An Ontology for Clinical Decision Support System to Predict Female's Fertile Period

Francisco Vaz[1], Rodrigo Rocha Silva[2,3] and Jorge Bernardino[1,3]
[1]Polytechnic of Coimbra - ISEC, Rua Pedro Nunes, Coimbra, Portugal
[2]FATEC Mogi das Cruzes, São Paulo State Technological College, Brazil
[3]CISUC – Centre for Informatics and Systems of the University of Coimbra, Coimbra, Portugal

Keywords: Ontology-based Clinical Decision Support System, Domain Ontologies, Knowledge Engineering.

Abstract: Nowadays, many women do not fully realize what the fertile period is, as well as what it represents in their life. The prediction of fertile period is quite complex and difficult to calculate accurately and this is undoubtedly a problem for women's. There are still some completely wrong ideas about reproductive health, especially about the fertile period and the menstrual cycle. A good example of the myths that persist is the fact that many women continue to believe that ovulation occurs precisely in the middle of their menstrual cycle, which is not always true. Therefore, to better understand the female cycle, we proposed a Clinical Decision Support System based on the use of an ontology. Our proposal can predict the female fertile period, based on certain factors that allow a calculation that is more accurate improving the quality of patient life.

1 INTRODUCTION

At a certain point in a woman's life, the fertile period gains a lot of importance. However, it is common for many women to not fully realize what the fertile period is and what it represents for their life. It is easily defined as the ideal time for the woman to conceive, which has a durability of approximately 6 days and occurs 14 days before the last of the cycle (Wilcox, Dunson, and Baird, 2000).

Women have begun to have more access to information about their reproductive health, but there are still many questions about the fertile period and the menstrual cycle. Some of these questions are, the factors influencing the menstrual cycle, the best way to calculate it, among others. With this questions, it is of great relevance to be able to unveil all the myths and all the ideas formed based on misconceptions and to encourage the women to know better their body.

In this paper we propose a Clinical Decision Support (CDS) system based in the use of an ontology to overcome the problem of predicting accurately the female fertile period and giving important information to the user. This system can work partially in place of the doctors in the prediction of the fertile period. The CDS system allows the user to have the possibility of sharing data with the doctor. This CDS system has several purposes, such as establishing a diagnosis, with access to the frequency of symptoms, for diseases that cause influences on the menstrual cycle or in cases of planning a pregnancy. The CDS system we propose reduces the frequency of physical consultations and allows a better doctor-patient relationship, since the doctor always has at his/her disposal the data of the user, being able to do a enhanced follow-up. Is evident that the system permits access to the patient historical data and thus the comparison between cycles. The system also allows the woman to acquire a better knowledge about their body and improve her quality of life. Of course, is aimed for female genre, from puberty, in which it has the first menstrual episode (commonly between 11 and 16 years), until menarche and menopause (Seeley, Stephens, and Tate, 2003). And with regard to the prediction of the fertile period, our system has the capacity to calculate different durations that a cycle can have, as well as the probability of occurrence of each factor (obesity, stress, pregnancy, among others) that has influence in the fertile period.

Vaz F., Silva R. and Bernardino J.
An Ontology for Clinical Decision Support System to Predict FemaleâĂŹs Fertile Period.
DOI: 10.5220/0006519102880293
In *Proceedings of the 9th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (KEOD 2017)*, pages 288-293
ISBN: 978-989-758-272-1

The main objective of this work is the construction of a system that can work without the support of health professionals at an early stage and that predicts the stages of the menstrual cycle for a woman. It is important that they be guided by some factors. The factors are the first day of the next menstruation, the risk of pregnancy for each day of the cycle, the recording of the days in which there were changes in mood, weight, symptoms, the day of ovulation, but always have medical follow-up through data sharing that the system will allow. Is important to use ontology to support the sharing and reuse of formally represented knowledge, and it is useful to define the common vocabulary in which shared knowledge is represented.

The rest of the paper is organized as follows. Section 2 discusses the existing work on the ontology in general and related to Ontology Based Clinical Decision Support systems. Section 3 discusses the proposed approach. Section 4 discusses the proposed ontology-based system. Finally, section 5 presents the conclusions and future work.

2 BACKGROUND AND LITERATURE REVIEW

In this section, we review some works to obtain more information and knowledge how to create the most correct and adequate ontology for a CDS system.

