

Artificial Intelligence in Personalized Healthcare Analysis for Women’s’ Menstrual Health Disorders

Łukasz Sosnowski
Systems Research Institute,
Polish Academy of Sciences
Newelska 6, 01-447 Warsaw, Poland
sosnowsl@ibspan.waw.pl

Joanna Żuławińska
OvuFriend Sp. z o.o.
Złota 61/100,
00-819 Warsaw, Poland
joanna.zulawinska@ovufriend.com

Soma Dutta
University of Warmia and Mazury
in Olsztyn
Słoneczna 54, 10-710 Olsztyn, Poland
soma.dutta@matman.uwm.edu.pl

Iwona Szymusik
Department of Obstetrics and Gynecology, Department of Obstetrics and Gynecology,
Medical University of Warsaw,
Zwirki and Wigury 61,
02-091 Warsaw, Poland
iwona.szymusik@gmail.com

Aleksandra Zyguła
Medical University of Warsaw,
Zwirki and Wigury 61,
02-091 Warsaw, Poland
am.zygula@gmail.com

Elżbieta Bambul-Mazurek
Uniwersyteckie Centrum Zdrowia
Kobiety i Noworodka WUM Sp. z o.o.
Plac Starynkiewicza 1/3;
02-015 Warsaw, Poland
ela.bambul@gmail.com

Abstract—The paper presents an AI-based model which depending on the input of a woman for a finite number of menstrual cycles helps in determining the possible ovulation dates as well as possibility of some health risks e.g., Premenstrual Syndrome, Luteal Phase Defect etc. The architecture of the model consists of three layers, namely analyzing and detecting the features from a single cycle, analyzing cycle level concepts based on the analyzed features, and analyzing the user’s health risks based on the cycle level concepts accumulated over a finitely many cycles.

I. INTRODUCTION

IN THE last decades, in parallel to the industrial and social progress, several healthcare paradigms have appeared in the scientific medical community. These paradigms are proposing changes in the way in which healthcare is deployed in our society. The image of Traditional Medicine, where physicians are artists that are isolated and taking decisions based only on their knowledge and experience, are changing to a new doctor always connected and with access to the last evidence existing in a globalized world. [1]

As examples of new paradigms of medical treatment author of [1] refers to Evidence-Based Medicine [2], Personalized Medicine [3] etc. The key issue in all of them is to create the protocols and guidelines for medical care by combining the knowledge from the existing literature of medicine, experience of the professionals, as well as the input parameters, habits, life style, and preferences of the individual patients. However, implantation of Evidence-Based medicine and personalized medicine together in a platform of healthcare have some challenges; one way it needs standardization of protocols at least with respect to the consensus of a group, and on the other hand, it needs to be case sensitive by considering treatments

Co-financed by the EU Smart Growth Operational Program 2014-2020 under the project "Developing innovative solutions in the domain of detection of frequent intimate and hormonal health disorders in women of procreative age based on artificial intelligence and machine learning - OvuFriend 2.0", POIR.01.01.01-00-0826/20.

out of the guidelines in the context of the patients who have different responses to the standard treatment.

The concern of this paper is related to what has been mentioned in the paradigmatic change in medical treatment, in particular in the context of infertility [4] which has become a civilization disease. According to statistics every fifth couple, trying to conceive (TTC), has a problem to achieve pregnancy in the first 12 months of efforts, and this tendency is increasing [5]. Moreover, the age of women trying for the first child statistically shifts towards 35, which increases a risk in pregnancy, including the birth of a child with defects.

This paper is in continuation of a series of papers [4], [6]–[8] that attempted to establish a platform, known as OvuFriend 1.0¹, for helping women in determining the possibility of conceiving and understanding the hidden risk of getting related health problems based on their data input. The platform of OvuFriend 1.0 is provided as a mobile app where an user can put the data related to her physical and mental states during a specific menstrual cycle, and the underlying algorithm of the app helps to get an analysis of the possibility of conceiving or not conceiving. As mentioned in [8], OvuFriend 1.0, the result of a R&D project finished in 2020, brought the company OvuFriend a big commercial success because of its underlying AI algorithm [7] dedicated to the prediction and confirmation of ovulation supporting the natural endeavour for family planning methods [9].

As a continuation to the above mentioned achievement the second R&D project, called as OvuFriend 2.0, is aimed at extending the previous platform by adding the ability of analyzing and assessing the risk of having certain health disorders based on the given input of a woman. Identifying the increased risk will give a chance to refer to the right doctor and heal the ailment faster. In particular, the project

¹www.ovufriend.pl

focuses on the analysis of whether a particular user has the possibility of having the risk of Premenstrual Syndrome (PMS²), Luteal Phase Defect (LPD³), benign growths like polyps, fibroids⁴ in the uterus, Polycystic Ovary Syndrome (PCOS⁵) or hypothyroidism⁶. PMS is a combination of symptoms that many women get about a week or two before their period. Severe PMS symptoms may be a sign of premenstrual dysphoric disorder (PMDD). On the other hand, LPD is a health condition that may play a role in infertility. Fibroids and polyps too may cause infertility or recurrent pregnancy loss. This paper focuses on the schemes for detecting PMS, LPD, and other anatomical changes like polyp and fibroids.

The general scheme in OvuFriend 2.0 for having an AI based app determining the possible days of ovulation as well as the possibility of the above mentioned health risks goes to a great extent in the line of Evidence-based Medicine and Personalized Medicine. In particular, the following features, that are included in the model proposed by OvuFriend 2.0, strengthen the support for a Personalized Medicine.

It has three hierarchical levels, known as *Detector level*, *Cycle level*, and *User level*.

- (i) At the *detector's level* the user can put information related to her mental and physical health over one complete cycle. A set of attributes are chosen by the medical experts. Based on the provided input by a particular user the values for those attributes are determined by a team of medical experts and they are tagged against the information details of the patient. So, while preprocessing the data, the model aggregates the perception of the user as well as the knowledge and experience of a team of experts.
- (ii) Based on the values of the attributes from a completed cycle, certain compound concepts such as *ovulation happened*, *days of ovulation*, *follicular phase interval*, *luteal phase interval*, *PMS score* etc are determined. These are called *cycle-level concepts* and for determining such concepts the system is fed with some relevant formulas involving the attributes prefixed at the detector level. These formulas are formulated by abstracting relationships among different attributes as described by a team of medical experts based on their knowledge from the literature and personal experiences. So, in the proposed model the mathematical formulations of the interrelationships among different attribute values are discovered by aggregating a team of medical experts' opinions.
- (iii) In the third level, the system aggregates the data related to

²<https://www.womenshealth.gov/menstrual-cycle/premenstrual-syndrome>

³<https://www.webmd.com/infertility-and-reproduction/guide/luteal-phase-defect>

⁴<https://progyny.com/education/female-infertility/understanding-uterine-fibroids-polyps/>

⁵In this condition the ovaries produce an abnormal amount of androgens, that are usually present in women in small amounts [10].

