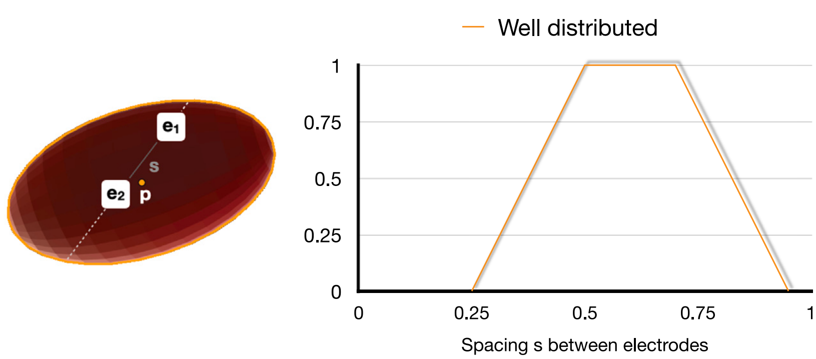



**Fig. 5.** Visual representation of the spacing between the electrodes ($s(E, e_1, e_2)$), and definition of the *well distributed* term.

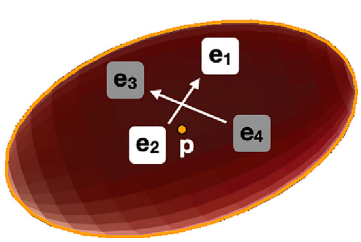



**Fig. 6.** Circuit intersection in the tetrapolar method.

### 3.3. Treatment evaluation

During the simulation of the treatment, the settings to be evaluated are the time of treatment and the intensity applied to the patient. These parameters are different from those of the previous section because although they can be preset, they can also be changed arbitrarily later during the treatment simulation.

The correct treatment time for each current, represented as a fuzzy set, is detailed in Table 4. The correct intensity level cannot, however, be specified as a fixed fuzzy set since it depends on factors such as the sex, weight and age of the patient. In practice, the valid intensity interval for a given current is associated with a *subjective intensity level*, as shown in Table 1. The practitioner has to increase the intensity gradually and identify when it has reached the correct level by taking into account the patient's feedback (Fig. 1(g)). The following subjective intensity levels have been established by agreement:

**Subliminal** The patient does not perceive any sensation.

**Liminal** Barely perceptible, associated with some tingling or prickling sensation.

**Supraliminal motor** The patient experiences involuntary muscle contractions.

**Supraliminal** Strong involuntary contractions. The patient can experience heat with certain currents.

**Table 4**
Current electrical parameters with a fuzzy representation for their correct values.

| Current | PW (ms) | RT (ms) | F (Hz) | AMF (Hz) | TT (minutes) |
|---|---|---|---|---|---|
| Galvanic direct /interrupted | – | – | – | – | T1(5, 10, 15, 15) |
| Iontophoresis | – | – | – | – | T1(5, 10, 15, 15]) |
| Dyadinamic DF/MF/LP/CP /RS | – | – | – | – | T1(5, 10, 15, 15) |
| Trabert | D(2) | T2(2, 5, 8) | T1(137, 141, 142, 146) | – | T1(5, 10, 20, 20) |
| TENS asymmetric biphasic | T1(40, 100, 200, 1000) | – | T1(60, 80, 120, 140) | – | T1(15, 60, 240, 240) |
| TENS symmetric biphasic | T1(40, 100, 200, 1000) | – | T1(60, 80, 120, 140) | – | T2(5, 10, 25, 35) |
| Interferential bipolar | – | – | D(4000) | T1(60, 80, 120, 140) | T1(15, 60, 240, 240) |
| Interferential tetrapolar | – | – | D(4000) | T1(60, 80, 120, 140) | T1(15, 60, 240, 240) |
| Interferential (russian stimulation) | – | – | D(2500) | T2(20, 50, 120) | T1(5, 10, 25, 35) |

PW: pulse width, RT: rest time, F: frequency, AMF: amplitude modulation frequency, TT: treatment time.

**Maximum level of tolerance** Very uncomfortable sensation, close to pain. The patient can suffer burns with certain currents.