Ontology is an information organization technique that has received special attention in the last years, mainly with respect to the formal representation of knowledge. Generally, they are created by specialists, having their structure based on the description of concepts and of the semantic relationships between them.

An ontology typically provides a vocabulary that describes a domain of interest and a specification of the meaning of terms used in the vocabulary (Gruber, 1993). Depending on the precision of this specification, the notion of ontology encompasses several data and conceptual models, including, sets of terms, classifications, thesauri, database schemas, or fully axiomatized theories (Euzenat and Shvaiko, 2007).

Otero-Cerdeira et al. (2015) present a literature review regarding articles on ontology matching published in the last decade. It serves the purpose of offering an up-to-date review of the field and showing its evolution trends. In this paper over 1600 papers have been sorted according to a classification framework defined by the authors. This framework helps in identifying the distribution of the load work in the last decade.

Giovannini et al. (2012) addressed the software part by proposing a product centric ontology in which concepts of products, processes and resources are associated to functions and sustainable manufacturing knowledge. In order to project a knowledge-based system by simulating a sustainable manufacturing expert, can automatically identify change opportunities and propose alternatives based on the existing production scenario.

For our work, the ontology has an enormous utility because in computing area, there is still no consensus on the semantics of the term "ontology". An ontology is also composed of a set of formal axioms that restrict the interpretation of concepts and relations, making a clear and unambiguous representation on the knowledge of the intended domain. With that we can build this system and create an ontology for our domain, which is female's fertile period.

The category we address in this paper is based on the domain ontology, cited by (Silva, Ferreira, and Vijaykumar, 2010) as being reusable ontologies in a certain domain providing vocabularies on the concepts and relationship of them within that domain, on the elementary theories, and principles that govern the domain and also on the activities involved in this area.

These days, in medical settings, doctors or other entities often make mistakes in several areas of health. These errors can easily be avoided considering the use of the CDS systems based on an ontology. The CDS system help doctors daily, increasing the quality of care provided to a patient. These systems support doctors in the decision-making process and suggest appropriate treatments. The ontology is best suited to encapsulate the concepts and relationship of terms associated with the medical domain. It is suitable for capturing medical knowledge in a formal way, allowing to share and reuse it whenever necessary.

Many studies have been done with CDS systems to help in several diseases. Manzoor, Balubaid, Usman, and Mueen (2015) proposed a system to predict high-risk pregnant woman. They proposed a framework based on a CDS system that can work in place of doctors to diagnose high-risk pregnant women and, if this happens, refer them to qualified doctors for the necessary treatment. In this way, high risk patients will receive the treatment at the correct time, and many lives can be saved. The main focus of the system that has been proposed is to build a diagnostic procedure that can work instead of qualified physicians and identify high-risk patients, once identified, can be treated by doctors. With this system it is possible to greatly reduce the workload of doctors, in addition to providing basic health care to more patients (Manzoor et al., 2015).

Highlighting also the authors in (Jostinah Lam, Abdullah, and Eko Supriyanto, 2015) that said there is a large shortage of doctors in rural areas and that is the main cause for maternal mortality to be quite high.

Then a solution was proposed introducing a new architecture of a CDS system in the field of high risk pregnancy. The proposed architecture is composed of seven main components. The need for CDS was investigated through the interview session, distribution of questionnaires and observation. The CDS architecture was categorized into 7 major components: knowledge base, inference engine, machine learning, case database, EMR, query engine and user interface (Jostinah Lam et al., 2015).

Raza, Chaundry, and Zaidi (2017) have done a study to accurately distinguish between finger tremors of Parkinson's disease using a tri-axial gyroscope. The study is an effort to provide physicians with a CDS system to facilitate them in accurate diagnosis of Parkinson's disease. They designed the hardware to acquire angular displacement from tri-axial gyroscope and apply a series of techniques to extract different features in time and frequency domains. A total of 104 people participated in their study, using resources from these data, they were able to create a CDS system with overall accuracy of 82.43%. They used the CDSS in a hospital with an accuracy of 77.78% (Raza, Chaundry, and Zaidi, 2017).

Semenov and Kopanitsa (2016) present a process of development and implementation of a decision support system for laboratory service patients. The system allows patients reading and understanding medical records in natural language. For the laboratory service the system allowed increasing the level of satisfaction of the patients and the number of patients who came back to the laboratory service for more detailed testing (Semenov and Kopanitsa, 2016).

