Analysis of the efficiency of physicians and students of medicine in the identification and classification of respiratory system auscultation sounds – a survey and practical research in Poland

Honorata Hafke-Dys[1,2], Anna Bręborowicz[3], Paweł Kleka[4], Jedrzej Kocinski[1,2], and Adam Biniakowski[2]

[1]Institute of Acoustics, Faculty of Physics, Adam Mickiewicz University in Poznań, Umultowska 85, 61-614 Poznań, Poland
[2]StethoMe™, MyWhizzy.com, Winogrady 18A, 61-663 Poznań, Poland
[3]Department of Pediatric Pneumonology, Allergology and Clinical Immunology, K. Jonscher Clinical Hospital in Poznań, Poznań University of Medical Sciences, Poland, Szpitalna 27/33, 60-572 Poznań
[4]Institute of Psychology, Adam Mickiewicz University, Wieniawskiego 1, 60-712 Poznań, Poland

March 14, 2018

Abstract

Rationale: Auscultation is the first examination of a patient in GP's office. It is totally subjective and depends on physician's abilities to interpret the sounds based on their psychoacoustical features. **Objectives:** A cross-sectional assessment of the skills of physicians and medical students in classification of respiratory system sounds in children is presented here. **Methods:** The experiment comprised 24 respiratory system sounds with different phenomena. 185 participants took part in the experiment. **Results:** We revealed difficulties in both recognition and description of respiratory sounds. The results significantly improved when sound classes were grouped to form more general ones. **Conclusions:** We confirm that this is a global problem which cannot be neglected and may result in ambiguities in diagnosis and mistreatment. Moreover the problem of insufficient training both during study and during medical practice is also highlighted here. There is also a perceived global need to standardize the nomenclature of auscultation sounds.

Introduction

Nowadays auscultation is still the first and most common examination carried out by every general practitioner (GP) or family doctor. Most of the diagnoses there are made based on it. It is fast, easy and does not need advanced technology. Nevertheless, one must keep in mind that auscultation is a medical examination that has been known since the time of Hippocrates. Its popularity was made possible thanks to Leannec, who invented the stethoscope in 1816. His invention was a rigid, cylindrical-shaped wooden headpiece with a central groove (as a sound tube) for listening to the lungs and heart. Thanks to this solution the doctor no longer had to put his ear to the patient's body, and examination became simpler and more hygienic. Apart from some minor changes, the form of the acoustic stethoscope has survived to this day. One modification was the introduction of bowls on both sides to adjust acoustic impedance, thereby increasing the volume. The British physician Bird (1840) introduced an elastic tube that connected the stethoscope chestpiece with a single earpiece. Since 1964 (Littmann, 1964) the stethoscope has also been equipped with an acoustic diaphragm. There are also various types of electronic stethoscopes nowadays, which operate different in different ways. They have a transducer (a piezo microphone) that converts the vibration on the chest surface to an electrical signal that is then reproduced to the physician via headphones and can be recorded for later analysis. Regardless of the type of stethoscope, the most important features of auscultation are non-invasiveness, simplicity, ease to carry out, and the low cost associated with the device. However, the results of this examination are subject to a relatively large rate of error due to the subjectivity of the assessment and the influence of many additional factors. First among these are the experience of doctors and their perceptual abilities. Physicians often differ in their assessments. Mangione and Nieman (1999) examined the pulmonary auscultatory skills of medical students, pulmonologists and interns in internal medicine and family practice. Research has shown that internal medicine and family practice doctors have not been statistically superior to medical students. Only pulmonologists have achieved statically better results than others.

Furthermore, differences also arise from the different nomenclature used and the division of different sound classes in the existing medical literature. This is an international problem, as has been shown by numerous authors (Bunin et al., 1979; Cugell, 1987; Francis et al., 2013; Wilkins et al., 1990). Pasterkamp et al. (2016), proposed a unified nomenclature for 6 languages and suggested a standardized terminology. Unfortunately, it is not wccepted worldwide in education nor everyday practice. As a consequence, there is still no uniformly standardized classification of the types of phenomena characteristic for the human respiratory system in the whole medical environment. Depending on the handbook and university, physicians often use other words with a completely different semantic meaning to describe the perceived pathology in the respiratory system. This leads to problems in preliminary auscultation courses, but primarily during later professional work, when doctors exchange or consult diagnoses, as they use different descriptions of the same sounds and semantically similar or even identical terms to describe different types of phenomena, which results in ambiguous descriptions or even makes the descriptions incomprehensible to other doctors, and the examination must be repeated. It is also worth noting that the sounds are not stored and there is no way to go back to the recordings or compare them to other sounds, so there is no way of verifying the description, which increases the ambiguity even more.

Aim

The goal of this work is to answer a few problems. First, the questionnaire was ascertain whether the medical community is aware of the problem of auscultation sounds classification, how they evaluate education in the area of auscultation, and how they assesses their abilities in it. The main goal of this paper, however, is to answer the question: How correctly and consistently do physicians and medical students evaluate respiratory sounds, and do they categorize them in the same way? In this context, other aspects of the problem have also been analyzed in details: Do pulmonologists perform better than other groups of physicians? How does a group of students fit into this community? And finally we try to conclude how to solve the problem of low efficiency of doctors by creating respiratory signal database and unifying nomenclature. This problem is especially crucial because most of patients that feel ill go to the doctor's office where they are examined by GPs or family doctors and the diagnosis and further treatment are mainly based on auscultation. The study was approved by the Bioethical Commision of the Poznań University of Medical Science.

