Rough Set Rules Determine Disease Progressions in Different Groups of Parkinson's Patients

Andrzej W. Przybyszewski[1,3(✉)], Stanislaw Szlufik[2], Piotr Habela[1], and Dariusz M. Koziorowski[2]

[1] Polish-Japanese Academy of Information Technology, 02-008 Warsaw, Poland
{przy, piotr.habela}@pja.edu.pl
[2] Department of Neurology, Faculty of Health Science, Medical University Warsaw, Warsaw, Poland
stanislaw.szlufik@gmail.com, dkoziorowski@esculap.pl
[3] Department of Neurology, University of Massachusetts Medical School, Worcester, MA 01655, USA

Abstract. Parkinson's disease (PD) is the second after Alzheimer most popular neurodegenerative disease (ND). We do not have cure for both NDs. Therefore the purpose of our study was to predict results of different PD patients' treatments in order to find an optimal one.

We have used rough sets (RS) and machine learning (ML) rules to describe and predict disease progression (UPDRS - Unified Parkinson's Disease Rating Scale) in three groups of Parkinson's patients: 23 BMT patients on medication; 24 DBS patients on medication and on DBS therapy (deep brain stimulation) after surgery performed during our study; and 15 POP patients that have surgery earlier (before beginning of our study). Every PD patient had three visits approximately every 6 months. The first visit for DBS patients was before surgery.

On the basis of the following condition attributes: disease duration, saccadic eye movement parameters, and neuropsychological tests: PDQ39, and Epworth tests we have estimated UPDRS changes (as the decision attribute).

By means of ML and RS rules obtained for the first visit of BMT/DBS/POP patients we have predicted UPDRS values in next year (two visits) with the global accuracy of 70% for both BMT visits; 56% for DBS, and 67, 79% for POP second and third visits.

We have used rules obtained in BMT patients to predict UPDRS of DBS patients; for first session DBSW1: global accuracy was 64%, for second DBSW2: 85% and the third DBSW3: 74% but only for DBS patients during stimulation-ON. These rules could not predict UPDRS in DBS patients during stimulation-OFF visits and in all conditions of POP patients.

Keywords: Neurodegenerative disease · Rough set · Decision rules · Granularity

B.U. Shankar et al. (Eds.): PReMI 2017, LNCS 10597, pp. 270–275, 2017.
https://doi.org/10.1007/978-3-319-69900-4_34

1 Introduction

Only very experience PD neurologists are successful in implementing individually adjusted therapy. In general doctors have very limited time for each patient and different approaches to patients that may lead to confusions and ineffective therapy. We propose to improve doctor's approach by additional more automatic measurements and intelligence symptom classification [1] that is similar to that found in the visual system for the complex objects recognition [2].

It is important to estimate the disease stage because it determines different sets of therapies. The neurological standards are based on Hoehn and Yahr and the UPDRS (Unified Parkinson's Disease Rating) scales. The last one is more precise and it will be used in this study. We would like to estimate disease progression in different groups of patients that were tested during three visits every half-year. Our method may lead to introduce more precise follow up and introduction of the possible internet-treatment.

2 Methods

All 62 PD patients were divided into three groups: BMT patients (only medication), and patients on medication and with implanted electrodes in the STN (subthalamic nucleus [3]) during our study: DBS group or before our study: POP group.

The Deep Brain Stimulation (DBS) surgery was performed in the Institute of Neurology and Psychiatry WUM. PD patients were tested in the following sessions: MedON/MedOFF sessions (sessions with or without medication). The other groups: DBS and POP patients were also tested in StimON/StimOFF session were DBS stimulation was switched ON or OFF. All combinations gave four sessions: (1) MedOFFStimOFF; (2) MedOFFStimON; (3) MedONStimOFF; (4) MedONStimON. Details of these procedures were described earlier [2]. The UPDRS tests and neuropsychological tests were performed by neurologists from Warsaw Medical University. Fast eye movements (EM) - reflexive saccades (RS) were recorded as described in details before [1, 3]. The following parameters of RS were measured: the delay (latency) related to time difference between the beginning of the light spot movements and the beginning of the eye movement; saccade's amplitude in comparison to the light spot amplitude; max velocity of the eye movement; duration of saccade defined as the time from the beginning to the end of the saccade.

2.1 Theoretical Basis

Our data mining analysis follows rough set (RS) theory after Zdzislaw Pawlak [4]) because RS gave the best results in PD symptoms classifications in comparison to other methodologies [1]. Our data are represented as a decision table where rows represented different measurements (may be obtained from the same or different patients) and columns were related to different attributes. An information system [4] is as a pair $S = (U, A)$, where U, A are finite sets: U is the universe of objects; and A is the set of attributes. The value $a(u)$ is a unique element of V (where V is a value set) for $a \in A$ *and* $u \in U$.