**Pain** Never used in electrotherapy. The patient can experience unbearable pain or suffer severe burns.

Depending on the type of current, the patient can perceive different sensations for the same subjective intensity interval. For instance, in the supraliminal intensity level the patient can experience a strong muscle contraction with a TENS current with biphasic symmetrical waveform or an intense heat with a galvanic current. This has been implemented in the application, so that the virtual patient's reactions vary with different currents and intensities. Also, the subjective levels are randomly generated within certain reasonable limits, simulating different patient's body characteristics, as in the sample depicted in Fig. 7. This makes the simulation of the clinical practice more realistic and prevents the student from memorizing the valid intensity for each current.

The evaluation of the intensity has more subtleties. The galvanic and iontophoresis currents are particularly dangerous if a relatively high intensity per unit area of the electrodes is applied, which may result in severe burns. The safety rule is that the intensity must not be higher than $0.15\ \mathrm{mA} \times \mathrm{cm}^2$. If this rule is violated at anytime during treatment, the overall score for this section is zero points. Beyond this, all currents cause severe pain or burns if the *pain* subjective intensity level is reached during treatment. Again if this occurs, the score is zero.

### 3.4. Overall score

The overall score could be obtained with a simple aggregation method such as an ordinary or weighted average, but our experts highlighted that as a general rule, bad scores in any variable should significantly penalize the overall score. For this reason we chose a pessimistic exponential ordered weighted average operator (PEOWA) (Filev & Yager, 1998), defined as follows:

$$f(a_1, \dots, a_n) = \sum_{j=1}^{n} w_j b_j \tag{5}$$

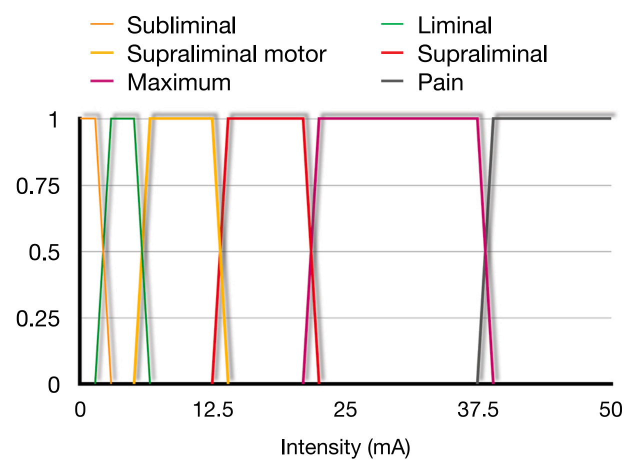



**Fig. 7.** Sample subjective intensity levels for a virtual patient.

where $a_1 \dots a_n$ are $n$ scores, $b_j$ is the $j$th largest score, and $w_i$ are the associated weights:

$$w_1 = \alpha^{n-1}; w_2 = (1-\alpha)^{n-2}; w_3 = (1-\alpha)^{n-3}; \dots ; w_{n-1} = (1-\alpha)\alpha; w_n = (1-\alpha)$$

The parameter $\alpha \in [0, 1]$ defines the importance of good scores respect to bad ones. A value $\alpha = 0$ always returns the worst score, and in general, values close to zero return overall scores closer to the average of the worst scores. After checking the output of the evaluation using different $\alpha$ values with the experts, we found that $\alpha = 0.4$ gives reasonable results.

In order to calculate the score of the electrotherapy treatment, the scores of each of the variables are aggregated into a partial section score. Then, the scores of the three sections are aggregated into the final overall score. The pseudocode for this calculation is illustrated, in simplified form, in Algorithm 1.