Method

The test was distributed among the academic medical community and in hospitals. It contained both signals and question about specialization of the participants. The test was anonymous and was conducted online via the Internet using Questionpro Professional. This program was chosen because enables uncompressed and high quality audio to be presented. Moreover, before the experiment was made available to the participants, the quality of the signals in this software was subjectively verified by two experienced acousticians (sound engineers) independently (without hearing loss) in terms of distortions and possible artifacts. No difference was found between direct and on-line listening. In addition, to minimize the possibility of a layperson completing the survey, the survey was distributed among the academic medical community and in hospitals. It was also passed on to interested people through lecturers at medical universities and through direct contact with doctors asking them to complete the test.

The experiment consisted of two parts.

Part I: Survey

In this part, each participant responded to a number of questions regarding:

- education, the specialization started or held,

- assessment of their own skills in adult and child auscultation,

- type of stethoscope (electronic / analog) and frequency of auscultation in their medical practice,

- opinion on scale a 5-grade on the number of hours devoted to studying auscultation during their study and specialization, the need for additional training, and the scale of the problem of ambiguity in the nomenclature used in the classification of auscultation sounds.

Part II: Classification of hearing sounds

This part consisted of 24 sounds (see Tab. 1 for details) that the test participant listened to (they could replay each sound), evaluated and assigned to specific classes (details below). Nine sounds were selected from the demonstration recordings included on a CD in the Fundamentals of Lung and Heart Sounds (Willkins et al., 2004). The rest were recorded with the Littmann 3200 electronic stethoscope. These were the records of the respiratory sounds of children aged 5 months to 14 years (average 7.6 years).

The choice of sounds for the test was two-step. First, sounds from a database of over 2000 sounds were selected by acousticians. Those sounds were chosen because they contained the smallest number of distortions and artifacts. The artifacts of Littmann 3200 stethoscopes appear mainly when the chestpiece is moved and during the application or deposition of the chestpiece from the body. These disturbances are mainly caused by the acousto-electric transducer. The other equally important criterion was the choice of sound class to make the sounds as diverse as possible, but at the same time unambiguous. This way 50 different sound examples were selected. The final selection of sounds was made by the team of specialists composed of eight experienced pediatricians and pulmonologists working at the Karol Jonscher Clinical Hospital in Poznań, Karol Marcinkowski Poznań University of Medical Sciences. At the meeting of those physicians, the Fostex PH-50 headset coupled with high quality professional headphones (Sennheiser HD600) enabled simultaneous listening of sound samples to all physicians. After listening to each signal, they classified the sound. Then there was a discussion about it and a common position was held. Finally, only the sounds which no physicians had any doubt over were chosen for the test. The set of those signals with their descriptions are called "standard" here.

The sounds represented certain classes (Table 1) and were presented to the participants as a collection of signals from 24 different patients. The terms used in the classes were chosen to include that the entire spectrum of nomenclature that is used in modern medical literature (Szczeklik i Szczeklik, 1979, Rowińska-Zakrzewska and Kus, 2004, Willkins et al., 2004, Mangione and Nieman 1999, Pasterkamp et al., 2016) , and thus by physicians in their daily practice. During the experiment more than one class could be assigned to each sound by a listener. Additionally, with each sound there was information about the place on the chest that the chestpiece of the stethoscope was placed on during recording, as well as the age, height and weight of the patient. Each participant was informed about the need for high quality headphones. No feedback was used, which means that the participants were not informed about the correctness of their response during the test. Moreover, the order of presentation was randomly chosen, but it was the same for all participants. As a result, the participants could not learn the correct answers and patterns as they performed the listening task.

the sound class described by the team of medical specialists (standard)	number of signals in test
vesicular breath sound	3
normal bronchial sound	2
abnormal bronchial sound	1
louder breath sound, prolonged expiratory phase and rhonchi	2
fine crackles	2
fine crackles and crepitus	1
fine crackles and abnormal bronchial sound	1
medium crackles	1
coarse crackles	2
crepitus	1
rhonchi and expiratory wheezes	1
rhonchi	1
stridor	1
expiratory wheezes	3
inspiratory wheezes and rhonchi	1
inspiratory and expiratory wheezes	1

Table 1: **The sound classes that were presented in the test (column 1) along with the number of sounds of particular category present in the test (column 2)**

Results

Participants

Both parts of the experiment were completed by 205 speaking participants. In the first step, the results from the questionnaires were analyzed for their quality. The average time for completing the questionnaires was 25 min. The total duration of the signals, excluding the survey part, was about 4 minutes. Therefore, the responses of the participants who completed the entire experiment in less than 5 minutes were rejected as unreliable. Therefore,185 questionnaires were approved for further analysis. Among those participants there were: 16 pulmonologists, 22 pediatricians, 29 doctors with other specializations, 50 doctors during the internship and 68 medical students.

Part I: Survey

Among respondents, the majority (67%) auscultates patients at least once a week, 14% do it at least once a month. The remaining group auscultates their patients less than once a month (19%). Only 4.9% of surveyed physicians use an electronic stethoscope in their daily practice. Fig. 1 shows how the physicians rated their skills in the auscultation of children and adults.