A decision table for S is the triplet: $S = (U, C, D)$ where: *C, D* are condition and decision attributes [5]. Each row of the information table gives a particular rule that connects condition and decision attributes for a single measurements of a particular patient. As there are many rows related to different patients and sessions, they gave many particular rules. Rough set approach allows generalizing these rules into universal hypotheses that may determine optimal treatment options for an individual PD patient. Different rules' granularities (abstraction) are similar to complex objects recognition [2] and may simulate association processes of the 'Golden Neurologist'.

In the present study, we are trying to use data from different groups of patients for training and testing. The purpose was to find what are limits of rules that may predict symptoms development of patients with different treatments in different disease stages.

We have used the RSES 2.2 (Rough System Exploration Program) [6] with implementation of RS rules to process our data.

3 Results

All 62 PD patients were divided into three groups: BMT patients (only medication), and patients on medication and with implanted electrodes in the STN (subthalamic nucleus [3]) during our study: DBS group or before our study: POP group.

In 23 patients of BMT group the mean age was 57.8+/− 13 (SD) years; disease duration was 7.1+/− 3.5 years, UPDRS was 36.1+/− 19.2. In 24 patients of DBS group the mean age of 53.7+/− 9.3 years, disease duration was 10.25+/− 3.9 years (stat. diff. than BMT-group: $p < 0.025$), UPDRS was 62.1+/− 16.1 (stat. diff. than BMT-group: $p < 0.0001$). In 15 patients of POP group the mean age was 56.2+/− 11.3 (SD) years and disease duration was 13.5+/− 3.6 years (stat. diff. than DBS-group: $p < 0.015$), UPDRS was 59.2+/− 24.5 (stat. diff. than BMT-group: $p < 0.0001$).

These statistical data are related to the data obtained during the first visit for each group: so-called BMT W1 (visit one), DBS W1 (visit one) and POP W1 (visit one).

3.1 BMT Patients' Rules for the Disease Progression

The BMT patients (only on medication) were tested in two sessions (session 1: without, and session 3: with medication) three times every half-year.

We have used ML and rough set theory [6] in order to obtain rules determining decision and condition attributed for the first visit BMTW1. On the basis of these rules we have predicted the UPDRS values obtained during the second (half -year later W2 – BMTW2) and the third (one year later BMTW3) visits. UPDRS was optimally divided by RSES into 4 ranges: *"(−Inf, 24.0)", "(24.0, 36.0)", "(36.0, 45.0)", "(45.0, Inf)"* for both visits (W2 and W3) the global coverage was 1.0 and the global accuracy was 0.7. Example of rules from BMTW1:

$$(Ses = 3)\&(PDQ39 = "(-Inf, 50.5)") => (UPDRS = "(-Inf, 33.5)"[12])\ 12 \quad (1)$$

$$(dur = \text{"}(-Inf, 5.65)\text{"})\&(Ses = 3)\&(Epworth = \text{"}(-Inf, 14.0)\text{"}) \Rightarrow (UPDRS = \text{"}(-Inf, 33.5)\text{"}[7])\ 7 \quad (2)$$

$$(dur = \text{"}(5.65, Inf\text{"}))\&(Ses = 3)\&(Epworth = \text{"}(14.0, Inf)\text{"}) \Rightarrow (UPDRS = \text{"}(-Inf, 33.5)\text{"}[4])\ 4 \quad (3)$$

In the first rule (1) if the session number 3 and PDQ39 = (−Inf, 50.5) then UPDRS was (−Inf, 33.5) in 12 cases. The second rule (2) was fulfilled in 7 cases and the third one (3) in 4 cases. There were 70 rules.

3.2 DBS and POP Patients' Rules for the Disease Progression

As **DBSW1** had only 2 sessions (before surgery) we could only predict session **DBSW3** on the basis of **DBSW2** (half of the year earlier) (Table 1).

Table 1. Confusion matrix for UPDRS of DBSW3 by rules obtained from DBSW2.

		Predicted					
		"(46.0, 72.0)"	"(38.0, 46.0)"	"(19.5, 38.0)"	"(72.0, Inf)"	"(Inf, 19.5)"	ACC
Actual	"(46.0, 72.0)"	12	5	2	5	1	0.48
	"(38.0, 46.0)"	2	5	1	2	2	0.42
	"(19.5, 38.0)"	0	4	13	3	7	0.48
	"(72.0, Inf)"	4	0	0	12	0	0.75
	"(-Inf, 19.5)"	0	0	4	0	12	0.75
	TPR	0.67	0.4	0.65	0.55	0.6	

TPR: True positive rates for decision classes; ACC: Accuracy for decision classes: the global coverage was 1 and the **global accuracy was 0.562**

POP patients' rules for the disease progression. As above, we have predicted UPDRS for visits **POPW2** and **POPW3** on the basis of visit **POPW1** with total accuracy: 0.667 and 0.793 with a coverage: 1 and 0.967.