⁶Hypothyroidism means the thyroid gland does not produce enough thyroid hormones, which can lead to changes in the menstrual cycle. (<https://helloclue.com/articles/cycle-a-z/hypothyroidism-and-the-menstrual-cycle>)

the detector level as well as the cycle level concepts of a particular user for a finitely many cycles. This level focusing on the user's history is known as the *user's level*. The examples of the user level concepts are *risk of PMS*, *risk of LPD*, *risk of infertility* etc. Here, the system calculates the probabilistic ratio of the above mentioned cycle level concepts over the total number of cycles considered for a particular user. Moreover, the system is also fed with a threshold value for each such user level concepts and these threshold values are learned or even adjusted based on the opinions of the medical experts and the histories of already recorded and analysed cases. If the respective ratio for a particular user level concept is greater than the prefixed threshold for that concept the system notifies the user about the possibility of such health risk. So, at this level the threshold, chosen for a particular health risk, is set based on both experts' knowledge and current existing evidences of such cases.

The above discussed general scheme is presented in Fig. 1. The process of determining ovulation was described in the previous publication as part of the scope of the previous project (Ovufriend 1.0) [7].

Thus, as a whole the model endorses a three-layered hierarchical learning and reasoning mechanism based on the knowledge and experiences of a team of medical experts, perceptions of the users, and already recorded evidences to the system. Furthermore, the hierarchy of approximating fuzzy concepts is developed by using the quantifiers of fuzzy linguistic summaries in the process of inferring and making local decisions [11], [12]. From this angle, the model designed in OvuFriend 2.0 complies to a great extent to the need of personalized and evidence based medicine. On the other hand, the model also endorses some features of Interactive Granular Computing (IGrC) [13], [14], by incorporating perception of the current health situation of a woman based on the individual spatio-temporal windows of the physical world and actual physical interactions in the form of measuring attributes in the given space and time windows.

The content of the chapter is organized as follows. Section II presents the development made under OvuFriend 2.0; it is divided into several subsections describing the schemes for determining PMS, analysing the risk of LPD, and indicating anatomical changes related to polyp, fibroids etc. Further in section III, the reference set, used for experiments, is described, and the obtained results are explained in section IV. The paper ends with a concluding section indicating future directions of developing the model.

II. AI ALGORITHMS DETERMINING WOMEN HEALTHCARE RISKS

In this section we would present the framework of OvuFriend 2.0 by describing the AI algorithms and schemes for determining whether an user has the risks of certain health diseases. Specifically, we focus on the health diseases such as PMS, LPD, Fibroids and Polyps. All these schemes are discussed below in separate subsections.

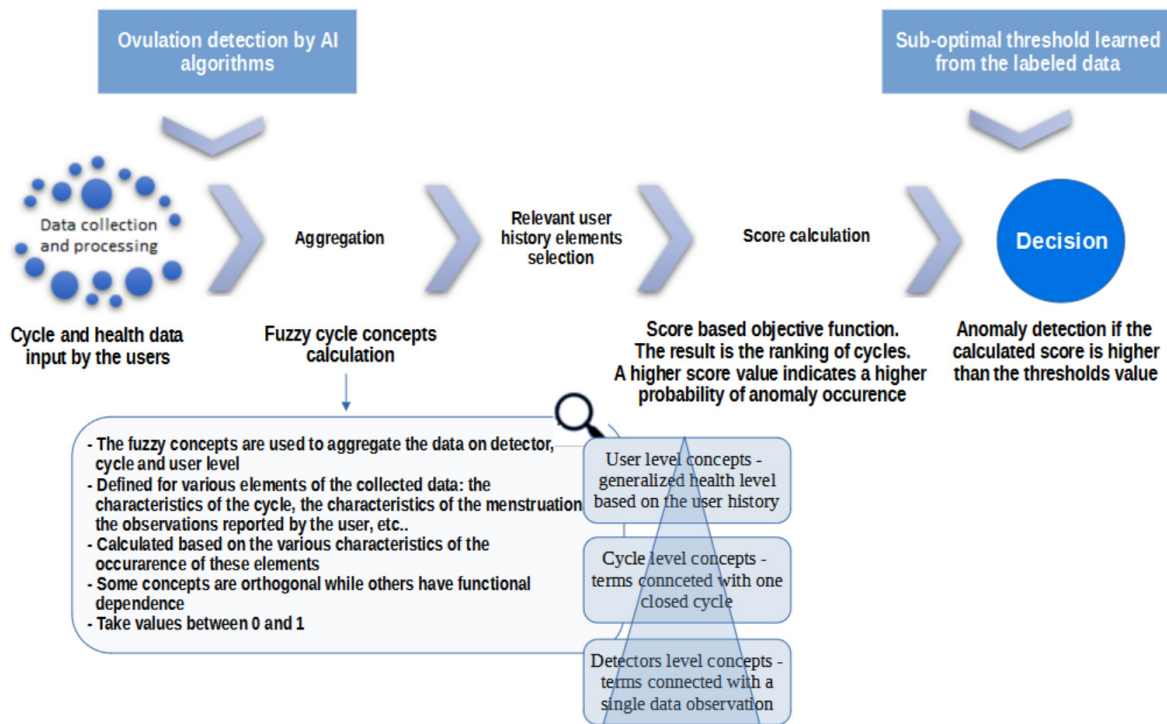


Fig. 1. General scheme for OvuFriend 2.0

A. Scheme to determine risk for PMS

The prerequisite to start this scheme is to collect data related to the physical and mental health of a woman before, during, and after a complete menstrual cycle. After the completion of a cycle, with the gathered data, analysis for the risk of PMS starts. Initially, the data is processed to investigate whether the ovulation has occurred and whether it is possible to determine the day of its occurrence. At this stage all concepts pertaining to the detector level are analysed and determined. For example, if ovulation has been determined, an attempt is made to indicate two intervals of equal length falling into the follicular phase and the luteal phase of the cycle respectively.

The length of the intervals depends on the length of menstruation, the day of ovulation, and the length of the total cycle. A complete cycle means number of days between starting of the menstruation in one month to the starting of the same in the next month. The beginning point of the first interval is chosen as the k -th day after the end of the monthly menstruation of the current month, where the value for k is prefixed in the algorithm. If the length of cycle is x and number of days of the current month menstruation is m , then each interval has to be of length $\frac{x-(k+m)}{2}$ and hence the beginning point and the end point of the first interval are respectively $m+k$ and $\frac{x+(k+m)}{2}$. Consequently, the beginning point of the second interval is $\frac{x+(k+m)}{2}$ and the the end point is x , the last day of the cycle.