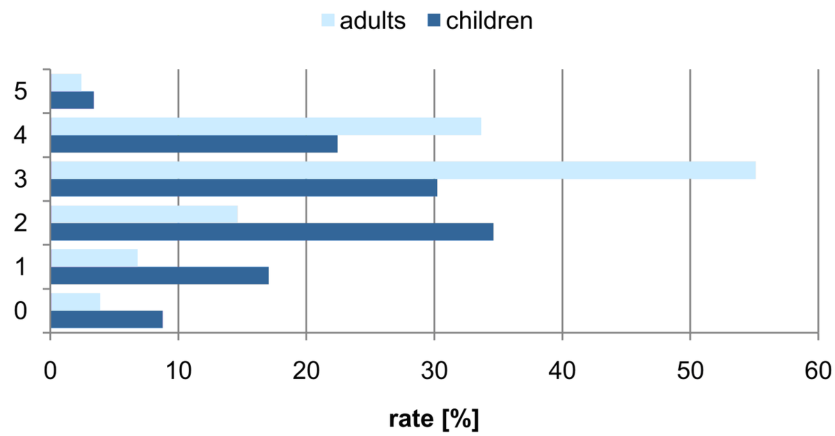

Figure 1: **The percentage distribution of responses to the question of assessing one's own skills (on a scale of 0-"very poorly" to 5 = "very well") for the auscultation of adult and child respiratory systems.**

Figs. 2a and 2b show the distribution of answers to questions about the number of hours devoted to the auscultation technique during their study program, internship and specialization. 64% of the respondents strongly agree or agree that the number of hours of instruction during their studies is insufficient. Only 11.3% of respondents disagree with this statement (strongly disagree and do not agree). The evaluation of training in auscultation during internships and specialization is more positive (Fig. 2b). Here 38.1% of the respondents strongly agree or agree that the number of hours devoted to learning of auscultation is low, while 37.7% partially agree, and 24.3% disagree or strongly disagree with this statement.

On the other hand, the medical community sees the need for additional auscultation training: more than 65% (strongly agree or agree) would like to improve their skills in this field regardless of their specialization (Fig.2c).

It must be emphasized that the medical community also sees an enormous problem in the ambiguous nomenclature used in the classification of respiratory sounds. More than 65% of respondents strongly agree or agree, and only 14.2% disagree or strongly disagree with the statement that the classification of respiratory sounds requires coherence and uniformity (Fig. 2d).

Part II: Classification of auscultation sounds

This part was related to the auscultation of the recorded sounds from the respiratory system and the classification of those sounds by the participants. First, the results of this experiment were divided according to compliance of the sound classification carried out by the participants with the standard developed by the team of medical specialists. For each sound presented to the participants, the responses were categorized according to Table 2. For each sound, the percentage distribution of each response category was determined (Fig. 3). The chart shows that doctors classify sounds with varying accuracy. In general, after averaging the responses for all the sounds it can be seen that there is 14.6% CSs, 26.2% PSs, and 59.3% (N), respectively.

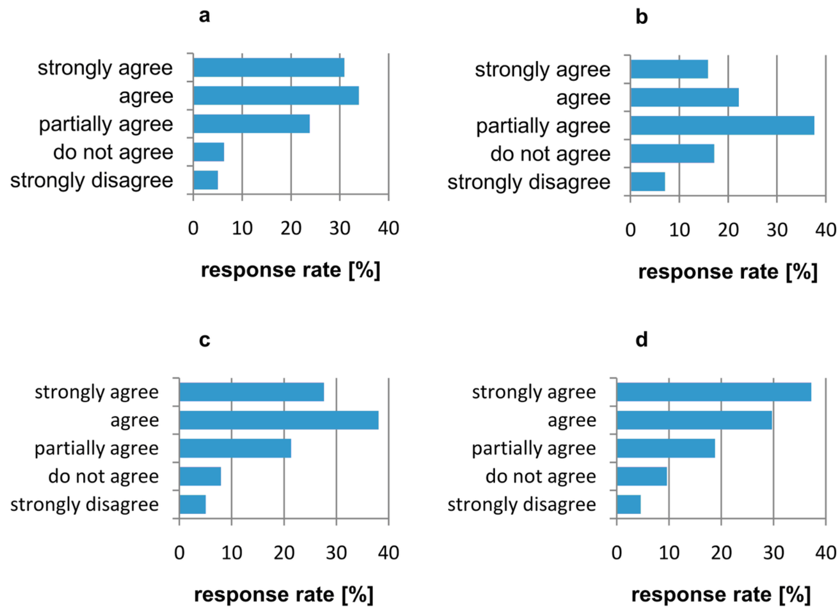

Figure 2: **(a). The percentage distribution of responses to the questions: (a) "To what extent do you agree with the statement: The number of hours spent on respiratory system auscultation training is insufficient.", (b) "To what extent do you agree with the statement: "The number of hours for respiratory system auscultation training during internship and specialization is insufficient for me." ,(c) "To what extent do you agree with the statement: Additional training in respiratory system auscultation is needed for the doctors of my specialization.", (d) "To what extent do you agree with the statement: The names associated with the auscultatory sounds of the respiratory system are inconsistent and need to be unified."**

abbreviation	response category	description of category
CS	the answer was compliant with the standard	full compliance with the standard and no incorrect answers
PS	the answer was partially compliant with the standard	at least one correct answer marked; additional phenomena that were not included in the standard were marked
N	incorrect answer	among the marked phenomena, the response from the standard was not marked

Table 2: **Assumptions used to evaluate the responses of participants and the abbreviation used.**

The most inaccurate diagnoses were statistically significant (chi2(2) = 59.8, p <0.001). Furthermore, if physicians are grouped according to their specialization, it can be seen from the Kruskal-Wallis test that there are statistical differences also between the groups of physicians of different specializations in the category CS (chi2(4) = 16.0 p <0.003). Pair analysis of the differences in answers between groups of physicians with different specializations using post hoc tests with a Holm correction showed that pulmonologists have statistically higher scores than students (p = 0.023) and interns (p = 0.060). In the case of correct answers, there were also statistically significant differences between specializations (chi2(4) = 16.8, p = 0.002). The analysis also showed that statistically students had significantly lower scores than pulmonologists (p = 0.022), pediatricians (p = 0.022) and other specializations (p = 0.073). For incorrect answers, the statistically significant differences (chi2 (4) = 69.9, p <0.001) were related to the lower scores obtained by students and interns.