Jabez Christopher, Khanna Nehemiah, and Kannan (2015) presents a CDS system to assist junior clinicians in the diagnosis of Allergic Rhinitis. In their study, they did intradermal skin tests were performed on patients who had plausible allergic symptoms. For their study 872 patients who had allergic symptoms were considered. The rule based classification approach and the clinical test results were used to develop and validate the CDSS (Jabez Christopher, Khanna Nehemiah, and Kannan, 2015).

Tams and Euliano (2015) share lessons learned from creating two respiratory CDS systems for ventilating patients in a critical care setting. They concluded that: when creating a CDS system you must seek input from trained clinicians who are willing and capable to make prompt and correct therapies; Clinical decisions are case sensitive; CDS system's may not have acess to all of the data required to make decisions, but sometimes simple modifications to the algorithms may dramatically improve performance; and that it is important to focus only on the values which prove to be relevant, because we have vast amounts of data available for a clinician to understand the overall scope of the patient (Tams and Euliano, 2015).

3 PROPOSED TECHNIQUE

For the construction of the CDS system we should take in consideration several factors: the day of the start of the last menstruation, the number of days in the cycle, number of days of menstruation, contraceptive methods, weight, calendar, pregnancy, mood and symptoms status, data sharing and notifications. With all these factors, the system meets what is necessary to achieve the resolution to our problem. We selected the main factors that stand out and have more influence the female's fertile period. In this section, it is explained the architecture, as well as the reason of each one of the chosen factors, based on the study realized. And since we understand that the question of the fertile period is so important for women, it must be addressed.

The proposed architecture for the CDS system is composed of three main components, which are inference engine (IE), knowledge base (KB) and machine learning (ML), as shown in Figure 1. The IE uses the knowledge on the system and the knowledge about the patient to draw conclusions regarding certain conditions. The IE controls what kind of actions need to be taken by the system. IE determines the reminders, alerts and conclusions to be displayed in the system. The knowledge represented by KB is used by IE and the KB may be built with the help of an automated process, like machine learning (ML) or field rules. In this automated process, knowledge is acquired from databases that have information about the users. The KB together with ML system will complete the whole decision-making process. The ML component will be done in the future work to improve our system.

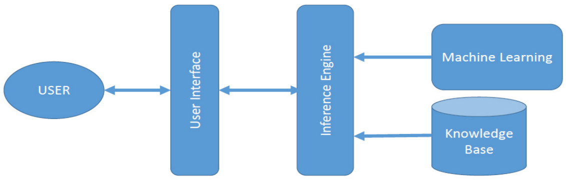

Figure 1: CDS system architecture.

Figure 2 shows the overall scheme for calculate female's fertile period. We can see in the middle of the figure, the central part of the system: the database and the server. It is there where all the communication takes place and all the information concerning the users and the doctors is stored.

From the central part of the scheme, there is the interaction of the two parts who can communicate in this system, user and doctors, represented to the left and to the right, respectively. On the left side, it was decided to put the image of a woman to be consulted by a doctor since, with our system a woman with a mobile phone can communicate with the doctor.

It should be noted that it is possible for a doctor to be aware of the condition of a patient without having to be physically with it, just in case of need.

Thus, Figure 2 aims to demonstrate in a simple but concrete way, the scheme for the proposed system.

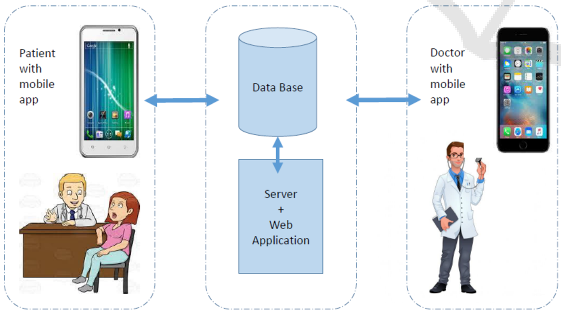

Figure 2: Overall scheme for calculate female's fertile period.

All the factors referenced above will now be explained and justified. These factors will be considered as input to the system and inserted by the user.