Now, if the intervals are successfully determined, the coefficients of occurrence of the physical symptoms and mood

symptoms characteristics of PMS are calculated. The set of mood symptoms is presented in the Fig. 5. This set of symptoms and formulas for calculating the coefficients based on them are defined based on the interactions with a team of medical experts and aggregating their consensus of gathered knowledge and experiences about variations of different moods, feelings and physical impacts observed in women during the menstrual cycles. Each such symptom related to physical or mood aspects is counted and it is checked whether they occur in both the phases or only in the second phase. If there is at least one physical symptom or mood symptom that occurs in both phases, the algorithm reduces the weights in the respective formula calculating the mood feel coefficient or physical feel coefficient. Usually the physical or mood impacts during the second phase are only reflected in the occurrences of PMS, and that is why, while some symptoms are observed in both the phases, the possibility for PMS is decreased by reducing the weights. Finally, by aggregating the number of physical symptoms and the mood symptoms in a particular phase the coefficients are calculated separately for the physical symptoms and the mood symptoms according to the following formulas. Fig. 2 shows the algorithmic flowchart behind the described process for determining the cycle level concept *PMS score*, denoted as PMS_{score} .

Let us denote the two phases as P_1 and P_2 respectively.

$$P_i MoodFeelCoeff = \frac{(SumOfOccurrenceP_iMood)}{K_1 \times PhaseLength} \times \alpha + (1-\alpha) \tag{1}$$

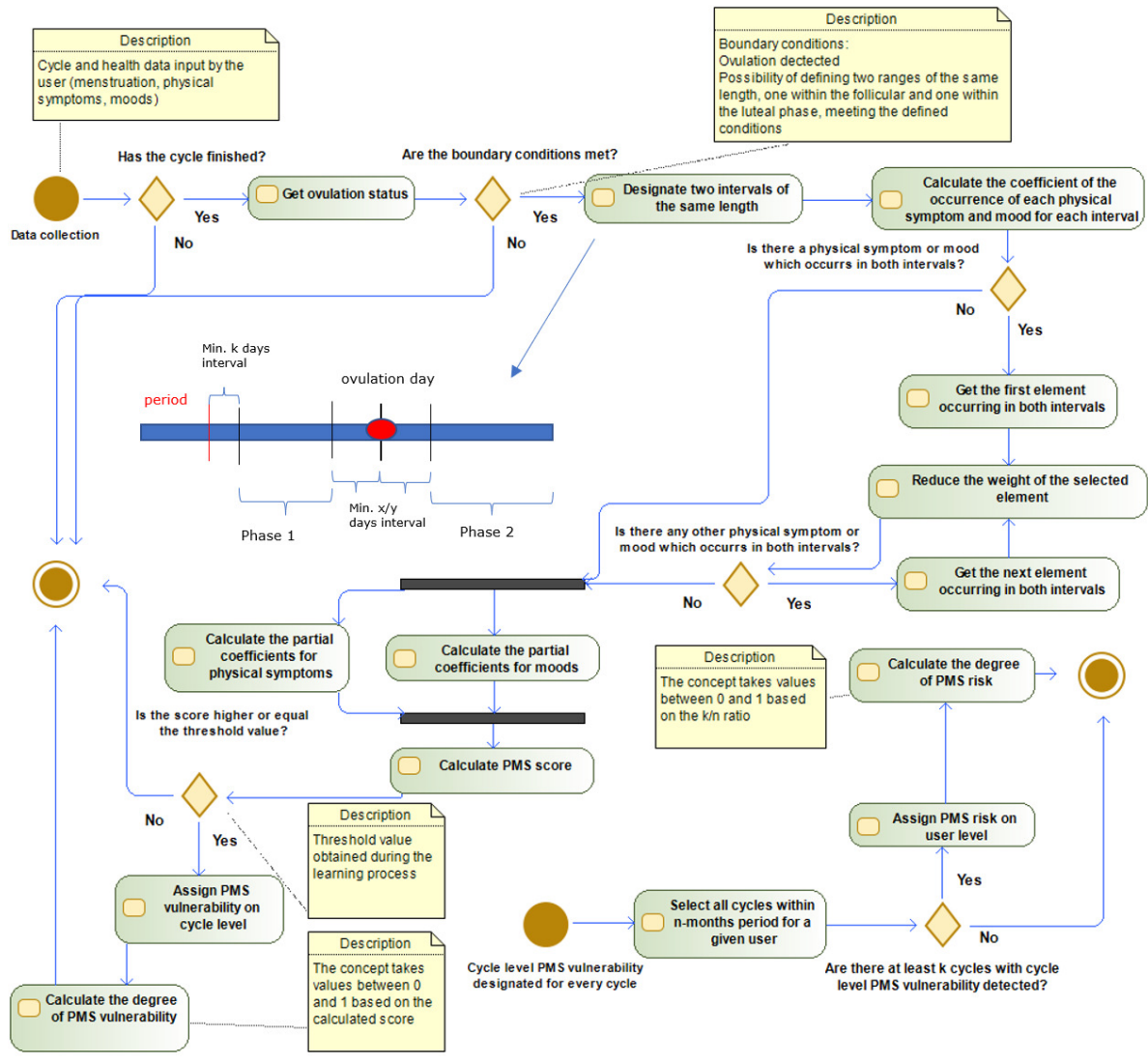


Fig. 2. Complete scheme for calculating PMS risk vulnerability (2 diagrams)

where $i = 1, 2$ and $\alpha \in (0, 1)$,

$$P_2PhysFeelCoeff = \frac{(SumOfOccurrenceP_2Phys)}{K_2 \times PhaseLength} \times \beta + (1 - \beta) \quad (2)$$

where $\beta \in (0, 1)$.

The symbols $SumOfOccurrenceP_iMood$ and $SumOfOccurrenceP_iPhys$, used in the above equations, respectively indicate the number of mood and the number of physical symptoms occurred in a particular phase P_i . The symbols K_1 and K_2 represent respectively the total number of all moods and physical symptoms listed in the system. The factors α and β are parameters to control the significance of the given components in the final calculation of the result. The formulas are designed by a team of scientific experts based on the general description given by the medical experts regarding the effect on physical and mental health of women during a cycle as well as keeping into account the observed

patterns of cases available in the record. Based on the above coefficients the cycle-level concept, namely $PMS\ score$ is calculated in accordance to the following formula.

$$PMS_{score} = \frac{P_2MoodFeelCoeff}{P_1MoodFeelCoeff} + \frac{P_2PhysFeelCoeff}{w_1} \quad (3)$$

where w_1 is the weight chosen by a team of medical experts.

At the beginning, the algorithm also fixes a threshold for PMS score by taking into account already available records of patients. The threshold may vary over time based on the changes in the patients' record, and thus, in some sense the underlying algorithm keeps a possibility of learning the threshold for PMS score based on the current evidences. Based on this threshold whether an user has the PMS susceptibility or not is determined just by checking if the obtained score is greater or equal to the prefixed threshold. If during the current

cycle the algorithm determines PMS susceptibility for an user the algorithm passes to the next level where the degree of PMS risk is calculated for a particular user based on the observations of finitely many cycles.

