Knowledge Engineering in Construction of Expert Systems on Hereditary Diseases

Boris A. Kobrinskii[1(✉)], Nataliya S. Demikova[2], and Nikolay A. Blagosklonov[1]

[1] Federal Research Center "Computer Science and Control" of RAS, Vavilova str. 44, kor. 2, Moscow 119333, Russian Federation
kba_05@mail.ru

[2] Russian Medical Academy of Continuous Professional Education, Barrikadnaya str. 2/1, Moscow 125993, Russian Federation

Abstract. The paper considers medical knowledge extraction for the diagnosis of rare hereditary diseases. The specific feature of the proposed approach is the combination of expert estimates of the symptom presence probability with three complementary factors of confidence: time of manifestation, severity and frequency of symptom manifestation in different age ranges. Preliminary textological cards used by experts and cognitive scientists are formed based on the linguistic analysis of related literature.

Keywords: Knowledge engineering · Confidence factors · Linguistic analysis
Textological cards · Hereditary diseases · Orphan diseases
Holistic visual images

1 Introduction

Unclear pathological manifestation is a characteristic feature of clinical medicine. This is most clearly manifested in hereditary diseases of progredient nature. The difficulty is in the assessment of intermediate states which gradually increase with age of changes. Even a skillful physician experiences difficulty due to the rarity of genetically determined diseases in his/her practice. At the same time orphan (rare) hereditary diseases are an important social problem [1]. The necessity for early detection is determined by the fact that the progression of a number of them can be prevented. However, the regression of pathological changes is impossible. Therefore, the delay in treatment due to late diagnosis leads to remediless results. However, on the one hand, hereditary diseases are very rare but, on the other hand, their total number is near 6,000 nosological forms [2]. It is clear that no physician can remember the manifestations of even several hundreds of diseases. Moreover, the heterogeneity of clinically similar diseases (variants with different genetic changes) is combined with polymorphism (polyvariety) and at the same time partial overlapping of the spectrum of clinical features. These characteristics of hereditary pathology lead to a large number of erroneous and late diagnosed cases. Thus, more than 30 years ago the development of computer diagnostic systems in the field of clinical genetics started. Some of systems were later upgraded [3–9]. They allowed increasing the quality of genetically determined diseases recognition. However, up to

S. O. Kuznetsov et al. (Eds.): RCAI 2018, CCIS 934, pp. 35–45, 2018.
https://doi.org/10.1007/978-3-030-00617-4_4

date the hypotheses for differential diagnosis of many similar diseases, especially in the early stages, with poorly expressed pathological changes are practically not effective. Therefore, it is of vital importance to search for new approaches to the construction of intelligent diagnostic systems in the subject area of hereditary diseases that will take into account the dynamics of the processes.

The complexity of the subject area determines the search in the direction of knowledge engineering including the evaluation of confidence measures regarding the occurrence of signs characterized by unclear manifestations, an additional problem being the reflection of experts.

2 Features of the Clinic and Diagnosis of Hereditary Diseases

Hereditary diseases, which are related to orphan ones, present serious difficulties for the diagnostics. In most cases physicians have little personal experience of observing such patients. At the same time, one of the main tasks is early detection. This is especially important for lysosomal storage diseases occurring at the frequency of 1 per 7,000 to 8,000 newborns while each disease from this group occurs in only 1 in 100,000 newborns [10]. These diseases are characterized by a worsening progression, and it is extremely difficult to identify manifesting clinical signs. This is connected to the vagueness and at the same time to the low frequency of occurrence of individual pathological features. In addition, some signs may change to the opposite ones in the progression. For example, loose joints due to the accumulation of deposited macromolecules (glycosaminoglycans) at the mucopolysaccharidosis leads to an ever-increasing stiffness of joints.