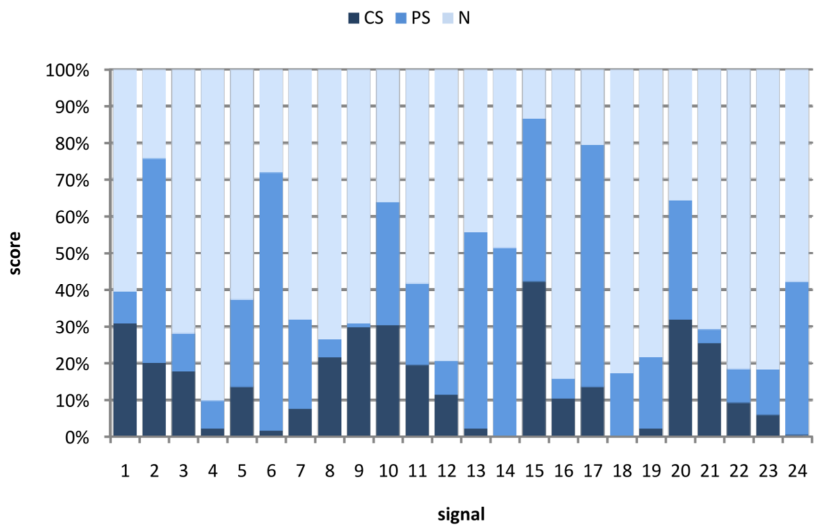

Figure 3: **The distribution of standard compliant (CS), partially compliant (PS) and incorrect (N) answers.**

In the next step, the sounds were grouped according to their class and the percentage of correct responses (P), depending on the medical specialization (Fig. 4). To give a general view, it was assumed that the correct answer category (P) contains both the standard compliant responses (CS) and the partially compliant responses (PS).

Analyzing false positives (defined as cases of indicating a sound class when this class was not indicated in the standard- this is the sum of all the incorrect selections that were not in the given sample, with the individual sound classes being treated independently of each other) the graph shows no differences between groups, except for the prolonged expiratory phase - the only sound class with statistically significant differences (chi2 (4) = 13.1, p = 0.011). This difference can be found in the results obtained by pulmonologists and students.

In the next step, the responses were grouped using the main classes proposed by the European Respiratory Society (ERS) (Pasterkamp et.al, 2016). In practice, this meant that if the participant had indicated any of the subgroups shown in Table 3 and any of them was marked as the correct one in the standard, the answer was treated as a correct one. In the case of the vesicular breath sound or the bronchical one, no other answer that is not part of the subgroup could be concurrently marked. In this way, five main classes of sounds were created. After the grouping described above, a graph of the correlation between the correct answers for each main class was obtained (Fig. 5). As one can see, for all the classes associated with pathological signals, the pulmonologist group is the most effective one. However, statistically significant differences based on the Kruskal-Wallis test and a comparison of the groups of participants in pairs were obtained only for the class of rhonchi between the group of pulmonologists who scored higher than pediatricians (p = 0.085), other physicians (p = 0.046), interns (p = 0.085) and medical students (p = 0.015). For wheezes, pulmonologists had statistically significantly higher scores than other specializations (p = 0.008). Also interns had better

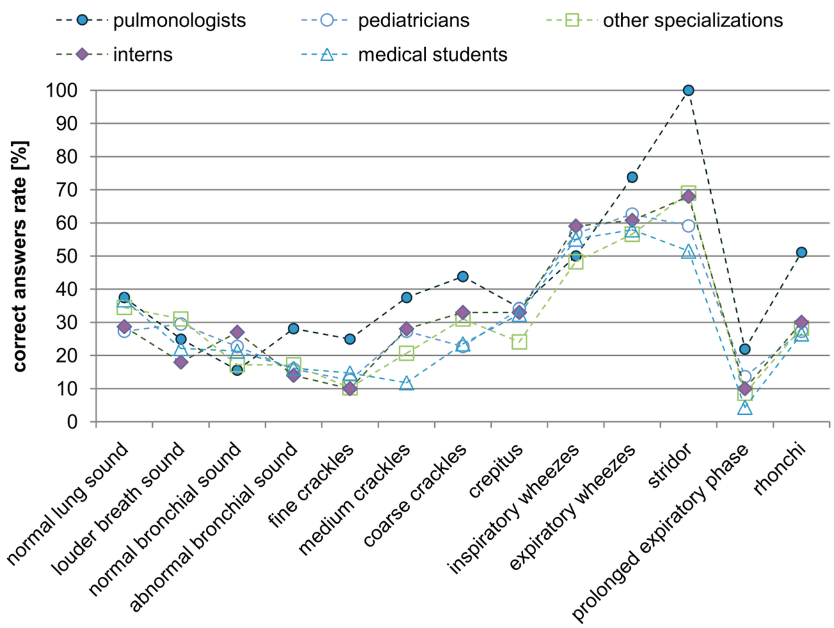

Figure 4: **The juxtaposition of correct (P=CS[?]PS) answers for each sound class for doctors of different specializations and medical students. The differences for individual sound classes were analyzed based on a Kruskal-Wallis test and post hoc tests with a Holm correction for statistically significant differences. According to the analysis, pulmonologists distinguish the sound classes in a better way when compared to the other groups - as many as three classes are significantly recognized differently by the group of pulmonologists relative to the other specializations: fine crackles (chi2(4) = 11.9, p = 0.018), stridor (chi2 = 14.4, p = 0.006) and rhonchi (chi2 (4) = 11.1, p = 0.026)**