The cycle level compliance data is used to investigate PMS risk at the user level. If over the selected period of n months there are at least k cycles with PMS susceptibility, then the PMS risk is assigned to the user, and its degree is calculated simply by the value of $\frac{k}{n}$ in the range $[0, 1]$. The flowchart, presented in the Figure 2, shows a complete overview of the algorithm specifying PMS susceptibility and PMS risk for a particular user.

B. Scheme to determine risk for LPD

The general prerequisite for running the algorithm to determine the risk of Luteal Phase Defect (LPD) [15] is common to all considered disorders but differs in details. At the beginning stage, preprocessing of the input data and analyzing the detector level concepts such as *whether ovulation has occurred* are performed, and then based on that the boundary conditions are calculated. These conditions are verified using fuzzy quantifiers of linguistic summaries operating on multivariate time series (e.g., the quantifier *exists*) [16]. The specific scheme of LPD differs from that of PMS in the formula that is fed to the algorithm in order to calculate the susceptibility of LPD and then consequently its degree of risk.

Similar like, PMS score, here the AI algorithm is fed with a formula for calculating LPD score, given by the following equation.

$$LPD_{score} = w_1 * LutParameters + w_2 * DecFer \quad (4)$$

Here both $LutParameters, DecFer \in [0, 1]$ and they respectively denote the values for *Luteal Phase Parameters* and *Decreased Fertility*. The *Luteal Phase Parameters* are determined based on the luteal phase length and various other factors related to the analysis of bleeding during the luteal phase. On the other hand, the *Decreased Fertility* concerns about the observation of the period of time in which the attempts are made for conceiving a child, the number of miscarriages (they are estimated on the basis of pregnancy tests performed and the length of the cycle compared to the typical lengths of the cycle and luteal phase for a given user), etc. Both the respective values for *LutParameters, DecFer* are obtained based on the input data of the particular user; whereas w_1, w_2 are some weights that are chosen by the team of experts based on their collective knowledge regarding which of *LutParameters* and *DecFer* are significant to what extent.

Once in the cycle level the algorithm determines the possibility of LPD, it passes to the next level and as in the case of PMS risk the algorithm calculates the degree of risk for LPD; that is, if in the selected period of n months there are at least k cycles with LPD susceptibility, then the LPD risk is simply the value of $\frac{k}{n}$. For an overview of the whole scheme the readers are referred to Fig. 3.

C. Scheme for indicating anatomical changes like polyps and fibroids

As in the cases for PMS and LPD, here also the analysis for the presence of anatomical anomalies starts with the data of a complete cycle of an user. The primary analysis is manifested by focusing on the data related to inter-menstrual bleedings or spots. As usual, initially, the data is processed to investigate whether the ovulation has occurred and whether it is possible to designate a possible ovulation date. At the same time, the detector level concepts are also determined and then cycle label concepts are analyzed in the same fashion as mentioned in the cases of PMS and LPD.

If the required data is obtained during a complete cycle so that the algorithm becomes able to determine the occurrence of ovulation or an-ovulation, the process proceeds to the next stage of the examination of the disease. The cycle level concepts, which are associated to the symptoms characterizing the particular diseases like polyp or fibroids, are selected. On their basis, a score is calculated in accordance to the formula presented below.

$$Score = w_1 * DisMens + w_2 * DecFer + w_3 * PhysSymp \quad (5)$$

Here all the weights w_1, w_2, w_3 are chosen by the team of experts, and *DisMens*, the value for the parameters corresponding to disordered menstruation, *DecFer*, the value for decreased fertility, and *PhysSymp*, the values corresponding to physical symptoms related to such diseases, are obtained from the input data of a particular user. All these values are scaled in the interval $[0, 1]$ based on the information related to inter-menstrual bleeding, long-lasting menstruation, intensity of menstruation, miscarriage, long trying time for conceiving, pelvis pain, polyuria etc.

As before, in this context also if a cycle's score is greater than or equal to the cut-off value, which is set through some learning process, the cycle is assigned anatomical susceptibility at that particular cycle level. Then its grade is calculated in the range of $[0, 1]$ depending on out of n cycles in how many cycles the algorithm agrees with the susceptibility of anatomical changes occurred in the case of a particular user; in other words, it is simply $\frac{k}{n}$ if in k such cycles susceptibility of anatomical changes is detected.

The full scheme for determining the possibility of such diseases as polyp, fibroids, can be visualized in Fig. 4.

III. REFERENCE SET BASED ON EXPERTS KNOWLEDGE

In order to estimate the effectiveness of the algorithms detecting anomalies, a reference set has been created consisting of cycles described by the experts. For each cycle the experts are provided the information e.g., the chance of an anomaly occurrence expressed by a value in the range $[0, 1]$ and a comment explaining the assessment made. They are also provided with a cycle visualization containing the basic data needed to determine the ovulation, as well as the information of already predicted ovulation (product of the Ovufriend 1.0 project). In addition, the visualization contains

