Dealing with Uncertainty in Ovarian Tumor Diagnosis

Andrzej Wójtowicz[1], Patryk Żywica[1], Krzysztof Szarzyński[1], Rafał Moszyński[2], Sebastian Szubert[2], Krzysztof Dyczkowski[1], Anna Stachowiak[1], Dariusz Szpurek[2] and Maciej Wygralak[1]

[1] Department of Imprecise Information Processing Methods
Faculty of Mathematics and Computer Science
Adam Mickiewicz University in Poznań
Umultowska 87, 61-614 Poznań, Poland
andrzej.wojtowicz@amu.edu.pl

[2] Division of Gynaecological Surgery
Poznan University of Medical Sciences
Fredry 10, 61-701 Poznań, Poland
rafalmoszynski@gmail.com

Abstract

In this paper we consider applications of bipolarity in modelling problems encountered in ovarian tumor diagnosis. We focus on imprecision of data obtained by a gynaecologist during examinations. We also present a wide range of predictive diagnostic models and propose a new idea for their improvement.

Keywords: gynaecology, ovarian tumor, computational intelligence, uncertainty, bipolarity, fuzzy set, IFS

1 Introduction

Despite a rapid development of medicine, tumors are still very dangerous. One of the most deadly is the ovarian tumor. In that case, the detectability

rate is still low and mortality rate remains alarmingly high. Reflections of these words are the numbers of women with diagnosed ovarian tumor and their death rate per year. For instance, in Poland it is 3000 and 2000, and in USA it is 22000 and 14000, respectively [10].

Since we distinguish between two types of tumors: malignant or benign, the diagnosis reduces to a binary classification problem. But for each of these two main classes there is a lot of subtypes (histological types). For example, the malignant tumor can be serous adenocarcinoma, mucinous adenocarcinoma, granulosa cell tumor, etc. The benign tumor might be endometrioid cyst, adult teratoma, corpus luteum cyst, etc. Their occurrence among the population varies and some of them are easier to detect by a gynaecologist than others. In consequence, we deal with a sub-classification of the tumor rather than a simple differentiation.

To diagnose malignant ovarian tumors is crucial because this implies the necessity of surgery and the choice of a person to perform the surgery (gynaecological oncologist or general gynaecologist) [7]. Obviously it is a high risk for the woman because the doctor interfere with her body. Secondly, improper treatment may worsen health of the patient. When we have certain diagnosis of a benign tumor we can control the situation. A tragedy occurs when the diagnosis is incorrect (the tumor is malignant in reality) and sometimes it is too late to save the patient.

2 Variety of possible examinations

A total number of factors and examinations essential for ovarian tumor diagnosis is more than fifty. Examples of those are given below:

- menopausal status: whether before or after;
- arrangement of vessels: regular or irregular;
- internal cyst walls: plain, irregular, has papillary projections, change mainly in solid element;
- biochemical blood markers such as CA-125 which stands for carbohydrate antigen which is a tumor marker;
- others: from basic medical history (such as age, contraception, tumors in family, etc.) through ultrasonographic examination to blood markers.

The selection of the most important attributes is crucial for the sensitivity and specificity of diagnosis [13]. Our analysis shown that each attribute has its own subjectivity, influence on differentiation, availability and integrity with other examinations. Some attributes are objective (such as levels of blood markers) but some of examinations need assessment from a gynaecologist who can be a source of subjectivity (the more experienced he/she is, the more confident the result). Some of examinations have less influence on differentiation of a tumor (e.g. number of child-birth) but some are crucial (e.g. vascularization in ultrasonography). It is not possible and not necessary to carry out all possible tests. Moreover, some of the required examinations cannot be carried out due to the technical limitations of the health care unit. Integrity of examination determines whether it can be performed together with others. To ensure the completeness of the diagnosis these factors must also be taken into account.