Lysosomal diseases are characterized by a combination of numerous pathological changes associated with the accumulation of macromolecules in various organs and systems of the body. These diseases include mucopolysaccharidosis, mucolipidosis, gangliosidosis and others. Depending on the location of abnormal focuses and the speed of macromolecular deposition, a different clinical presentation or clinical variants ("portraits") of the disease takes place including the age of the manifestation, the severity of the symptoms, the degree of incidence at a particular age. In this regard, there are fuzzy intermediate states in the dynamics of the pathological process, various combinations of features, and varying degrees of their severity.

At the same time the need for the earliest possible detection of lysosomal accumulation diseases characterized by a deficiency of enzymes can be explained by the recent availability of enzyme replacement therapy.

The presented features of clinical manifestations of rare hereditary diseases create certain limitations in the application of computer decision support methods. At the same time, the requirements for knowledge about the diagnosis of orphan pathology are increasing.

Cumulative knowledge including its own verbalized and implicit (intuitive) knowledge in combination with well-known knowledge from various sources is an important factor for the knowledge base formation for a diagnostic system by hereditary diseases.

3 Fuzzy Sets of Transition States of Diseases

Each group of diseases can be formally represented by sets and subsets (multisets - for a higher level of associations of similar diseases). For orphan pathology, fuzzy sets (with inaccurately defined boundaries) are characteristic. This is particularly evident with progressive accumulation diseases. They are characterized by a constant process of increasing the severity of previously manifested pathological changes and the appearance of pathological changes in other organs and systems of the organism. In clinical practice, doctors constantly encounter situations that have minor differences. In addition, there can be fuzzy attributes, and fuzzy references to certain classes. In the course of chronic diseases, the signs tend to undergo changes, which are differently assessed and described by each of the physicians. A great diagnostic value in hereditary diseases can also have appearance (visual images) [11]. The idea of a visual series should be perceived as a visual quasi-continuum of close or relatively close images, which implies the vagueness of transitions between individual representatives (Fig. 1). However, an appearance is characterized by fuzzy transitions, which form a different visual row of patients of one subset, and sometimes of one set. This creates additional diagnostic difficulties and requires a verbal appreciation based on a fuzzy scale, similar to linguistic features. A fairly broad range of opinions of fuzzy verbal definitions to the views expressed by doctors is typical, which can be often combined with the modality "it seems to me".

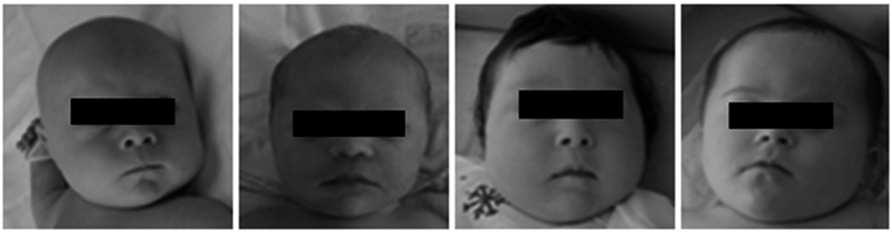

Fig. 1. Visual images of patients of the same type

To model the allowable level of inaccuracy in solving a particular task, Lotfi Zadeh proposed to consider degrees of granulation of measuring information [12], meaning by a granule a group of objects united by the relations of indistinguishability, similarity, proximity, or functionality. The granulation of information is based on the nonclassical representation of the set. Different types of fuzzy measures correspond to different points of view on the evaluation of the certainty of events. The theory of possibilities of Zadeh [13], allowing several (may be continuum set of) degrees of possibility, actually departs from modal semantics. In medicine, this corresponds to the existence of many variants of transitional states of manifestations of individual signs or health status. The problem of separating the continuum into diagnostically significant intervals is that the integrating estimates include numerous fuzziness. In the integral analysis of the organism's conditions, such concepts as trend, dynamics, which are almost always characterized by fuzzy marks, prevail. One of the variants of the subdivision of the health continuum into