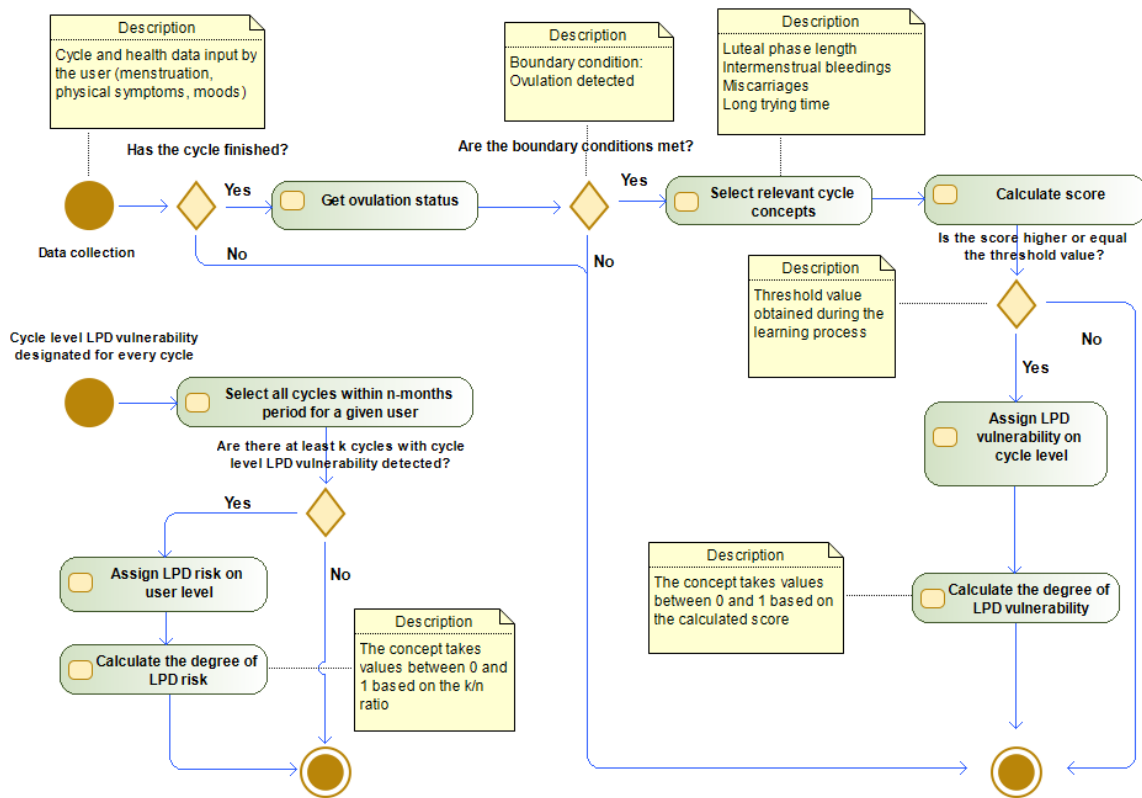


Fig. 3. Complete scheme for LPD risk vulnerability (2 diagrams)

a series of low level data (e.g., group of symptoms) broken down by observation days. As a whole it can be considered as a multidimensional time series indexed with the days of the ovulatory cycle. An example of a visualization labelling form for *PMS* can be found in Fig. 5 [17], and the same for luteal phase deficiency (*LPD*) can be found in Fig. 6. The form for tagging polyps and fibroids are similar to the one for *LPD*. It is extended with a few additional attributes and it is shown in Fig. 6. The reference set consists of 900 menstrual cycles, and in the category of "Anatomical or hormonal abnormalities with inter-menstrual spots" each cycle is tagged for both the presence of *NFL* and anatomical changes. The number of items in the set has grown steadily as successive sets of cycles are submitted for tagging. Subsequent sets of cycles are drawn

TABLE I
THE SAMPLE SIZE BY CLASS AND TYPE OF ANOMALY. PC - POSITIVE CLASS, NC - NEGATIVE CLASS

Anomaly	#	PC	% PC	NC	% NC
PMS	300	164	55%	136	45%
LPD	300	147	49%	153	51%
Polyps and Fibroids	300	160	53%	140	47%

on the basis of the given criteria, selected in such a way that it helps to obtain a similar size in the positive class (minimum score 0.5) and the negative class (score below 0.5). The size

of each class, broken down by a group and the types of anomaly, is presented in Table I. In each of the three groups of anomalies, the sample is well-balanced, where the share of the positive class ranges from 49% to 55% of the sample.

IV. EXPERIMENTS AND RESULTS

In order to evaluate the effectiveness of the prototypes of the algorithms, four experiments have been conducted for each of the three disorders. Overall the methodology looks as follows. Each of the experiments involves a different number of repetitions such as 1000, 500, 100, and 10, respectively for ReSample evaluation [18]. Each time 33% of cases from the reference set are drawn. Training is performed on the selected subset, and testing is performed on the remaining 67%. On each iteration, the intersection of both sets remains empty. The values of the cut-off thresholds of rankings for all described disorders are learned in each iteration from training set that consists of 100 cycles (33% of 300 tagged cycles). The test procedure that returned the values for contingency table are processed on 200 cycles each time. Single iteration results are stored in the contingency table. Finally, all TP (True Positive), TN (True Negative), FP (False Positive) and FN (False Negative) are summed and then measures of effectiveness are calculated. The obtained results for 1000 repetitions are presented in the Table II.

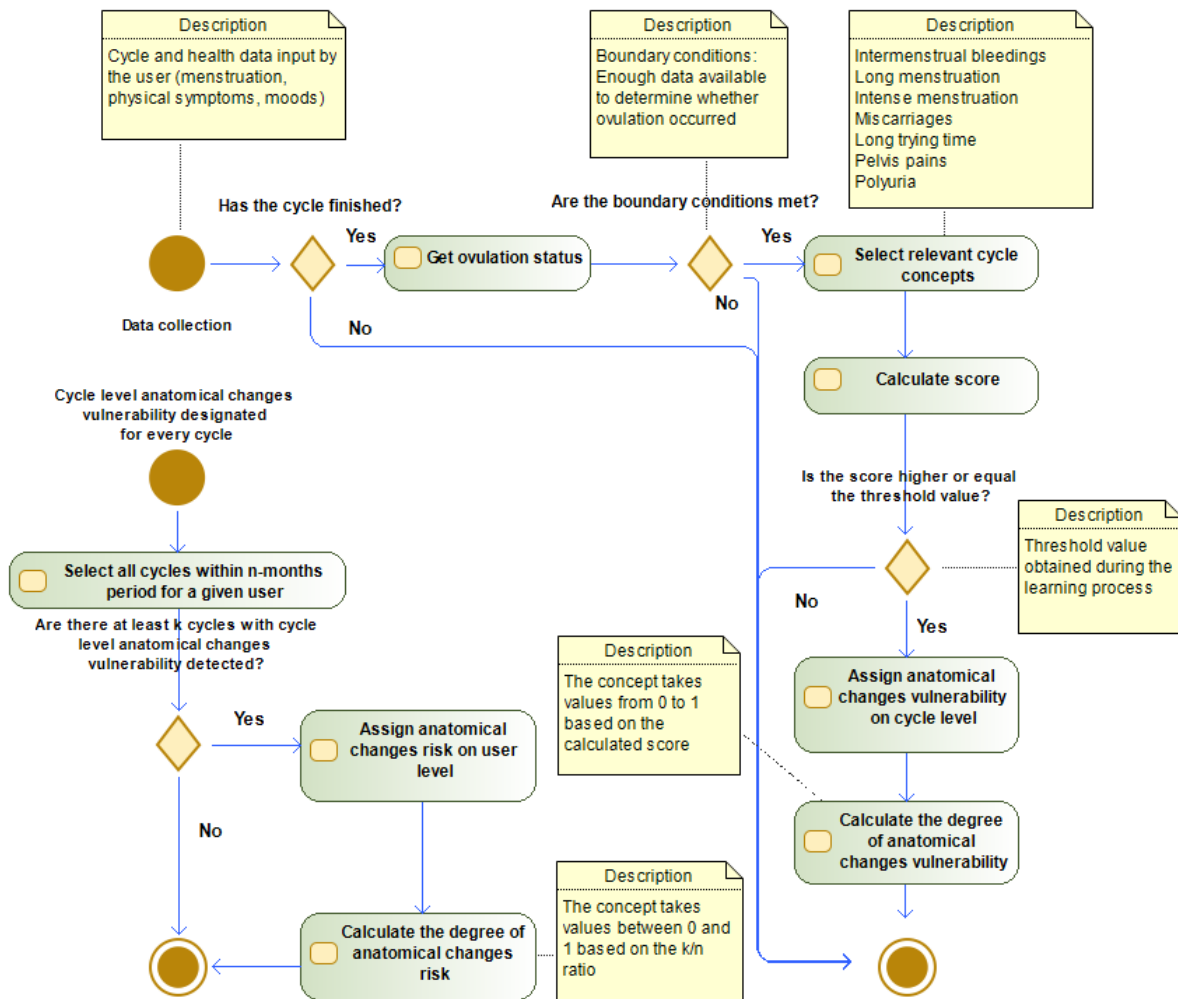


Fig. 4. Complete scheme for determining risk vulnerability of polyps and fibroids (2 diagrams)