3 Methods of diagnosis

Popularization of ultrasound resulted in the need to standardize a method for assessing adnexal pathology, which in turn resulted in formation of the first diagnostic models. An important motivation for such models is to help young, inexperienced clinicians in the selection of significant features or examinations for diagnosis. Summary of clinical, ultrasound and biochemical data became the basis for the construction of a predictive model – Risk Malignancy Index (RMI). Publication of the results obtained in the model by Jacobs took place in 1990 and immediately gained a bunch of followers [8].

Over 10 years ago, the International Ovarian Tumor Analysis (IOTA) Group has stated a project to develop a predictive model for differentiation between benign and malignant ovarian tumors. Many years of comprehensive and broad studies resulted in development of number of predictive models. Among them, two models (LR1 and LR2) based on logistic regression are the most important [15].

Both previously mentioned models (LR1 and LR2) are linear regression models. Another approach was presented in sonomorphological index (SM) [11], where the authors identified the most important characteristics of cancer, assigning each appropriate number of points depending on the impact on diagnosis. If the tumor receives score higher than the threshold, it is classified as malicious. A similar approach is presented in the GI-RADS [1] scale but instead of assigning specific points only five levels were defined.

Each of them sets criteria to be met by the tumor at that level. Only tumors identified as GI-RADS 5 are considered to be certainly malignant.

Another interesting approach is offered by Simple Rules Method proposed by IOTA [14]. Five features of malignant and benign tumors were there extracted from the database – five M-rules (malignant) and five B-rules (benign). On the basis of these results the model was build. It was able to precisely classify 76% of tumors with a sensitivity of 93% and specificity of 90%. Recently developed diagnostic scales use state of the art machine learning and classification algorithms such as neural networks [5, 6, 12, 16], Bayesian networks [2] or SVMs [9].

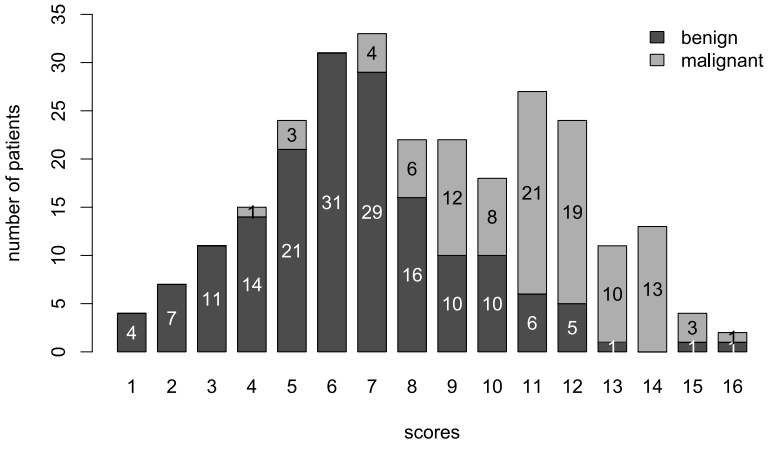
It is also worth noting that a diagnostic scale is almost always biased in either benign or malicious tumor favour (can be liberal or conservative, respectively). More precisely, the scale may have either high precision or specificity. The gynaecologists mostly choose conservative scales because they seem to be safer for patients.

4 Uncertainty in diagnostic scales

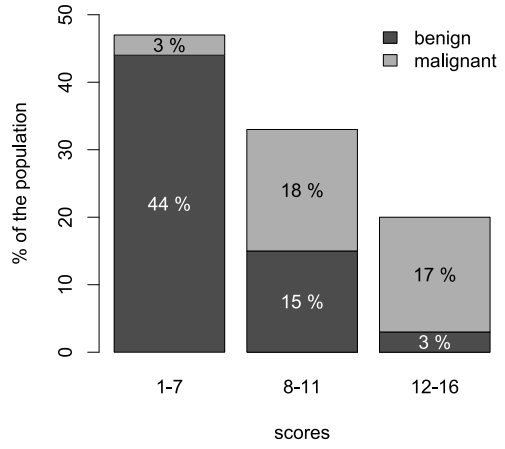
Our study group was 268 women diagnosed and treated due to ovarian tumor in the Division of Gynecological Surgery, Poznan University of Medical Sciences between 2005 and 2012. Among them, 62% was diagnosed with a benign tumor and 38% with malignant. Let us look at sonomorphological index (SM) with cut-off at 8 points. Fig. 1 presents a distribution of SM scores among the population. Clearly, between 8 and 11 the prediction is just a toss of a coin. The question that we try to answer is, if it is possible to improve the results.