main class	subclasses included in the main class
breath sound / bronchial sound	vesicular breath sound
	diminished breath sound
	louder breath sound
	normal bronchial sound
abnormal bronchial sound	no subclasses
crackles	fine crackles
	medium crackles
	coarse crackles
	crepitus
wheezes	inspiratory wheezes
	expiratory wheezes
	stridor
rhonchi	no subclasses

Table 3: **Main classes of respiratory sounds and subclasses that are part of them** .

scores than physicians of other specializations (p = 0.026). Generally, it can be stated that this grouping highlighted the advantage of pulmonologists over the rest of the groups in the correct recognition of the respiratory sounds phenomena.

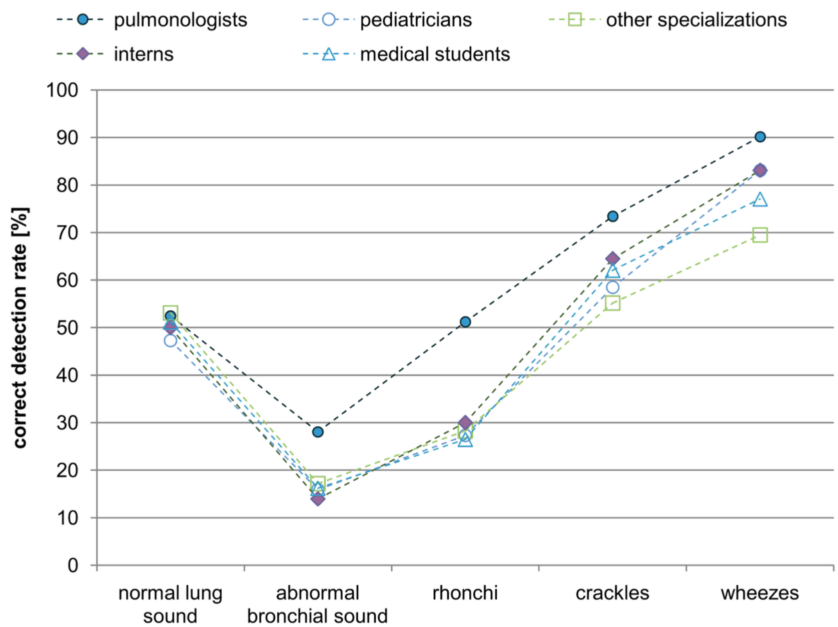

Figure 5: **The percentage of correct detection for grouped sound classes (main classes) for physicians of various specializations. To emphasize the differences, the points for individual groups are connected by lines.**

In the test, both the reference sounds included on the CD of the physician handbook (Wilkins et al., 2004) and real-life sounds recorded using the Littmann 3200 stethoscope were used. The effect of the type of sound recording on the results was analyzed. The results for correct, partially correct, and incorrect detections were compared, taking into account the specialization of the participants. Based on the results of the Wilcoxon test, no significant difference in the percentage of any type of diagnoses between the two kinds of recording (CD and Littmann 3200) was observed in any of the participants ($p > 0.05$). This means that the type of recording did not affect the results obtained by the participants, and they perform similarly with the classification of sounds in the case of different origins.

In the next step, we analyzed what phenomena other than the standard ones were marked by subjects. It should be emphasized that the purpose of this analysis was first of all to draw attention to the grouping of classes, which may be the first step in further unification of the nomenclature of sounds. It was assumed that phenomena that are most similar would also most often be confused with each other. Therefore, the obtained relationships were grouped by classes and the analysis of those are depicted, namely: crackle class (Fig. 6), breath sounds (vesicular and louder) and bronchial sounds (Fig. 7), wheezes and rhonci (Fig. 8). This approach is in line with the assumptions of Pasterkamp et al. (2016) for other languages.

In the case of the juxtaposition of coarse, medium, fine crackles and crepitus (Fig. 6), it can be seen that these are classes that are confused. Coarse crackles are most often confused with other types of crackles and crepitus, and even rhonchi. In the medium crackle class the results are similar to those for coarse crackles, but the most noticeable additional class marked for this phenomenon was crepitus.