- Day of the start of the last menstruation: this corresponds to the first day of the cycle. It is of utmost relevance for the calculation of the fertile period. Over time, the system will be more suitable for the user.
- Number of days in the cycle: perceive if it is a short, long or normal cycle and allows to have the perception of the regularity of the cycles. Also having great relevance for the calculation of the fertile period, because at the last day of the cycle will be taken 14 days and will be obtained the probability of the day on which ovulation occurs. The mean duration of the cycle is 22 to 36 days (Fehring, Schneider, and Raviele, 2006).
- Number of days of menstruation: For woman to have more knowledge about her body. The average duration of menstruation is from 2 to 7 days (Seeley et al., 2003).
- Contraceptive method: in relation to the pill, the existing calendar allows the user to note whether the pill has been taken, whether it was late or note. It should be noted that it is possible to use any contraceptive method and that the system will be adapted.
- Calendar: this is a very important feature of the system, due to the fact that the ovulation day, the days of the fertile period, the risk of pregnancy (low, medium and high risk) for each day of the menstrual cycle, the days of menstruation inserted by the user and access to the history of the previous months, are all shown in the calendar.
- Weight: it is a factor that influences the oscillations of the fertile period. People with low weight do not have sufficient amounts of fat, with this, the cycles become increasingly irregular. Often menstruation may not even come. Usually weight gain helps reverse the previous scenario, which is low weight. This condition is quite common in athletes and in overworked women. It was found that body weight and dissatisfaction with the body varies during the menstrual cycle (Teixeira, Damasceno, Dias, Lamounier, and Gardner, 2013).
- Pregnancy: there is the possibility of applying the mode of pregnancy in the system, if the woman gets pregnant, serving as reference and keeping the record of each month of gestation. The pregnancy mode allows to see the countdown to the baby's birth and allows to receive a reminder to record the cycle after pregnancy. Taking into account the doctor's knowledge to make a more present follow-up to the patient.
- Mood and Symptoms Status: it allows the woman to make comparisons later between the phase of the cycle and a particular

symptom. Turns out to be her clinical history related to symptoms favouring the knowledge of her body. With regard to pre-menstrual tension it turns out to be a recurrent pattern of emotional, physical and behavioural changes in the days before menstruation (Seeley et al., 2003). When they perceive that they are the same with the cycles, it will cause the week before menstruation more easily. The mood and symptoms status can cause changes in the calculation of the fertile period and for the system it is of important relevance that one can adapt to these factors, so that the calculation of the fertile period is the most reliable possible.

- Data Sharing: this allows an exchange of data between the doctor and anyone the user wishes to have access to. That is, the user can always have the doctor's care without physical presence. And regarding the sharing between the couple, it is important that the man or woman also follow and that has knowledge of the body of his partner.
- Notifications: it will inform the user about taking the pill (if applicable) and predicts the day on which the next menstruation will occur. It promotes the fact of the daily intake at the same time, taking to the maximum efficiency of the same. In the case of another contraceptive, or even none at all, the system will adapt it and notify, for example, only the prediction of the day of the next menstruation.

4 PROPOSED ONTOLOGY-BASED SYSTEM

In this section, the ontology is discussed based on the factors identified and justified above. Through the above factors, an ontology capable of solving this problem of the calculation of the female's fertile period was implemented.

With regard to the proposed ontology for this domain, Figure 3, it consists of 6 main classes which are the symptoms, mood states, contraceptive methods, calendar, weight and pregnancy, the choice of these six classes comes from the study carried out in section 3.

In the symptoms status class, there are several subclasses like headaches and abdominal cramps, being able to have more according to the most relevant symptoms for the calculation.

In the mood status class, it happens as in the case of symptoms, where we have all the states that influence the calculation.

In the class of contraceptive methods are considered those contraceptives that currently exist.

The Calendar class has two subclasses, one of which allows the sharing of data between users of our system, as well as the subclass of notifications to enable better effectiveness, for example in taking a contraceptive method at a specific time, or warning about the next menstruation.

We also have the pregnancy classes that serves to indicate whether or not a woman is pregnant. This is a very important factor for the calculations.

Finally, we have the weight class, which according to some studies, (Teixeira et al., 2013) proved to be a necessary factor in our ontology, since the weight variations interfere in the calculation of the fertile period.

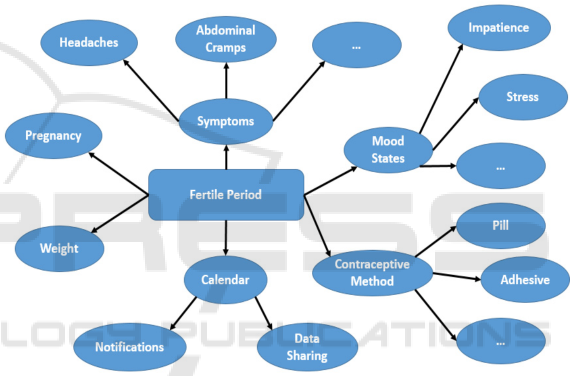

Figure 3: Snapshot of female's fertile period ontology.