We have an uncertainty level in SM scale. What we can do is to check compatibility of the prediction with the results given by another scale. We take a combination of two diagnostic scales and check their conjunction. If both say "it is malignant", then our outcome is "malignant". If they say "it is benign", then our outcome is "benign". Assigning membership/non-membership values to the malignancy and uncertainty, we can easily express our results in the language of IF-sets (Atanassov's intuitionistic fuzzy sets, see [3, 4]).

Some combinations of scales significantly improves differentiation and reduction of uncertainty. For instance, we took a conservative scale (SM) and a liberal one (GI-RADS). Fig. 2 presents the results of such combination. The proper classification of benign tumors stayed at the same level but in case of malignant it rose from 17% to 25%. Moreover, the uncer-



(a)



(b)

Figure 1: Results of sonomorphological index scores among the population. (a) Detailed distribution of scores. (b) Aggregated scoring. 3% of tumors were incorrectly classified as malignant and 3% incorrectly as benign. 33% of cases got scores in range 8 and 11, which is around the cut-off point, and it should be treated as an uncertainty area.

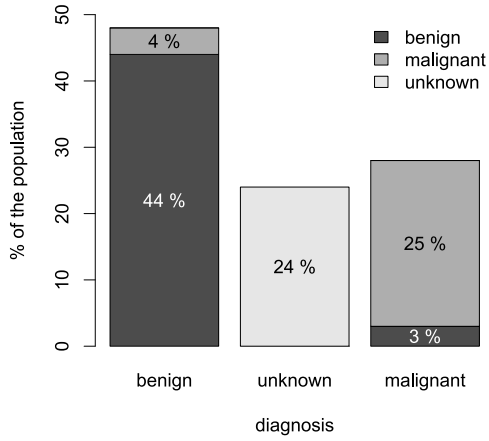


Figure 2: Results of classification of the combined SM and GI-RADS scales. 3% of tumors were incorrectly classified as malignant and 4% incorrectly as benign. In 24% of cases the decisions could not be made.

tainty decreased from 33% to 24%. What is worth noticing, this approach preserves low miss-classification error. McNemar's test we have performed showed that with 95% confidence SM classifies the dataset differently than the combination of SM and GI-RADS does ($p < 0.001$).

5 Further research and conclusions

Investigating the problem of ovarian tumors, we often have to deal with lack of knowledge or with uncertainty. Thus, many of terms might be modelled by IF-sets, especially those with high subjectivity.

Our further goal is to design a system which is a virtual assistant for young and inexperienced gynaecologists. We want to model both a positive and negative prediction of each possible sub-class of a tumor, and in general a malignancy, taking into account both completeness and certainty of overall diagnosis. Another worth mentioning module is a time-line of diagnosis. When new data or examinations are put into the system, such module would suggest currently possible tumors with some possibility information. What is more, the system will display information about where further examination may lead.

Certainly, topics mentioned here are noteworthy but any single research can not exhaust the subject. We still gather new information about patients which will be used to build a validation set for our models.

References

[1] Amor F., Vaccaro H., Alcázar J. L., León M., Craig J. M. and Martinez J. (2009) Gynecologic imaging reporting and data system a new proposal for classifying adnexal masses on the basis of sonographic findings. *Journal of Ultrasound in Medicine 28*, 3, 285–291.