In turn, fine crackles are not confused so often with rhonchi and medium and coarse crackles, but are often confused with the crepitus category, which is a class that is much more often marked than that described by

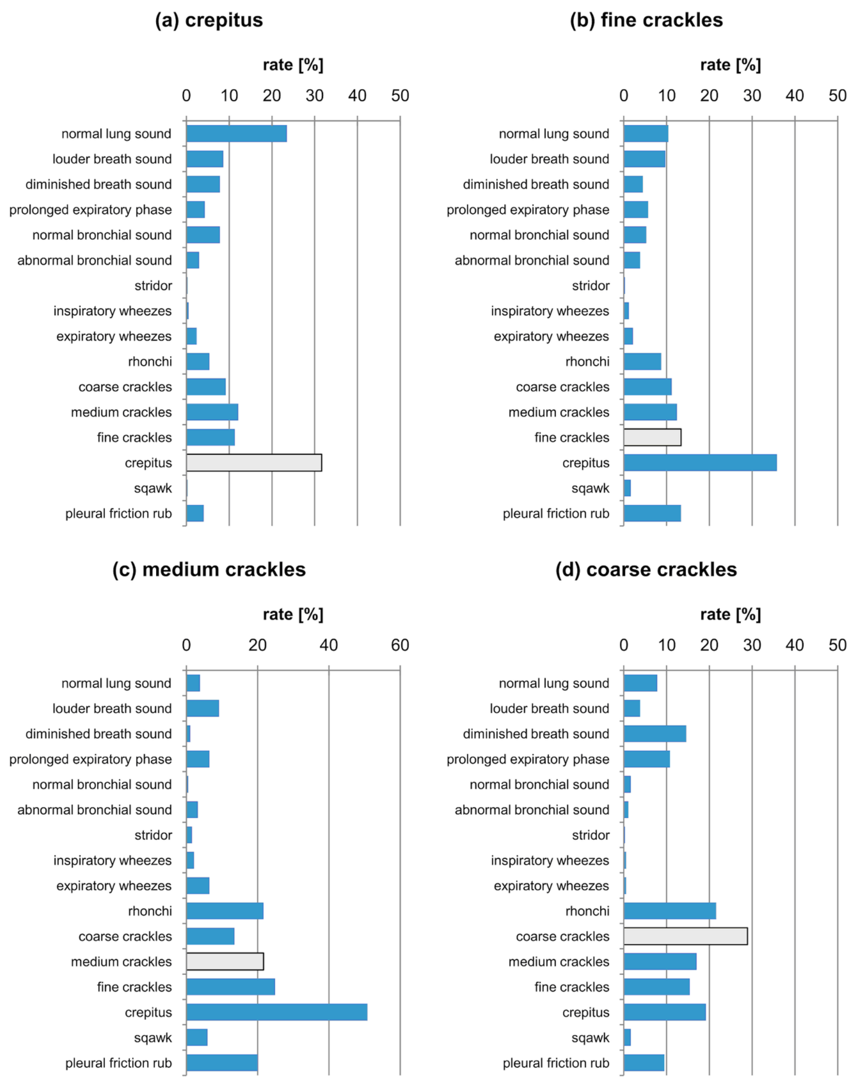

Figure 6: **The percentage distribution of responses marked for the classes (a) crepitus (b) fine crackles (c) medium crackles (d) coarse crackles. The light grey bar depicts the correct answer.**

the standard (fine crackles). This is a situation that seems natural, because in many medical communities worldwide crepitus is equivalent to fine crackles, or is a subgroup for this class (Szczeklik, 1979). In the case of crepitus, the most common category marked was normal lung sound, which may mean that the respondents did not notice the sound class and treated the extra sound as an artifact occurring during the recording. This is confirmed by Fig. 8b, in which the answers other than those from the standard are shown for the normal lung sound, which was the correct (standard) response. It can be seen that the respondents, in addition to the correct answer, most often marked louder breath sound and crepitus. In the case of normal bronchial sound (Fig. 7a), the louder breath sound was more frequent than the standard one. In addition, the respondents also often noted a prolonged expiration phase.

In the samples that had louder breath sound (Fig. 7c), the subjects often recognized the respiratory pathology classes - inspiratory and expiratory wheezes or rhonchi. In the case of normal sounds, the selection of these classes was not observed.