intervals meaningful for a particular subject area, corresponding to the range of values of some notion by the criterion of maximizing the "information gain" measure, is proposed by J.R. Quinlan [14]. The decomposition of fuzzy sets allows us to represent an arbitrary fuzzy set as a sum of fuzzy sets [15], corresponding to separate nosological forms or subclasses of diseases. These fuzzy sets can result from formalized expert evaluations in linguistic form or in the form of digital labels corresponding to the parameters at the linguistic scale. Combinations of features represented by fuzzy sets/subsets allow us to characterize various variants of independent diseases or separate clinical forms. The method of ordering objects is based on an assessment of their proximity to some "ideal" object in a multivalued space. The method of ordering objects is based on an assessment of their proximity to some "ideal" object in a multivalued space.

The metric problem for measuring the distances between objects in this space can be solved by representing multisign objects based on the multiset formalism [16]. In medicine, as an "ideal" object, a case can be considered, characterized by a typical combination of symptoms most often repeated in a subclass of diseases.

4 Linguistic Analysis as the First Stage of Knowledge Extraction

The unclear representation of descriptions in the literature is determined by a number of factors: clinical picture of the disease, age characteristics of the patient, rate of the pathological process, knowledge of physicians and their subjective preferences based on past events that have been observed in the past and known from other studies. At the same time, various diseases may be hidden behind these differences. This fact requires special attention. Therefore, linguistic scales of medical parameters and fuzzy logic have particular importance for supporting effective differential diagnosis of diseases and conditions of the organism [17, 18].

Knowledge engineering assumes different approaches to the knowledge representation. At the first (preliminary) stage this can be the analysis of special (medical) literature sources. The data will serve as the additional material for the experts at the next stage.

Taking into account inhomogeneity and multilinguality of initial sources, an integrated approach combining the semantic, textual, and linguistic methods of text analysis was used to extract knowledge from literature.

Semantic analysis allows one to evaluate the meaning of the text. Based on results of the semantic analysis, it is possible to determine the frequency of various terms occurrence. In the problem area of orphan diseases diagnosis, it is necessary to identify and rank certain indicators. This is the age of manifestation of symptoms, the dynamics of their changes and the degree of incidence.

Textual analysis uses methods of studying the text based on the comparative research method. As applied to the problem under study, it should be a comparison of various publications on the diagnosis of hereditary diseases. Among others, a cognitive comparative analysis of manifestations of diseases in various ethnic groups is carried out.

Linguistic analysis involves the study of various aspects of diseases (etiology, pathogenesis, clinical manifestations), as well as the correlation of multilingual terms, the

identification of terms-metaphors, etc. The result of a complex analysis of literature sources using the mentioned-above approaches is a structured representation of the extracted knowledge in the form of textological cards (see Table 1).

Table 1. Fragment of the textological card of the Hunter syndrome

Symptom	Presence	Level of manifestation	Age group	Degree of incidence	Source
Mental retardation	+	Strong	1 to 3 years	Frequently	*
	+	Very mild	over 6 years	Rare	**
	+	Strong	–	Frequently	***
Coarse facial features	+	Very strong	1 to 3 years	Very frequently	*
	+	Strong	1 to 3 years	Frequently	**
	+	Strong	–	Frequently	***
Macroglossia	+	Strong	1 to 3 years	Very frequently	*
	+	Strong	1 to 3 years	–	**
	+	Strong	–	Frequently	***

* Wraith, J.E., Scarpa, M., Beck, M., Bodamer, O.A., De Meirleir, L., Guffon, N., Meldgaard Lund, A., Malm, G., Van Der Ploeg, A.T., Zeman, J.: Mucopolysaccharidosis type II (Hunter syndrome): A clinical review and recommendations for treatment in the era of enzyme replacement therapy. Eur. J. Pediatr., **167**(3), 267–277 (2008)

** Namazova-Baranova, L.S., Vashakmadze, N.D., Gevorkyan, A.K., Altunin, V.V., Kuzenkova, L.M., Chernavina, E.G., Babaykina, M.A., Podkletnova, T.V., Kozhevnikova, O.V.: Obstructive Sleep Apnea Syndrome in Children with Type II Mucopolysaccharidosis (Hunter Syndrome). Pediatric pharmacology, **10**(6), 76–81 (2013) (In Russian)