**Algorithm 1:** Computation of the overall score by aggregating the different partial scores through the PEOWA operator

**if** $electrodeTechnique = correct$ **then**
  $electrodePlacementScore \leftarrow PEOWA(\alpha, electrodeOrientationScore, electrodeCenteringScore, electrodeDistributionScore)$
  $electrodeSectionScore \leftarrow PEOWA(\alpha, electrodeChoiceScore, electrodePlacementScore)$
**else**
  $electrodeSectionScore \leftarrow 0$
**if** $currentType \neq 0$ **then**
  $currentSectionScore \leftarrow PEOWA(\alpha, currentType, currentCCCV, currentPulseWidth, currentRestTime, \ldots, currentModulationScore)$
**else**
  $currentSectionScore \leftarrow 0$
**if** not $safetyLimitsExceeded$ and not $patientBurned$ **then**
  $treatmentSectionScore \leftarrow PEOWA(\alpha, treatmentIntensityScore, treatmentTimeScore)$
**else**
  $treatmentSectionScore \leftarrow 0$
$overallScore \leftarrow PEOWA(\alpha, electrodeSectionScore, currentSectionScore, treatmentSectionScore)$

## 4. Validation of the results

An assessment of the reliability of the student evaluation made by the application was performed. The method used was the study of interobserver agreement between the evaluation of a clinical practice performed by the described system and that of an expert professor in the subject with more than 15 years of experience. It is important to point out that as we explained in Section 3, the teaching of this subject in the Physical Therapy program uses the resolution of simulated clinical cases as one of the main learning methods.

### 4.1. Methods

The participants were a group of 36 students that recently completed the course on electrotherapy *of the first year of the Bachelor's Degree on Physical Therapy*. The experiments were carried out with two iOS devices: an iPad 2 WLAN+3G 64 GB of 2011 and an iPad Mini 4 64 GB of 2015. Both devices ran the application smoothly and with good readability of the texts displayed by the user interface. Two students were assessed in parallel, maintaining the necessary distance to avoid interactions between them that could affect the result of the evaluation.

Before starting the assessment, the students were taught how to use the application, and they had the opportunity to explore and play with it for five minutes. Then they selected the game mode of the application, that starts by presenting a textual description of a random clinical case out of 12 different options, covering different conditions in different areas of the body. This description is similar to that of the case study that illustrates Section 1, and the corresponding application screen is depicted in shown in Fig. 1(b). After taking the necessary time to fully understand the clinical case the student is prompted to choose the correct electrotherapy treatment, selecting the proper electrodes, technique, type of current, current parameters, time and intensity during the treatment, based on what they learned during the course (Figs. 1c–g). In parallel, the professor observed the decisions made by the student and, based on her expertise, wrote down the scores in a form with the same evaluable items as the application, including the overall score for the clinical practice. This was a blind evaluation: the expert was never aware of the partial or overall scores given by the application (Fig. 1h), that were wrote down in a different form by a second professor. The students did not have a time limit to solve the proposed case, but none needed more than 5 min for this task.

**Table 5**
Inter-rater agreement for categorical variables (correct/incorrect) using the Kappa coefficient ($\kappa$).

| Section | Variable | $\kappa$ | Degree of agreement |
|---|---|---|---|
| Electrodes | Technique | 1 | Excellent |
| | Choice | 1 | Excellent |
| Current | CC/CV | 1 | Excellent |

Data handling and analysis were conducted using the statistical package for social sciences (SPSS) version 21. In order to assess the agreement between the two observers (expert professor and application) in the categorical items (correct/incorrect in Table 3) the Cohen's Kappa coefficient ($\kappa$) (Cohen, 1960) was used. The values of this coefficient are interpreted as degrees of agreement according to Landis and Koch (1977): non-existent if $\kappa < 0$, non-significant when $0.0 \leq \kappa \leq 0.2$, discrete when $0.2 < \kappa \leq 0.4$, moderate if $0.4 < \kappa \leq 0.6$, substantial if $0.6 < \kappa \leq 0.8$, and excellent if $0.8 < \kappa \leq 1.0$. For items rated in a continuous scale (0%–100%), the intraclass correlation coefficient of Shrout and Fleiss (1979) (ICC) was used. Reliability or agreement was considered poor when $ICC < 0.4$, moderate when $0.4 \leq ICC < 0.75$, substantial when $0.75 \leq ICC < 0.9$ and excellent when $ICC \geq 0.9$.