The case for TP is assigned when the cycle is tagged with at least 0.5 by the medical experts and the algorithm has calculated the score that positioning the cycle in the positive set of a given disease; on the other hand, the case for FP is assigned when the experts have given mark below 0.5 but the algorithm has classified the case into positive class. The case for TN is obtained when the experts have assigned less than 0.5 and the algorithm has calculated score under the learned threshold. Finally the cases for FN is indicated when the algorithm has calculated the score value under the learned threshold but from the experts it receives a mark greater or equal 0.5. The evaluation results for 500 repetitions are presented in the Table III.

The obtained results for 100 and 10 repetitions are presented in the Table IV and Table V respectively.

In the project, the team of medical experts consists of three

highly qualified medical scientists, and the decisions regarding tagging cycles have been made based on discussion within the team. As the values, selected after tagging, are already agreed with the consensus of the whole team they do not require additional processing in order to be used for evaluation.

The presented results of the experiments concern the first stage of the project, in which the parameters of the algorithms are learned from the tagging of medical experts. The labels were made on the data constituting the subjective observations of the system users and the measurements of the physical parameters (e.g., BBT, cervical mucus, cervix position, etc.) and the subjective labels of the team of experts. In the next stage, the data and the final evaluation will be extended on the real test results, that are used by the doctors to make the diagnosis. In this way, the algorithm will be tested against the real hardcore medical data.

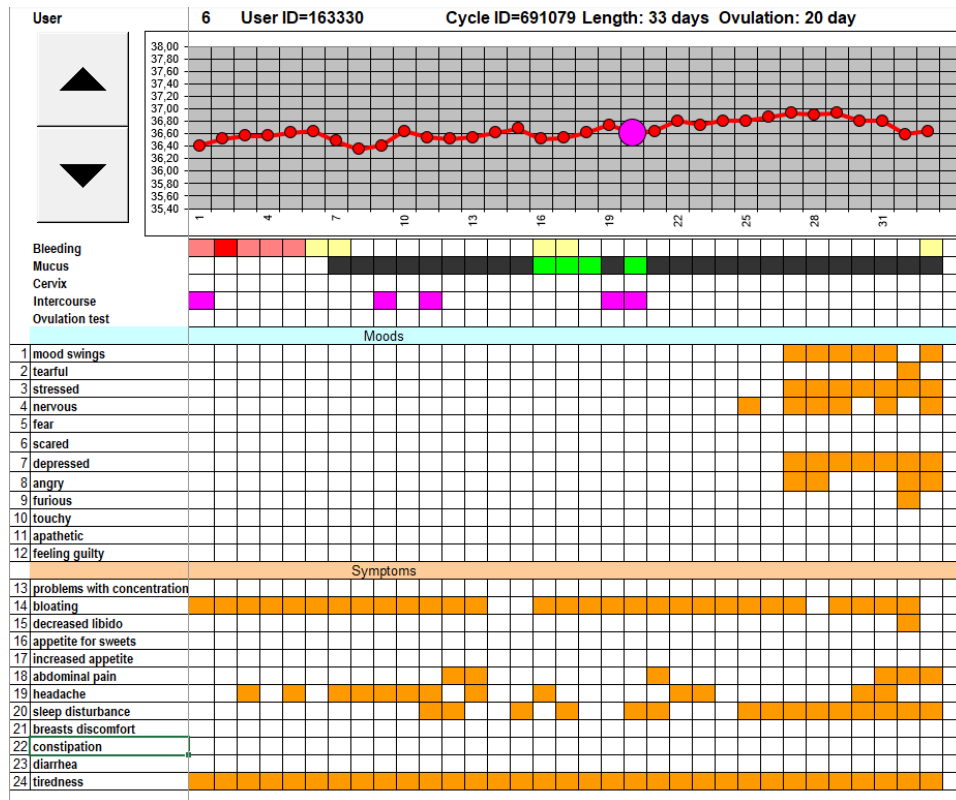


Fig. 5. PMS labelling form prepared for medical experts to evaluate susceptibility of selected cycles. Form is based on such attributes as: bleeding, mucus, bbt, cervix, mood swings, tearful, stressed, nervous, depressed, angry, furious, tiredness, bloating, problem with the concentration, appetite for sweets, sleep disturbance, breasts pains, constipation, etc.

TABLE II

RESULTS AVERAGED OVER 1000 ITERATIONS OF THE ReSAMPLE ROUTINE. ABBREVIATIONS: # - SAMPLE, TP - TRUE POSITIVES, TN - TRUE NEGATIVES, FP - FALSE POSITIVES, FN - FALSE NEGATIVES, PR - PRECISION, RE - RECALL, F1 - F1 SCORE, ACC - ACCURACY, POL-FIBR - POLYPS AND FIBROIDS

Type	#	TP	TN	FP	FN	PR	RE	F1	ACC	min_F1	max_F1	min_ACC	max_ACC
LPD	200000	88650	87920	13991	9439	0.864	0.904	0.883	0.883	0.781	0.925	0.795	0.920
PMS	200000	96685	70710	19921	12684	0.829	0.884	0.856	0.837	0.715	0.900	0.725	0.885
POL-FIBR	200000	86714	82157	10990	20139	0.888	0.812	0.848	0.844	0.731	0.900	0.760	0.895

V. CONCLUSION

AI and machine learning based techniques are nowadays prevalent in every sphere of life and the healthcare industry is also not out of that influence of automated decision support. In the Introduction, the terms like personalized medicine and evidence based medicine are presented in the context of the need for a new paradigm of medical treatment. It is expected that in the new context, treating a patient should not be an isolated process conducted by an individual doctor based on his/her knowledge and experience about a particular field of medicine. Moreover, there should be a standardization in the process of treating a particular disease by different doctors.