3.3 BMT Patients' Rules for Estimation of DBS Patients' Disease Progression

As BMT patients have only two sessions (S1 – MedOff, and S3 – MedON) and DBS patients four sessions (see Methods) we have divided them to two sets: one with StimON set-up and another one with StimOFF set-up. We were not successful in prediction SimOFF sessions as DBS patients were in more advanced stage than BMT group. Our UPDRS predictions for **DBSW1** had global accuracy 0.64 (coverage 0.5);

for **DBSW2** - global accuracy was 0.85 (coverage 0.3); for **DBSW3** - global accuracy was 0.74 (coverage 0.6).

3.4 DBS Patients' Rules for Estimation of POP Patients' Disease Progression

We could not predict UPDRS of POP patients from rules obtained from DBS patients probably because many years of DBS have changed some brain circuits.

4 Discussion

There are novel technologies and data constantly improve PD patients' treatments, but are also still doubts if the actual procedures are optimal for a particular individual case. Our long time purpose is to use the data mining and machine learning in order to *compare different neurological protocols and their effectiveness*. We think that the best future approach will be to perform all tests automatically at home, process them with intelligent algorithms and to submit results to the doctor for his/her decision. Another, more advanced approach that we were testing in this work, would be to create the standard treatment for each new case on the basis of already successfully treated patients and correct treatment as symptoms are developing in time. We have demonstrated that relatively easy to estimate symptoms and their time development in populations treated in a different ways (e.g. only medication treatment). This result may give the basic (locally optimal) follow-ups. If patient is doing significantly worse then others (rules), his/her treatment is not optimal, and should be changed. In the next step, we may use rules obtained from different clinics to get them even more universal and optimal. Our new approach is related not only to longitudinal study but also test different patient population with different treatments. Can we in this case find optimal way of different treatments? The second group of patients were in more advanced stage of disease so it was not possible to get 100% coverage like in the first case. The second group with longitudinal study had a new treatment (brain stimulation) that started from the second visit. We have tested if the same treatment in different populations gives similar results. Patients got two treatments: medication (medication ON and OFF) and electric brain stimulation (ON and OFF). We have analyzed these treatments as two different sets: (1) StimOFF: medication ON and OFF; (2) StimON: medication ON and OFF. As a result, it was not possible to get sufficient accuracy in the first situation, but we got good accuracy in the second case- with the brain stimulation. However, our third POP group was different than two other as we did not succeeded to obtain good prediction by rules obtained by other groups BMT or DBS. It maybe related to the longer period of brain stimulation (DBS) that has changed some central mechanisms. It is an important negative result that needs more study. In the near future, we may look for additional condition attributes in order to improve a global accuracy. The reason that our rules did not apply to symptoms of patients without brain stimulation might be related to the surgery. Inserting electrodes through the brain till the basal ganglia probably partly

destroys some of these pathways. Functions of these connections are expressed by our rules, damaging them changes their functionality. We have demonstrated that the DBS (electric STN stimulation) procedure revoked and improved rules that became similar to rules of early stage Parkinson's disease patients.

5 Conclusions

This work is a continuation of our previous findings [1, 3], comparing classical approach used by most neurologists and based on their partly subjective experience and intuitions with the intelligent data processing (machine learning, data mining) classifications. We have demonstrated that the parameters of eye movements and neuropsychological data are sufficient to predict longitudinal symptom developments in different therapy related groups of PD patients.

Acknowledgements. This work was partly supported by projects Dec-2011/03/B/ST6/03816, from the Polish National Science Centre.

References

1. Przybyszewski, A.W., Kon, M., Szlufik, S., Szymanski, A., Koziorowski, D.M.: Multimodal learning and intelligent prediction of symptom development in individual parkinson's patients. Sensors **16**(9), 1498 (2016). doi:10.3390/s16091498
2. Przybyszewski, A.W.: Logical rules of visual brain: from anatomy through neurophysiology to cognition. Cogn. Syst. Res. **11**, 53–66 (2010)
3. Przybyszewski, A.W., Kon, M., Szlufik, S., Dutkiewicz, J., Habela, P., Koziorowski, D.M.: Data mining and machine learning on the basis from reflexive eye movements can predict symptom development in individual Parkinson's patients. In: Gelbukh, A., Espinoza, F.C., Galicia-Haro, S.N. (eds.) MICAI 2014, Part II. LNCS, vol. 8857, pp. 499–509. Springer, Cham (2014). doi:10.1007/978-3-319-13650-9_43
4. Pawlak, Z.: Rough sets: Theoretical Aspects of Reasoning About Data. Kluwer, Dordrecht (1991). p. 509. Springer (2014)
5. Bazan, J., Son, Nguyen, H., Trung, T., Nguyen, Skowron A., Stepaniuk, J.: Decision rules synthesis for object classification. In: Orłowska, E. (ed.) Incomplete Information: Rough Set Analysis, pp. 23–57. Physica–Verlag, Heidelberg (1998)
6. Bazan, J.G., Szczuka, M.: RSES and RSESlib - a collection of tools for rough set computations. In: Ziarko, W., Yao, Y. (eds.) RSCTC 2000. LNCS(LNAI), vol. 2005, pp. 106–113. Springer, Heidelberg (2001). doi:10.1007/3-540-45554-X_12