### 4.2. Results

Tables 5 and 6 summarize the results of the experiment. There is a Kappa value of 1 for all the categorical variables (correct/incorrect). The agreement for these items between the expert and the application was perfect. The dichotomous variables *technique* and *choice* in the *electrodes* section are related with very important and basic aspects in the clinical treatment with electrical currents. They are part of the fundamental contents of the matter and are well understood by the students. Their evaluation es clear, simple and without ambiguity: correct or wrong, and as a result there is a total agreement between the expert and the evaluation system of the application. The remaining dichotomic variable *current CC/CV* refers to a characteristic of the electrical currents that is only rarely set to the *CV* value. For this reason in all the proposed clinical cases during the experimentation the *CC* value was chosen by the participants, achieving a 100% correct for the variable and therefore a total agreement between the expert evaluation and the application.

Regarding the items with a 0%–100% score, the agreement varied between moderate and excellent. This agreement was moderate for *electrode placement*, *current modulation* and *intensity applied*, which are the ones with the greatest room for improvement. The choice of the current type, one of the most important aspects of the evaluation, showed a substantial agreement. An excellent agreement was found for the *impulse time*, *frequency* and *treatment time* variables. Finally it is noteworthy that the aggregation of the partial scores into the overall score computed by the application is substantially in line with that of the expert.

As evidenced above, there is room for improvement for several of the variables. The most important is the *electrode placement*. The design of the application organizes the body into relatively wide treatment regions, such as shoulder, calf region or hip, but we found that many students have difficulty to center and distribute the electrodes correctly in small areas such as the elbow, wrist or knee. The expert rater also had problems to match the application score in these areas. This leaded to a less than expected (moderate) inter-rater agreement. Moreover, many students use their anatomical knowledge to put the electrodes on a particular nerve, ligament or tendon, which may not necessarily fall in the center of the region. As a result, although this is perfectly right, the application would not give the highest score. In the experiments, the expert rater also tended to score according to these anatomical structures more than to the treatment region in general. Again this

**Table 6**
Inter-rater agreement for continuous variables (0%–100%) using the intraclass correlation coefficient (ICC).

| Section | Variable | ICC (95%CI) | Degree of agreement |
|---|---|---|---|
| Electrodes | Placement | 0.522 (0.059-0.758) | Moderate |
| Current | Current type | 0.867 (0.741–0.932) | Substantial |
| | Impulse time | 0.908 (0.814–0.954) | Excellent |
| | Frequency | 0.910 (0.818–0.955) | Excellent |
| | Modulation | 0.649 (0.292–0.827) | Moderate |
| Treatment | Intensity | 0.671 (0.339–0.837) | Moderate |
| | Time | 0.962 (0.923–0.981) | Excellent |
| Overall score | | 0.766 (0.470–0.899) | Substantial |

penalized the ICC. These two problems will be addressed in a future version by defining more relaxed fuzzy sets for the evaluation of small regions, and adding more specific and localized treatment regions matching the most relevant anatomical structures.

The *modulation* is a current parameter that shows certain intricacy. Each type of current has different modulation possibilities and each treatment can accept several options with different scores. It is possible that some particularities of this parameter have not been correctly transferred to the application, resulting in a lower agreement than expected. This has to be analyzed and improved in the next version of the application.

As explained in Section 3.3, the correct intensity for a treatment is not fixed: it has to be deduced from the physical sensations manifested by the patient. This adds a extra complexity to the evaluation and can lead to different interpretations of the messages of the patient by the student and expert. These messages and the corresponding fuzzy intensity intervals have to be reviewed with the guidance of our experts to uncover any possible inconsistencies.

Despite these potential areas of improvement, we can conclude that, in general terms, the assessment of the student's knowledge performed by the application is comparable with that of a human expert, and therefore, it is reliable for self-assessment, and potentially, as a learning tool.