In this regard, the attempt of OvuFriend 2.0 has been to develop an AI-based model for women health care where based on the input of a particular user the model can suggest the possibility of certain health risks. The architecture of the model is developed in such a way that the system has an

interface of user in order to gather input data as well as an interface of a team of medical experts who based on a consensus creates a protocol for standardizing lowest level concepts, known as detector level concepts, and determining their values. Based on a complete cycle of data and values selected for detector level concepts the next level concepts, known as cycle level concepts, are analyzed and evaluated. Finally, at the highest level, known as user level, the degree of risk for certain cycle level concepts are computed by considering the values obtained for those concepts for a finitely many cycles.

Firstly, the user interface of the model keeps it sensitive to the user's perceptions and thus incorporates the aim of personalized medicine. Secondly, the interface for a team of medical experts keeps the possibility open for discussion, standardization, and revision of the defining criterion for medical concepts, and at the same time the process of treatment

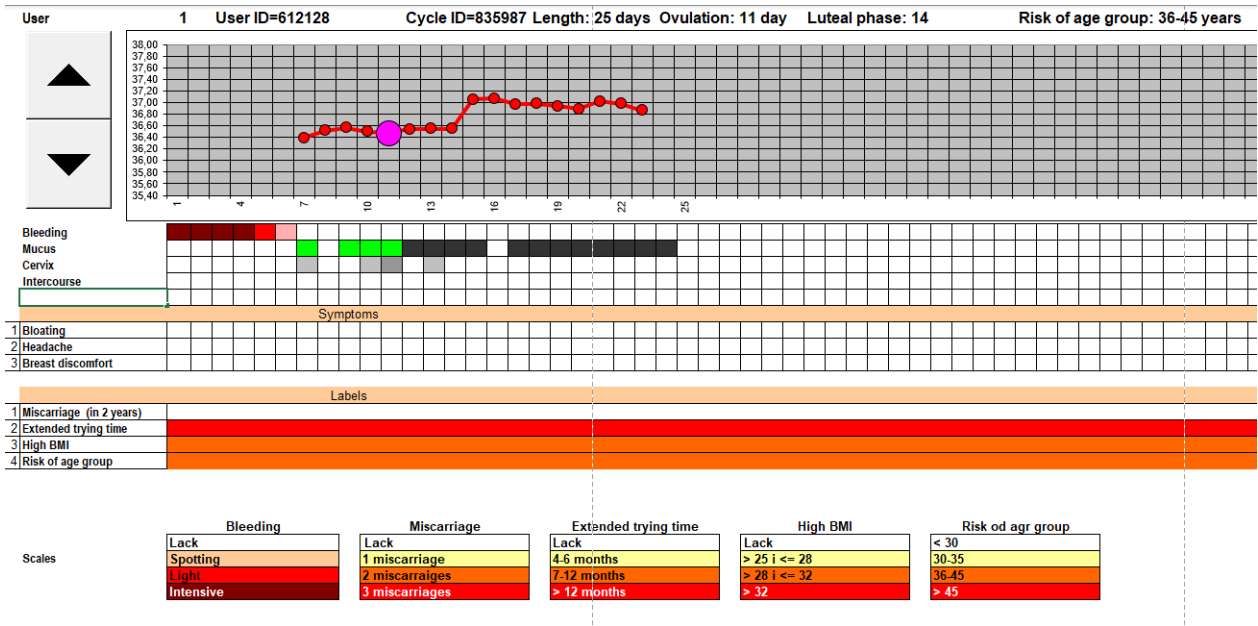


Fig. 6. Luteal phase deficiency and polyps tagging form prepared for medical experts to evaluate susceptibility of selected cycles. Form is based on such attributes as: bleeding, mucus, bbt, cervix, age group, increased BMI, extended trying time, miscarriages in history, etc.

TABLE III

RESULTS AVERAGED OVER 500 ITERATIONS OF THE ReSAMPLE ROUTINE. ABBREVIATIONS: # - SAMPLE, TP - TRUE POSITIVES, TN - TRUE NEGATIVES, FP - FALSE POSITIVES, FN - FALSE NEGATIVES, PR - PRECISION, RE - RECALL, F1 - F1 SCORE, ACC - ACCURACY, POL-FIBR - POLYPS AND FIBROIDS

Type	#	TP	TN	FP	FN	PR	RE	F1	ACC	min_F1	max_F1	min_ACC	max_ACC
LPD	100000	44317	44071	7035	4577	0.863	0.906	0.884	0.884	0.789	0.921	0.805	0.915
PMS	100000	48043	35454	9965	6538	0.828	0.880	0.853	0.835	0.754	0.902	0.745	0.885
POL-FIBR	100000	43523	41078	5618	9781	0.886	0.817	0.850	0.846	0.747	0.910	0.770	0.905

TABLE IV

RESULTS AVERAGED OVER 100 ITERATIONS OF THE ReSAMPLE ROUTINE. ABBREVIATIONS: # - SAMPLE, TP - TRUE POSITIVES, TN - TRUE NEGATIVES, FP - FALSE POSITIVES, FN - FALSE NEGATIVES, PR - PRECISION, RE - RECALL, F1 - F1 SCORE, ACC - ACCURACY, POL-FIBR - POLYPS AND FIBROIDS

Type	#	TP	TN	FP	FN	PR	RE	F1	ACC	min_F1	max_F1	min_ACC	max_ACC
LPD	20000	8808	8833	1394	965	0.863	0.901	0.882	0.882	0.783	0.921	0.795	0.915
PMS	20000	9687	7025	2049	1239	0.825	0.887	0.855	0.836	0.760	0.911	0.760	0.895
POL-FIBR	20000	8672	8245	1122	1961	0.885	0.816	0.849	0.846	0.751	0.900	0.775	0.890

does not remain in the hand of one individual. Thirdly, the underlying AI algorithm has a updating mechanism which with time changes certain thresholds for analyzing health risks based on the already available records of the patients. Thus, the model incorporates a learning mechanism as well as supports the idea of evidence based medicine.

In the present version of the model, there are some aspects where lie the possibility of extension and improvement. Let us list them as immediate directions for future research.

- One of them is to make the two above mentioned interfaces interactive by introducing a language of dialogue [19], [20] so that the underlying treatment protocol can be updated or revised based on even a particular user’s input. In the present model, certain weights for a particular health disease does not depend on the input of a

particular user. Incorporating this direction may make the model more dynamic in learning the optimized care for a particular user.

- In the present context, the formulas determining the values for the cycle level concepts and the user level concepts are fixed. In this context also there is a possibility of using machine learning techniques in order to learn a set of possible rules or formulas for diagnosing certain health risks based on the already existing evidences and making the process of diagnosing more flexible and evidence driven.