*** Scarpa, M., Almássy, Z., Beck, M., Bodamer, O., Bruce, I.A., De Meirleir, L., Guffon, N., Guillén-Navarro, E., Hensman, P., Jones, S., Kamin, W., Kampmann, C., Lampe, C., Lavery, C.A., Leão Teles, E., Link, B., Lund, A.M., Malm, G., Pitz, S., Rothera, M., Stewart, C., Tylki-Szymaska, A., Van Der Ploeg, A., Walker, R., Zeman, J., Wraith, J.E.: Mucopolysaccharidosis type II: European recommendations for the diagnosis and multidisciplinary management of a rare disease. Orphanet J. Rare Dis., **6**(1), 1–18 (2011)

Knowledge engineer filling out a textological card uses the scales specially developed for this study.

The scale of symptom presence is presented in the form of " + " and "−" showing presence or absence of the indicator.

The rank scale of levels of symptom manifestation includes 5 linguistic estimates: very strongly expressed, strongly expressed, moderately expressed, mildly expressed, expressed very mildly.

The age scale contains five gradations: (1) the neonatal period i.e. the first 28 days of life, (2) the first year of the child life, (3) 1 to 3 years, (4) 4 to 6 years, (5) over 6 years.

The linguistic scale of symptom incidence is the following: very frequent (more than 80%), frequent, (60–80%), relatively often (from more than 30% to 60%), rare (from more than 15 to 30%), very rare (less than 15%).

Textual cards are used by experts and knowledge engineer in the process of forming a "disease-signs" matrix.

5 Probability and Factors of Confidence in the Diagnosis of Hereditary Storage Diseases

In medicine, signs are in most cases stable concepts, but the attributes that characterize them indicate a significant variety of their manifestations and severity. Matrices for the formalized presentation of diseases with expert estimates for each nosological form at certain age intervals allow to be taken into account for the pathological process dynamics.

In this study the probability of developing a sign (attribute), called "the presence of a sign", was first indicated in each age group. This evaluation is based on the personal knowledge of the experts who took into account the views of other researchers presented in the textological cards. The scale for the expert evaluation of the probability of the sign presence includes 7 gradations in the interval from −1 to +5: −1 stays for the impossibility of the symptom manifestation in a certain age range, 0 stays for the normal range (the absence of a pathological feature being theoretically possible at this age), 1 stays for minimal possibility, 2 for remotely possible, 3 for possible, 4 for highly possible, 5 for maximum possibility. Such a linguistic scale corresponds to the accepted estimates of observed signs in medical practice. As an example of the gradation usage, −1 may be a sign of "tooth pathology in a newborn" because of the impossibility of its presence at a given age. When experts analyzed textological cards, the opinions of medical specialists (for example, ophthalmologists in assessing the pathological symptoms from the side of the vision organ) were of great importance.

The objectivity of the conclusion in each case was determined by a joint assessment of confidence measures by two experts who jointly make decisions. At first, one of the experts offered his assessment. If the second expert agreed the decision was taken. In case of disagreement, a dialogue of experts was carried out with the argumentation of their opinions. If necessary, additional literature sources were used in the discussion process. In the discussion, a cognitive scientist participated. He could comment on textological cards for these diseases.

Considerable part of the subjective knowledge is determined by the dispositions of the expert including those which are modified by the influence of their personal experience [19]. The result of personal experience is a measure of confidence in one's own and other people's knowledge.

Fuzzy logical conclusions create a model of rough reflection of a person, in particular a doctor, in the moment when a decision has to be made. A measure of confidence or an expert confidence factor is an informal assessment attached to the conclusions. In the Stanford model the certainty factor is a value from the interval [−1; 1], which numerically defines the confidence measure of the expert decision [20]. In linguistic form it can be presented, e.g., as follows: "almost impossible", "very doubtful", "doubtful", "can be suspected", "cannot be excluded", "most likely" or a large number of degrees of freedom.