## 5. Practical implications

The developed tool improves the assimilation of the practical concepts within a complex field of knowledge, such as clinical treatments using electrotherapy. In addition, students can get better training with it because it solves some of the main problems that they usually face:

(a) Each electric current has different characteristics or parameters that are difficult for students to learn with only theoretical learning. Practicing and/or studying with real electrotherapy devices (outside of a few hours at the laboratory) is practically impossible. This tool allows this learning outside the teaching space with a very similar resource to a conventional electrotherapy device that is expensive and not very accessible for students.
(b) The correct dosage of the intensity is a key point of an electrotherapy treatment. It is always necessary to have feedback of the patient's sensations; this tool provides these important messages from the patient, resembling the subjectivity and diversity that patients present in clinical practice. This objective is impossible to obtain with merely theoretical learning.
(c) This resource has proven to be reliable to learn, practice and evaluate the knowledge acquired autonomously, without being under the supervision and judgment of an expert teacher.

These are the most direct practical implications of this tool. Moreover, incorporating this evaluation system increases the pedagogical power of the application and implies a novel advance in the learning process in general that could be extrapolated to many other contents in the field of health sciences. All subjects that involve well-defined and developed protocols, a common aspect in health disciplines, could potentially benefit from e-learning tools incorporating adapted versions of the evaluation methodology described here.

## 6. Conclusions and future work

The present work has described an assessment system integrated in a serious game for learning and training in electrotherapy techniques. The system mimics the expert's flexible evaluation by modeling each aspect of the electrode placement and current configuration through a fuzzy set representing the range of correct values. The overall score for the treatment is computed through a pessimistic exponential OWA operator, that penalizes a low score in any of the steps of the clinical practice. Our experiment have shown a good agreement of the system with a human expert, which confirms its value as a study and training tool for the students of physical therapy.

Future work will focus first on tuning the system to improve its agreement with the human expert even more, solving the shortcomings identified in Section 4.2. Although we showed that the system provides a similar assessment for a simulated clinical practice to that of a human expert, this does not allows us to conclude that it helps to improve the academic performance of the student, and as a result, his/her proficiency in the subject. Therefore, we have planned to assess this by conducting a more extensive experimentation, allowing a set of students to use the application throughout the electrotherapy course for studying and training, and comparing their academic performance with that of the students that prepared the course in the traditional way.

From a more technical point of view, all or part of the aspects involving the fuzzy evaluation of the variables and the aggregation rules to compute the overall score may be implemented as a fuzzy inference system using a fuzzy logic control framework. Many of these frameworks allow to encode the fuzzy term membership functions and rules in an external configuration file, which is a more flexible implementation and facilitates quick and easy adjustments of the system.

## CRediT authorship contribution statement

**Antonio J. Rueda:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review & editing, Supervision. **Carmen Martínez-Cruz:** Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Visualization, Writing – review & editing. **Ángeles Díaz-Fernández:** Validation, Formal analysis, Investigation, Resources,Writing – original draft. **María Catalina Osuna-Pérez:** Validation, Formal analysis, Investigation, Resources, Writing – original draft.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

We would like to thank professor Juan José Jimenez Delgado, from the Computer Science Department of the University of Jaén (Spain), for the ideas he contributed during the early stages of the development of this work. We would also like to thank the students of the first year of the Physiotherapy Degree of the 2020/2021 academic year at the University of Jaén (Spain) for their kindly participation in the evaluation of the app.

Finally, this project is partially supported by project B-TIC-744-UGR20 "ADIM: Accesibilidad de Datos para Investigación Médica of the Junta de Andalucía".

## References

Andayani, S., Hartati, S., Wardoyo, R., & Mardapi, D. (2017). Decision-making model for student assessment by unifying numerical and linguistic data. *International Journal of Electrical and Computer Engineering (IJECE)*, *1*, 363–373.

Annabestani, M., Rowhanimanesh, A., Mizani, A., & Rezaei, A. (2020). Fuzzy descriptive evaluation system: real, complete and fair evaluation of students. *Soft Computing*, *24*, 3025–3035.

Aziz, A., Golap, M. A., & Hashem, M. M. A. (2019). Student's academic performance evaluation method using fuzzy logic system. In *2019 1st international conference on advances in science, engineering and robotics technology (ICASERT)* (pp. 1–6).

Capuano, N., Loia, V., & Orciuoli, F. (2017). A fuzzy group decision making model for ordinal peer assessment. *IEEE Transactions on Learning Technologies*, *10*, 247–259.