[2] Antal P., Fannes G., Timmerman D., Moreau Y. and De Moor B. (2003) Bayesian applications of belief networks and multilayer perceptrons for ovarian tumor classification with rejection. *Artificial Intelligence in Medicine 29*, 1, 39–60.

[3] Atanassov K. T. (1999) *Intuitionistic Fuzzy Sets: Theory and Applications.* Springer-Verlag.

[4] Atanassov K. T. (2012) *On Intuitionistic Fuzzy Sets Theory.* Springer Publishing Company, Inc.

[5] Biagiotti R., Desii C., Vanzi E. and Gacci G. (1999) Predicting ovarian malignancy: application of artificial neural networks to transvaginal and color doppler flow us. *Radiology 210*, 2, 399–403.

[6] Clayton R., Snowden S., Weston M., Mogensen O., Eastaugh J. and Lane G. (1999) Neural networks in the diagnosis of malignant ovarian tumours. *BJOG: An International Journal of Obstetrics & Gynaecology 106*, 10, 1078–1082.

[7] du Bois A., Rochon J., Pfisterer J. and Hoskins W. J. (2009) Variations in institutional infrastructure, physician specialization and experience, and outcome in ovarian cancer: a systematic review. *Gynecologic Oncology 112*, 2, 422–436.

[8] Jacobs I., Oram D., Fairbanks J., Turner J., Frost C. and Grudzinskas J. (1990) A risk of malignancy index incorporating CA-125, ultrasound and menopausal status for the accurate preoperative diagnosis of ovarian cancer. *BJOG: An International Journal of Obstetrics & Gynaecology 97*, 10, 922–929.

[9] Lu C., Van Gestel T., Suykens J. A., Van Huffel S., Vergote I. and Timmerman D. (2003) Preoperative prediction of malignancy of ovarian tumors using Least Squares Support Vector Machines. *Artificial Intelligence in Medicine 28*, 3, 281–306.

[10] Siegel R., Naishadham D. and Jemal A. (2013) Cancer statistics, 2013. *CA: A Cancer Journal for Clinicians 63*, 1, 11–30.

[11] Szpurek D., Moszynski R., Zietkowiak W., Spaczynski M. and Sajdak S. (2005) An ultrasonographic morphological index for prediction of ovarian tumor malignancy. *European Journal of Gynaecological Oncology 26*, 1, 51–54.

[12] Tailor A., Jurkovic D., Bourne T. H., Collins W. P., and Campbell S. (1999) Sonographic prediction of malignancy in adnexal masses using an artificial neural network. *BJOG: An International Journal of Obstetrics & Gynaecology 106*, 1, 21–30.

[13] Timmerman D. (2004) The use of mathematical models to evaluate pelvic masses; can they beat an expert operator? *Best Practice & Research Clinical Obstetrics & Gynaecology 18*, 1, 91–104.

[14] Timmerman D., Testa A. C., Bourne T., Ameye L., Jurkovic D., Van Holsbeke C., Paladini D., Van Calster B., Vergote I., Van Huffel S. et al. (2008) Simple ultrasound-based rules for the diagnosis of ovarian cancer. *Ultrasound in Obstetrics & Gynecology 31*, 6, 681–690.

[15] Timmerman D., Testa A. C., Bourne T., Ferrazzi E., Ameye L., Konstantinovic M. L., Van Calster B., Collins W. P., Vergote I., Van Huffel S. et al. (2005) Logistic regression model to distinguish between the benign and malignant adnexal mass before surgery: a multicenter study by the international ovarian tumor analysis group. *Journal of Clinical Oncology 23*, 34, 8794–8801.

[16] Timmerman D., Verrelst H., Bourne T., De Moor B., Collins W., Vergote I. and Vandewalle J. (1999) Artificial neural network models for the preoperative discrimination between malignant and benign adnexal masses. *Ultrasound in Obstetrics & Gynecology 13*, 1, 17–25.