In the medical diagnosis of diseases with progredient nature certainty factors by expert are specific for different age groups in terms of the severity of signs (symptoms) characterized by a fuzzy scale of changes, the timing of their manifestation and the degree of incidence.

In the clinical picture of diseases with progredient nature signs in most cases are stable concepts but the meanings of attributes that characterize the severity and time of manifestation are determined by the age-related dynamics of the pathological process.

"Sign presence" attributes are accompanied by three certainty factors: (a) the manifestation of the sign at a certain age, (b) the degree of incidence of the sign in the age group, (c) the severity of the sign. Estimates of the certainty factors lie in the area between 0.1 and 1.0. A confidence measure of 1.0 was used by experts only in cases where the presence of a sign is −1 or 0. The introduction of three complementary certainty factors allows us to obtain a complete ("three-dimensional") picture of the confidence measure of the analyzed sign including the time of its manifestation, the changing severity of manifestations and the frequency in each of the age groups. When deciding on the issue of the certainty factors the expert physicians turn to their intuitive knowledge, which expresses itself in the process of deciding on the choice of the estimates to be presented in verbally inexplicable changes.

Certainty factors significantly increase the likelihood of differentiation between certain diseases. It is especially important to introduce three confidence measures to the signs characterizing the clinical forms of diseases, for example, mucopolysaccharidosis type I, which according to the international classification of hereditary diseases [2] is represented by three types (IH – OMIM 607014, IH/S – OMIM 607015, IS – OMIM 607016). In fact, in this case there is a continuum of values of each characteristic, which was conditionally quantized by experts. The intermediate nosological form IH/S includes manifestations from severe (close to the subtype IH) to relatively light (close to the subtype IS). As another example, the "hepatomegaly (enlarged liver)" sign is considered in the case of mucopolysaccharidosis type II (Hunter syndrome – OMIM 309900), which includes two clinical forms, severe and light. In the severe form the sign presence in the first year of life corresponds to the "cannot be excluded" concept. Three certainty factors agreed by the experts were as follows: the sign manifestation at this age is 0.6, the degree of incidence is 0.4; level of manifestation is 0.3. The certainty factors in the next age range increase to 0.6; 0.5; 0.4, respectively. Another example concerns the "cuboidal vertebral bodies" sign at the age of 2 to 3 determined by radiological examination of the patients with mucopolysaccharidosis type IH (Hurler syndrome). Given the fact that the experience of the radiologist influences the detection of this sign, experts came to the conclusion that the manifestation of this sign can be characterized by the following certainty factors: manifestation – 0.4, the degree of incidence – 0.6, degree of severity – 0.7. An example of a table with signs (the top line for each sign-characteristic indicates presence in five age groups, respectively) and their confidence measure (three positions in the next line) for the Hunter syndrome (severe subtype) is presented in Table 2.

It should be noted that for Hurler syndrome and a number of other diseases, holistic signs (for example, rough facial features previously metaphorically named "gargoyle-like face") are of great importance in the diagnostic but their defragmentation reduces the ability to recognize the disease. Therefore, holistic visual images should be included in the knowledge base with expert confidence measures [11].