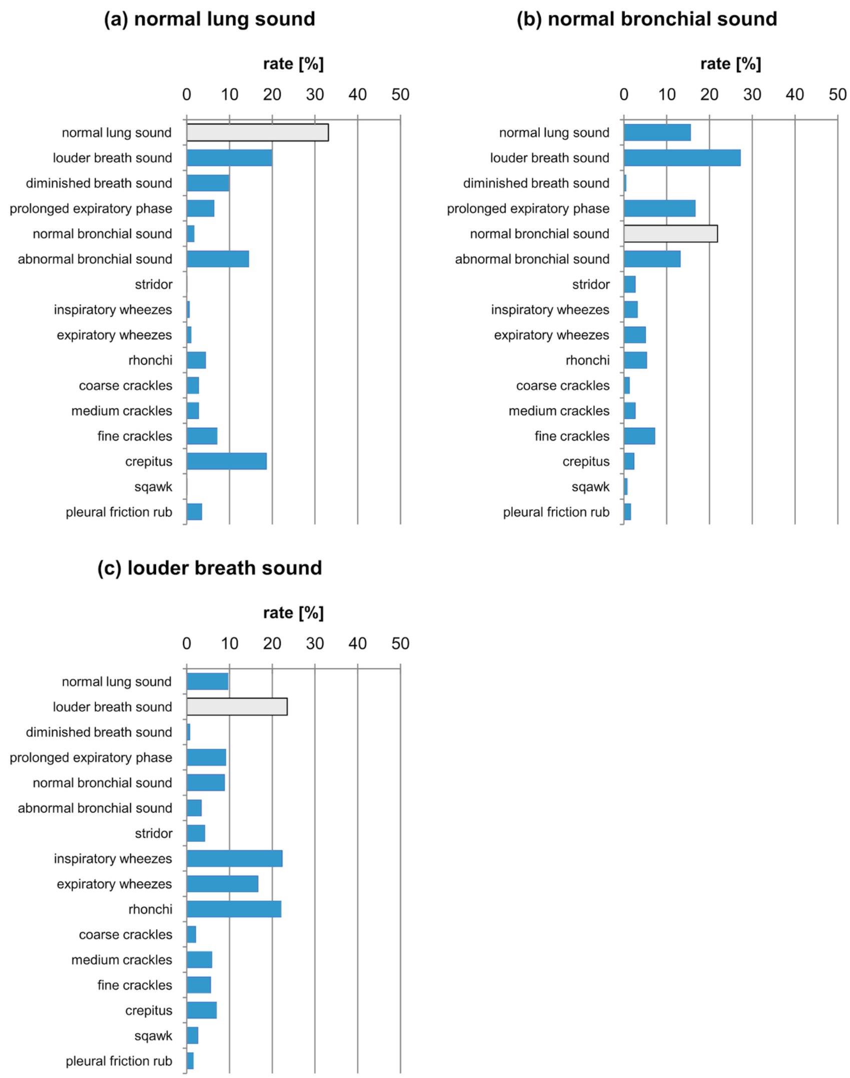

Figure 7: **The percentage distribution of responses marked for the classes (a) normal bronchial sound; (b) normal lung sound; (c) louder breath sound. The light grey bar depicts the correct answer, n is the number of this kind in the survey.**

In the case of the juxtaposition of inspiratory and expiratory wheezes (Fig. 8), it can be noticed that they are generally recognized correctly, i.e. the percentage of those answers is the highest, so those are the most likely answers. Most often they were confused, i.e. inspiratory wheezes were confused with expiratory wheezes and vice versa. In addition, only in the case of inspiratory wheezes, was rhonchi marked as an additional class. This means that in the case of evaluating previously recorded sounds physicians have difficulty identifying the inspiratory and expiratory phases.

In the case of rhonchi (Fig. 8c), it can be seen that this is the class closest to the wheezes and it is most often confused with them. The class of inspiratory and expiratory wheezes is marked almost as often as the

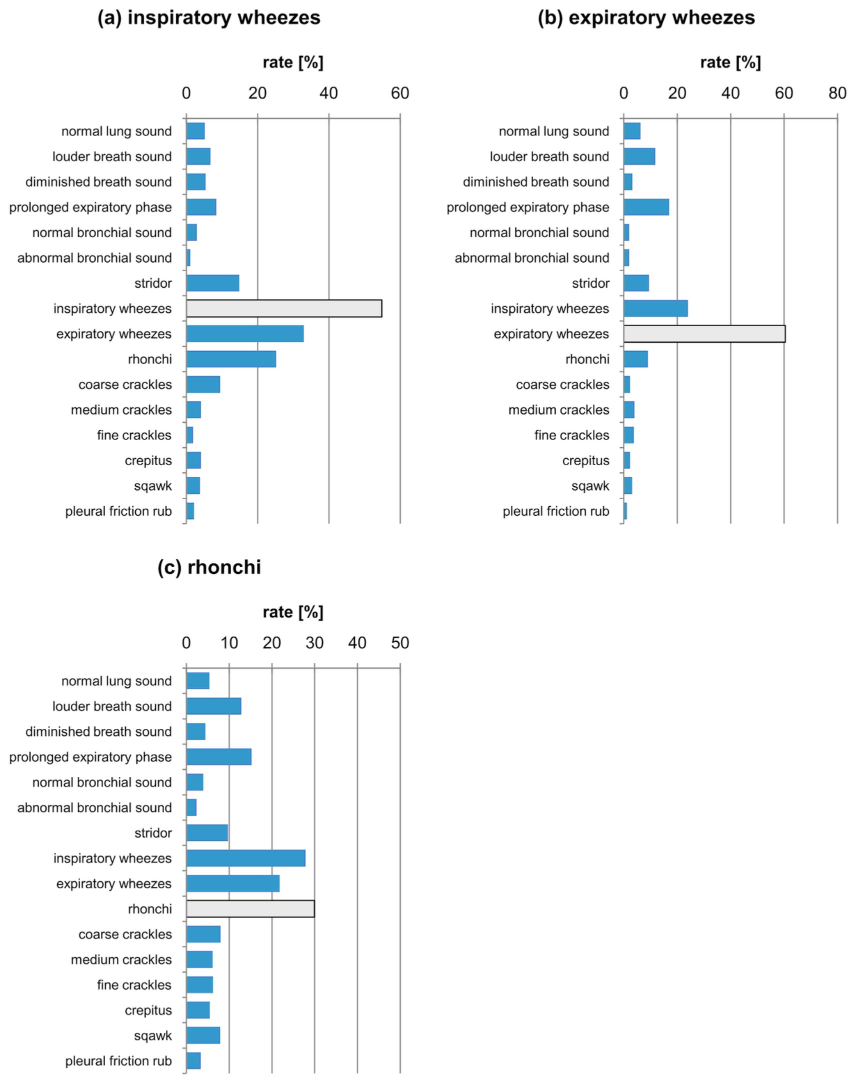

Figure 8: **The percentage distribution of responses marked for (a) inspiratory wheezes, (b) expiratory wheezes, (c) rhonchi class. The light grey bar depicts the correct answer, n is the number of this kind in the survey**

rhonchi class, which was the correct answer here. Additionally, for the rhonchi class a louder breath sound and prolonged expiration phase were also often marked.