The current state of the algorithms shows a very good quality of the results achieved by the algorithms, while the real test will be their modernization and incorporation of data from medical examinations into operation. It will be a

TABLE V

RESULTS AVERAGED OVER 100 ITERATIONS OF THE RESAMPLE ROUTINE. ABBREVIATIONS: # - SAMPLE, TP - TRUE POSITIVES, TN - TRUE NEGATIVES, FP - FALSE POSITIVES, FN - FALSE NEGATIVES, PR - PRECISION, RE - RECALL, F1 - F1 SCORE, ACC - ACCURACY, POL-FIBR - POLYPS AND FIBROIDS

Type	#	TP	TN	FP	FN	PR	RE	F1	ACC	min_F1	max_F1	min_ACC	max_ACC
LPD	2000	902	873	120	105	0.883	0.896	0.889	0.888	0.845	0.924	0.855	0.920
PMS	2000	950	709	209	132	0.820	0.878	0.848	0.830	0.831	0.869	0.815	0.845
POL-FIBR	2000	868	824	115	193	0.883	0.818	0.849	0.846	0.791	0.877	0.810	0.875

milestone which, if achieved, will guarantee another success for applications and system users.

REFERENCES

- [1] Carlos Fernández-Llatas and Jorge Muñoz-Gama *et al.* "Process Mining in Healthcare", pages 41–52. Springer, 2020.
- [2] Brian Haynes and Andrew Haines. "Barriers and bridges to evidence based clinical practice". *BMJ*, 317(7153):273–276, 1998.
- [3] Margaret A. Hamburg and Francis S. Collins. "The Path to Personalized Medicine". *New England Journal of Medicine*, 363(4):301–304, 2010.
- [4] Lukasz Sosnowski and Tomasz Penza. "Generating Fuzzy Linguistic Summaries for Menstrual Cycles". volume 21 of *Annals of Computer Science and Information Systems*, pages 119–128, 2020.
- [5] L Bablok, W Dziadecki, I Szymusik, and *et al.* "Patterns of infertility in Poland - multicenter study". *Neuro Endocrinol Lett.*, 32(6):799–804, 2011.
- [6] Joanna Fedorowicz, Lukasz Sosnowski, Dominik Slezak, Iwona Szymusik, Wojciech Chaber, Lukasz Milobedzki, Tomasz Penza, Jadwiga Sosnowska, Katarzyna Wójcicka, and Karol Zaleski. "Multivariate Ovulation Window Detection at Ovufriend". In Tamás Mihálydeák, Fan Min, Guoyin Wang, Mohua Banerjee, Ivo Düntsch, Zbigniew Suraj, and Davide Ciucci, editors, *Rough Sets - International Joint Conference, IJCRS 2019, Debrecen, Hungary, June 17-21, 2019, Proceedings*, volume 11499 of *Lecture Notes in Computer Science*, pages 395–408. Springer, 2019.
- [7] Lukasz Sosnowski, Iwona Szymusik, and Tomasz Penza. "Network of Fuzzy Comparators for Ovulation Window Prediction". volume 1239 of *Communications in Computer and Information Science*, pages 800–813. Springer, 2020.
- [8] Lukasz Sosnowski and Jakub Wróblewski. "Toward automatic assessment of a risk of women's health disorders based on ontology decision models and menstrual cycle analysis". In Yixin Chen, Heiko Ludwig, Yicheng Tu, Usama M. Fayyad, Xingquan Zhu, Xiaohua Hu, Suren Byna, Xiong Liu, Jianping Zhang, Shirui Pan, Vagelis Papalexakis, Jianwu Wang, Alfredo Cuzzocrea, and Carlos Ordonez, editors, *2021 IEEE International Conference on Big Data (Big Data), Orlando, FL, USA, December 15-18, 2021*, pages 5544–5552. IEEE, 2021.
- [9] B Smoley and C Robinson. "Natural family planning". *Am Fam Physician*, 86(10):924–928, 2012.
- [10] Neil F. Goodman, Rhoda H. Cobin, Walter Futterweit, Jennifer S. Glueck, Richard S. Legro, and Enrico Carmina. "American Association of Clinical Endocrinologists, American College of Endocrinology, and Androgen Excess and PCOS Society Disease State Clinical Review: Guide to the Best Practices in the Evaluation and Treatment of Polycystic Ovary Syndrome - Part 1". *Endocrine Practice*, 21(11):1291–1300, 2015.
- [11] Janusz Kacprzyk, Ronald R. Yager, and José M. Merigó. "Towards Human-Centric Aggregation via Ordered Weighted Aggregation Operators and Linguistic Data Summaries: A New Perspective on Zadeh's Inspirations". *IEEE Comput. Intell. Mag.*, 14(1):16–30, 2019.
- [12] Janusz Kacprzyk and Sławomir Zadrozny. "Fuzzy logic-based linguistic summaries of time series: a powerful tool for discovering knowledge on time varying processes and systems under imprecision". *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, 6(1).
- [13] Soma Dutta and Andrzej Skowron. "Toward a Computing Model Dealing with Complex Phenomena: Interactive Granular Computing". In Ngoc Thanh Nguyen, Lazaros Iliadis, Ilias Maglogiannis, and Bogdan Trawinski, editors, *Computational Collective Intelligence - 13th International Conference, ICCCI 2021, Rhodes, Greece, September 29 - October 1, 2021, Proceedings*, volume 12876 of *Lecture Notes in Computer Science*, pages 199–214. Springer, 2021.
- [14] Andrzej Jankowski, Andrzej Skowron, and Roman W. Swiniarski. "Interactive Complex Granules". *Fundam. Informaticae*, 133(2-3):181–196, 2014.
- [15] Kenneth A. Ginsburg. "Luteal Phase Defect: Etiology, Diagnosis, and Management". *Endocrinology and Metabolism Clinics of North America*, 21(1):85–104, 1992. Reproductive Endocrinology.
- [16] Janusz Kacprzyk, Jan W. Owsinski, Eulalia Szmidi, and Sławomir Zadrozny. "Fuzzy Linguistic Summaries for Human Centric Analyses of Sustainable Development Goals (SDG) Related to Technological Innovations". In José L. Verdegay, Julio Brito, and Carlos Cruz, editors, *Computational Intelligence Methodologies Applied to Sustainable Development Goals*, volume 1036 of *Studies in Computational Intelligence*, pages 19–35. Springer, 2022.
- [17] Mehri Kalhor, Samaneh Yusefloo, Behroz Kaveii, Fatemeh Mohammadi, and Homa and Javadi. "Effect of Yarrow (*Achillea millefolium* L.) Extract on Premenstrual Syndrome in Female Students Living in Dormitory of Qazvin University of Medical Sciences". *Journal of Medicinal Plants*, 18(72), 2019.
- [18] P.I. Good. "Resampling Methods: A Practical Guide to Data Analysis". Birkhäuser Boston, 2005.
- [19] Soma Dutta and Andrzej Skowron. "Concepts Approximation Through Dialogue With User". volume LNCS 11499 of *IJCRS 2019*, pages 295–311, 2019.
- [20] S. Dutta and P. Wasilewski. "Dialogue in Hierarchical Concept Learning using Prototypes and Counterexamples". *Fundamenta Informaticae*, 162:17–36, 2018.