Table 2. Fragment of the table with expert confidence factors

Signs	Hunter syndrome (severe subtype)														
Macrocephaly	−1			0			2			3			4		
	1	1	1	1	1	1	0,9	0,8	0,6	1	0,9	0,8	1	0,9	0,9
Coarse facial features	0			0			3			4			4		
	1	1	1	1	1	1	0,4	0,4	0,3	0,6	0,6	0,5	0,8	0,8	0,9
Macroglossia	−1			0			3			4			4		
	1	1	1	1	1	1	0,3	0,4	0,4	0,5	0,6	0,6	0,8	0,8	0,7
Teeth widely spaced	−1			−1			3			3			3		
	1	1	1	1	1	1	0,3	0,3	0,3	0,5	0,5	0,5	0,7	0,8	0,8
Mental retardation	−1			0			2			4			4		
	1	1	1	1	1	1	0,3	0,3	0,3	0,5	0,5	0,5	0,2	0,8	0,9
Chest Funnel-shaped	−1			0			1			1			1		
	1	1	1	1	1	1	0,1	0,1	0,2	0,1	0,1	0,2	0,1	0,1	0,2
Joints hypermobility	0			0			0			0			0		
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Joint contractures	−1			−1			3			5			5		
	1	1	1	1	1	1	0,3	0,4	0,2	0,5	0,6	0,4	0,7	0,7	0,8
Carpal tunnel syndrome	−1			−1			0			4			5		
	1	1	1	1	1	1	1	1	1	0,5	0,5	0,6	0,7	0,7	0,8
Claw-like hand	−1			−1			0			3			4		
	1	1	1	1	1	1	1	1	1	0,5	0,5	0,6	0,7	0,8	0,8

Judgments about the plausibility of the signs accompanied by the certainty factors in the level and time of manifestation and degree of incidence can serve as a further basis for effective incensement of the suggested expert hypotheses system.

6 Reflection of Experts in the Evaluation of Fuzzy Signs

Personal features of cognitive style affect the quality of knowledge structures [21]. This should be taken into account, since the true expert never views his ideas about the subject domain as an axiom. He mentally analyzes the picture of the world he offers. In medicine, this is the significance of each of the signs, their specificity for this pathology, relevance for differential diagnosis, and much more. And cognitive scientists often observe elements of doubt among experts when they formulate their ideas about the subject domain. This is manifested in assessing the difference in situations in the proposed decisions based on hypotheses that are finally formed under the influence of the expert's reflexive system.

Reflection show at activity of expert-physicians in the process of extracting knowledge from him for creating an intelligent system for diagnostic and/or therapeutic decisions support. The greatest problems are estimates of indistinctly manifested states of the disease.

The results of the knowledge extraction sessions, including the reflexive composition of experts, can be presented in a formalized form [22]. The process of the intellectual activity of the expert as a whole, including the reflex expert, can be represented on the basis of a formula including a fuzzy variable, a confidence factor, relevance of the characteristics, and a temporal characteristic of the changes in the course of a disease.

7 Discussion

Nowadays computer systems are used in medical practice to diagnose hereditary diseases [5, 8, 23] based on the analysis of information databases, which does not allow to take into account the progression of the pathological process to the necessary extent. The basis of these systems is mainly databases containing various combinations of clinical signs (including post-diagnostic demonstration of photographs). Some systems are designed to demonstrate and analyze a large number of photographs, i.e. visual diagnosis of decision support system [24–28]. This is also a symbolic picture of the world. However, the presentation of photographs without a proper intellectual context does not allow the doctor to effectively solve the diagnostic problems of most orphan hereditary storage diseases, excluding certain congenital dysmorphisms. Inclusion of visual images (coarse facial features, a skull in the form of a shamrock, the shape of ears, and the like) at the stage of the formation of diagnostic hypotheses may allow for a new approach to the problem of recognizing orphan diseases characterized by specific external features. This was partially revealed in the DIAGEN [6] system, when it was discovered that doctors mistaken in determining, for example, the shape of the nose (beak-shaped, pear-shaped, plum-shaped, etc.).

Systems such as Human Phenotype Ontology allow one to focus on the phenotypic similarity between diseases [29] and on the semantic metrics to measure the phenotypic similarity between queries and hereditary diseases. However, the classification in the study on ontologies does not take into account the dynamic change of characteristics. To rank the expected diagnoses at ontology system used also a statistical model of diseases [30]. But the statistical approach allows diagnostics only for relatively common hereditary diseases.