In this section, a figure of great relevance was explained (Figure 3). Ontology has the advantage where the knowledge can be changed easily if knowledge of the domain has been outdated. By using ontological approach, the knowledge repository becomes more flexible to changes.

The proposed ontology provides the controlled vocabulary required for the annotation of our dataset with the woman and doctor's information, facilitating the retrieval of and, more generally, access to information. Such standardization also facilitates the exchange of information and contributes to semantic interoperability among system.

Our proposed ontology are also critical to hypothesis generation and knowledge discovery in a data-driven approach to predict the delay of the woman cycle for the next month.

5 CONCLUSIONS AND FUTURE WORK

This work describes the importance of ontology-based CDS systems in calculate female's fertile period. It also improves the quality of health care service helping the society. The fact that this system has the option of data sharing is an increased value, since it allows a doctor-patient communication, quite efficient. We must not forget that the woman only shares what she wants, so that there is privacy for her. It also allows to save and optimize resources, such as the time spent on visits and consultations, as well as optimize the diagnosis by the doctor, because it has a history of the patient in question. The ontological approach had been used as a foundation in the development of CDS systems.

As future work we intend to develop a CDS system based on the proposed architecture. This prototype will be evaluated for its capability and usability, and it is necessary to put into practice using machine learning algorithms. Thus, with the role of machine learning in the proposed CDS system architecture, it will be possible to calculate more accurately the female's fertile period, and also, for example, calculate the probability of delaying menstruation in a particular user based on the past data thereof. The possibility of using field rules in this CDS system may also be studied, since it could also be included here.

With all these implementations of the future work made, it will be possible to do one of the most important purposes, which are tests with women and doctors, to prove the effectiveness of our system.

REFERENCES

Euzenat, J. and Shvaiko, P., Ontology matching. Springer, 2007

Fehring, R. J., Schneider, M., and Raviele, K. (2006). Variability in the Phases of the Menstrual Cycle. *Journal of Obstetric, Gynecologic and Neonatal Nursing*, *35*(3), 376–384.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*(2), 199–220.

Jabez Christopher, J., Khanna Nehemiah, H., and Kannan, A. (2015). A clinical decision support system for diagnosis of Allergic Rhinitis based on intradermal skin tests. *Computers in Biology and Medicine*, *65*, 76–84.

Jostinah Lam, Abdullah, M. S., and Eko Supriyanto. (2015). Architecture for clinical decision support system (cdss) using high risk pregnancy ontology. *ARPN Journal of Engineering and Applied Sciences*, *10*, 9.

Manzoor, U., Balubaid, M. A., Usman, M., and Mueen, A. (2015). Ontology-Based Clinical Decision Support System for Predicting High-Risk Pregnant Woman. *International Journal of Advanced Computer Science and Applications*, *6*, 6.

Raza, M. A., Chaundry, Q., and Zaidi, S. M. T. (2017). Clinical Decision Support System For Parkinson's Disease and Related Movement Disorders, 1108–1112.

Seeley, R. R., Stephens, T. D., and Tate, P. (2003). *Anatomia e Fisiologia* (6ª Edição).

Semenov, I., and Kopanitsa, G. (2016). Development of a Clinical Decision Support System for the Patients of a Laboratory Service. *Studies in Health Technology and Informatics*, *228*, 90–94.

Silva, R., Ferreira, M. G., and Vijaykumar, N. (2010). An Object-Oriented Meta-Model as Ontology for Describing Domains and Problems for Planning Space Applications Planning. *SpaceOps 2010 Conference*, (August).

Tams, C. G., and Euliano, N. R. (2015). Creating clinical decision support systems for respiratory medicine. *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, *32669*, 5335–5338.

Teixeira, A. L. S., Damasceno, V. O., Dias, M. R. C., Lamounier, J. A., and Gardner, R. M. (2013). Association between Different Phases of Menstrual Cycle and Body Image Measures of Perceived Size, Ideal Size, and Body Dissatisfaction. *Perceptual and Motor Skills*, *117*(3), 892–902.

Wilcox, A. J., Dunson, D., and Baird, D. D. (2000). The timing of the "fertile window" in the menstrual cycle: day specific estimates from a prospective study. *Bmj*, *321*, 1259–1263.