Carlsson, C., & Fullér, R. (1997). Owa operators for decision support. In *Proceedings of EUFIT'97 conference* (pp. 8–11).

Chai, K., Tay, K. M., & Lim, C. (2015). A new fuzzy peer assessment methodology for cooperative learning of students. *Applied Soft Computing*, *32*.

Chang, T.-Y., & Chen, Y.-T. (2009). Cooperative learning in e-learning: A peer assessment of student-centered using consistent fuzzy preference. *Expert Systems with Applications*, *36*, 8342–8349.

Chuang, T.-Y., Zhi-Feng Liu, E., & Shiu, W.-Y. (2015). Game-based creativity assessment system: The application of fuzzy theory. *Multimedia Tools and Applications*, *74*, 9141–9155.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46.

Díaz-Fernández, A., Jiménez-Delgado, J. J., Osuna-Pérez, M. C., Rueda-Ruiz, A. J., & Paulano-Godino, F. (2016). Development and implementation of a mobile application to improve university teaching of electrotherapy. In *2016 international conference on interactive mobile communication, technologies and learning (IMCL) (pp. 33–37). San Diego (USA), 17-19 Octubre*. http://dx.doi.org/10.1109/IMCTL.2016.7753766.

Dubois, D. H. P. (2000). Fundamentals of fuzzy sets. In *The handbook of fuzzy sets series*, Kluwer Academic Publishers.

Echauz, J. R., & Vachtsevanos, G. J. (1995). Fuzzy grading system. *IEEE Transactions on Education*, *38*, 158–165. http://dx.doi.org/10.1109/13.387218.

Filev, D., & Yager, R. R. (1998). On the issue of obtaining owa operator weights. *Fuzzy Sets and Systems*, *94*, 157–169.

Jamsandekar, S. S., & Mudholkar, R. R. (2013). Performance evaluation by fuzzy inference technique. In *Proceedings of the international journal of soft computing and engineering (IJSCE), Vol. 3*.

Kapp, K. M. (2012). *The gamification of learning and instruction: game-based methods and strategies for training and education* (1st ed.). Pfeiffer & Company.

Kwok, R., Zhou, D., Zhang, Q., & Ma, J. (2007). A fuzzy multi-criteria decision making model for is student group project assessment. *Group Decision and Negotiation*, *16*, 25–42.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174.

Ma, J., & Zhou, D. (2000). Fuzzy set approach to the assessment of student-centered learning. *IEEE Transactions on Education*, *43*, 237–241.

Mamdani, E., & Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies*, *7*, 1–13.

Nguyen, J., Sánchez-Hernández, G., Agell, N., Armisen, A., Rovira, X., & Angulo, C. (2017). Evaluating student-internship fit using fuzzy linguistic terms and a fuzzy owa operator. In *2017 IEEE international conference on fuzzy systems (FUZZ-IEEE)* (pp. 1–5).

Pangaro, L. (2000). Investing in descriptive evaluation: A vision for the future of assessment. *Medical Teacher*, *22*, 478–481.

Rasmani, K., & Shen, Q. (2006). Data-driven fuzzy rule generation and its application for student academic performance evaluation. *Applied Intelligence*, *25*.

Samarakou, M., Prentakis, P., Mitsoudis, D., Karolidis, D., & Athinaios, S. (2017). Application of fuzzy logic for the assessment of engineering students. In *2017 IEEE global engineering education conference (EDUCON)* (pp. 646–650).

Shrout, P., & Fleiss, J. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, *86*, 420–428.

Takagi, T., & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-15*, 116–132.

Vince, J. (2017). *Mathematics for computer graphics* (5th ed.). Springer.

Yager, R. R. (1988). On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Transactions on Systems, Man, and Cybernetics*, *18*, 183–190.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, *8*, 338–353.

Zhou, D., Kwok, R., Zhang, Q., & Ma, J. (2001). A new method for student project assessment using fuzzy sets. In *Proceedings joint 9th IFSA world congress and 20th NAFIPS international conference* (pp. 586–591).