This study is a continuation of the work on the creation of an intelligent decision support system for diagnostic rare diseases. The difficulties in identifying hereditary diseases, known from literary sources, imply that one can identify and assess the characteristics of clinical manifestations of orphan diseases in their age dynamics only with the help of the knowledge of medical experts. This is especially important, since early diagnosis is a prerequisite for the timely initiation of treatment and prevention of abnormalities from various organs and organism systems. But this is possible only if the clinical decision support system used at all stages of the development of children, beginning with the early neonatal period. The expert knowledge presented in the present work on the age-related manifestations of signs, accompanied by factors of confidence measure in their manifestation, severity and frequency of manifestations at different periods of life, is a condition for the formation of rules and logical conclusions of the knowledge base in the dynamics of the pathological process.

In contrast to other work the main attention of this study is paid to the formation of a field of knowledge taking into account the signs of the disease manifestation at different ages. The introduction of the three confidence factors allowed experts to reflect on their views, their confidence in the manifestation, and severity of signs at different ages and also in the frequency of symptom occurrence in each age group. A preliminary analysis of literary sources with the formation of textological cards provides a deeper subsequent analysis of clinical manifestations by experts. Thus, symptom assessments in

combination with the triple certainty factors will help to increase the effectiveness of expert systems for different situations caused by the dynamics of hereditary accumulation diseases.

8 Conclusion

Due to the lack of statistically reliable samples of patients with orphan diseases, which are necessary for the statistical analysis, knowledge-based systems should be considered as means of creating support diagnostic solutions. The expert certainty factors include three aspects: (1) in relation to the probability of sign manifestation at a certain age ("time characteristic" including negative values when it is not possible to manifest at certain age limits), (2) in relation to the severity of signs ("characteristic on the depth of pathology") and (3) in relation to the degree of incidence of signs with a certain pathology ("frequency response characteristic"). This approach ensures the three-dimensionality of the observed sigh evaluation for each of the studied age periods. The proposed "triune" system of certainty factors opens up new opportunities for knowledge bases of the expert systems. In addition, the knowledge of engineering must take into account the vagueness of definitions in medicine, the presence of holistic images and the reflection of doctors, which creates serious difficulties in the diagnosis of hereditary orphan diseases. The created expert system will be implemented on ontologies.

References

1. Schieppati, A., Henter, J.I., Daiana, E., Aperia, A.: Why rare diseases are an important medical and social issue. Lancet **371**(9629), 2039–2041 (2008)
2. Online Mendelian Inheritance in Man® An Online Catalog of Human Genes and Genetic Disorders. https://www.omim.org/. Accessed 30 Mar 2018
3. Ayme, S., Caraboenf, M., Gouvernet, J.: GENDIAG: a computer assisted syndrome identification system. Clin. Genet. **28**(5), 410–411 (1985)
4. Pitt, D.B., Bankier, A., Haan, E.A.: A visual verbal computer assisted syndrome identification system. Aust. Paediat. J. **21**(4), 306–307 (1985)
5. POSSUM web 2018. Pictures of Standard Syndromes and Undiagnosed Malformations. https://www.possum.net.au/. Accessed 15 Apr 2018
6. Kobrinsky, B., Kazantseva, L., Feldman, A., Veltishchev, J.: Computer diagnosis of hereditary childhood diseases. Med. Audit News **4**(1), 52–53 (1991)
7. Guest, S.S., Evans, C.D., Winter, R.M.: The online London dysmorphology database. Genet. Med. **5**(1), 207–212 (1999)
8. Baraitser, M., Winter, R.M.: London Dysmorphology Database, London Neurogenetics Database and Dysmorphology Photo Library on CD-ROM, 3rd edn. Oxford University Press, Oxford (2001)
9. Ayme, S.: Orphanet, an information site on rare diseases. Soins **672**, 46–47 (2003)
10. La Marca, G.: Lysosomals. In: Blau, N., Duran, M., Gibson, K.M., Vici, C.D. (eds.) Physician's Guide to the Diagnosis, Treatment, and Follow-Up of Inherited Metabolic Diseases, pp. 785–793. Springer, Heidelberg. (2014). https://doi.org/10.1007/978-3-642-40337-8_52

11. Kobrinskii, B.A.: Approaches to the construction of cognitive linguistic–image models of knowledge representation for medical intelligent systems. Sci. Tech. Inf. Process. **43**(5–6), 289–295 (2016)
12. Zadeh, L.A.: Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. Fuzzy Sets Syst. **90**(2), 111–127 (1997)
13. Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets Syst. **1**(1), 3–28 (1978)
14. Quinlan, J.R.: C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers Inc., San Mateo (1993)
15. Malik, D.S., Mordeson, J.N.: Fuzzy Discrete Structures. Physica-Verlag, New York (2000)
16. Petrovsky, A.B.: Multi-attribute classification of credit cardholders: multiset approach. Int. J. Manage. Decis. Mak. **7**(2/3), 166–179 (2006)
17. Torres, A., Nieto, J.J.: Fuzzy logic in medicine and bioinformatics. J. Biomed. Biotechnol. **2006**(2), 1–7 (2006)
18. Prasath, V., Lakshmi, N., Nathiya, M., Bharathan, N., Neetha, N.P.: A survey on the applications of fuzzy logic in medical diagnosis. Int. J. Sci. Eng. Res. **4**(4), 1199–1203 (2013)
19. Popper, K.: Knowledge and the Body-Mind Problem: In Defence of Interaction. Routledge, Abingdon (2013)
20. Shortliffe, E.H., Buchanan, B.G.: A model of inexact reasoning in medicine. In: Buchanan, B.G., Shortliffe, E.H. (eds.) Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project, pp. 233–262. Addison-Wesley Publishing Company, London (1984)
21. Gavrilova, T., Leshcheva, I.: The interplay of knowledge engineering and cognitive psychology: learning ontologies creating. Int. J. Knowl. Learn. **10**(2), 182–197 (2015)
22. Kobrinskii, B.: Expert reflection in the process of diagnosis of diseases at the extraction of knowledge. In: Proceedings of the IV International research conference "Information technologies in Science, Management, Social sphere and Medicine" (ITSMSSM 2017), vol. 72, pp. 321–323. Atlantis Press, Paris (2017)
23. Rat, A., Olry, A., Dhombres, F., Brandt, M.M., Urbero, B., Ayme, S.: Representation of rare diseases in health information systems: the orphanet approach to serve a wide range of end users. Hum. Mutat.: Var. Inf. Dis. Special Issue: Deep Phenotyp. Precis. Med. **33**(5), 803–808 (2012)
24. Hammond, P., et al.: 3D analysis of facial morphology. Am. J. Med. Genet. **126A**(4), 339–348 (2004)
25. Hammond, P., et al.: Discriminating power of localized three-dimensional facial morphology. Am. J. Hum. Genet. **77**(6), 999–1010 (2005)
26. Vardell, E., Bou-Crick, C.: VisualDx: a visual diagnostic decision support tool. Med. Ref. Serv. Q. **31**(4), 414–424 (2012)
27. Kuru, K., Niranjan, M., Tunca, Y., Osvank, E., Azim, T.: Biomedical visual data analysis to build an intelligent diagnostic decision support system in medical genetics. Artif. Intell. Med. **62**(2), 105–118 (2014)
28. Ferry, Q., et al.: Diagnostically relevant facial gestalt information from ordinary photos. eLife **3**, e02020 (2014). https://doi.org/10.7554/eLife.02020
29. Robinson, P.N., Köhler, S., Bauer, S., Seelow, D., Horn, D., Mundlos, S.: The human phenotype ontology: a tool for annotating and analyzing human hereditary disease. Am. J. Hum. Genet. **83**, 610–615 (2008)
30. Köhler, S., et al.: Clinical diagnostics in human genetics with semantic similarity searches in ontologies. Am. J. Hum. Genet. **4**(85), 457–464 (2009)