In general, comparing the results in Figs. 8a and 8c, it can be concluded that wheezes and rhonchi groups are subjectively close to each other and are often confused with similar acoustic features. Fig. 8c. shows that in sound samples which contained rhonchi, according to the standard, the respondents mistook them mainly for the class of wheezes. At the same time, it can be seen that the class of rhonchi was also often chosen when the standard response was medium or coarse crackles, and thus the rhonchi class possesses the acoustic features characteristic for the crackles class, causing confusion among the participants. Finally, it can be stated that the class of rhonchi is quite ambiguous and differently classified by the respondents due to the fact that it has the features of both wheezes and crackles.

Conclusions

Auscultation is an important part of a medical examination. It allows rapid screening as an essential step in practically every visit. From the analysis of the data described in this paper, it can be stated that the survey results clearly indicate the need for more training at different stages of learning and medical practice in this area. They also indicate the awareness of the medical community concerning the lack of a unified nomenclature of respiratory sounds. The results of the questionnaire are confirmed by the results obtained in the test section, in which the participants of the experiment classified the previously recorded respiratory sounds. These results were compared with the standard developed by the team of specialist physicians. The average number of correct detections of auscultation phenomena in the test (not taking into account the sound classes) is only from 24.1% for students to 36.5% for pulmonologists. It should be borne in mind that correct detection means that additional phenomena may have been identified which, according to the standard, were not present in the sound sample (P). If one looks more restrictively at the results and takes into account only those responses that are in full compliance with the standard, the average number of responses compliant with the standard (CS) is just 14.7% for all groups. If one splits the results into detailed and main sound classes, it can be noticed that the highest number of correct answers was obtained for the wheeze classes. This is confirmed by both the correct answers received by participants of different professions and students (Fig. 4 and Fig. 5), as well as the lowest number of false positives noticed for this class (Fig. 6). This is not surprising, since they are the most characteristic sounds: continuous, tonal, periodic and relatively loud, e.g. in the case of stridor, a characteristic type of inspiratory wheeze, 100% correct answers for pulmonologists was found. More problematic are the classes of crackles and rhonchi. For crackles, in the case of detailed, coarse, medium and fine crackle classes, the number of correct detections was relatively low and for the best group of physicians (pulmonologists) the number ranged between 40% of correct responses, depending on the type of crackles. This is comparable and consistent with the results obtained by US doctors (Mangione and Nieman, 1999). If all types of classes along with crackles are grouped together (i.e. each of the answers for the class of crackles: fine, medium, coarse crackles and crepitus are considered to be correct, regardless of which one was correct according to the standard), then the result is significantly improved and reaches 73.4% of the correct answers for the group of pulmonologists. For the remaining groups, this value is around 60% of the correct answers. In the case of rhonchi, which are the most capacious class of respiratory sounds, the advantage of pulmonologists is clearly visible. In their case, the number of correct detections is 51.2%, while for other groups this value does not exceed 30%. The lowest number of correct indications is visible for the breath sound. The results show that few physicians can unequivocally assess the appearance of a louder breath sound, prolongation of the expiratory phase or diminished breath sound. These are phenomena whose unambiguous analysis would require more recordings, from different recording points, and preferably the reference sound recorded in a healthy patient.

For all classes of sounds, one can see the advantage of the pulmonologist group in terms of their correct answers, which is also consistent with the results of Mangione and Nieman (1999). This shows indirectly that it is possible to improve the skills of doctors in this field, through training and practice.

The comparison of the results of the correct answers for the detailed classification (Fig. 4) and the classification with the grouped (main) classes (Fig. 5) shows that this grouping significantly increases the percentage of correct responses. It must be stressed that this grouping results from the fact that if doctors are mistaken in the detailed classification, they usually confuse the subgroups of a given class, as is also shown in Figs. 6-8. This can be interpreted on the one hand as the possibility of the semantic ambiguity of classes (nomenclature), and on the other hand as a too detailed division, which due to the acoustic characteristics of distinctive features of sounds is unrealistic, because doctors in most cases are unable to distinguish these sounds using their hearing only.

Analyzes of the sound classes that the respondents marked, together with the correct answer, show how they grouped the individual auscultation phenomena into the main sound classes. It can be seen that in the case of normal sounds the category of crackles often appears in the supplementary answers. This is most likely related to the quality of the recording. With the use of piezoelectric transducers in the Litmann

3200 electronic stethoscope, it is not possible to eliminate artifacts associated with moving the stethoscope chestpiece. These artifacts are akin to crackles and are in most cases perceptually indistinguishable from pathological sounds. In the case of the sounds chosen for the test, these samples were almost devoid of these artifacts, but as the results show, some people perceived silent sounds resembling crackles and qualified them as pathological sounds. This problem seems particularly important when listening to a remote patient, where the doctor is unable to verify whether the sound is an artifact, or if it is a cause for concern. The technical solution to this problem is to use a microphone instead of a piezo transducer.

The remaining results from the analysis of additional phenomena indicate that the responses group sounds to some general classes and this is consistent with the standardized nomenclature (Pasterkamp et al., 2016). On this basis one can distinguish three main classes of sounds: crackles (which are divided into classes of fine, medium and coarse), wheezes and rhonchi. Then in these categories physicians (mainly pulmonologists) may try to distinguish specific, detailed classes. In the case of the crackles and crepitus class, it is noteworthy that, for nomenclature and semantic meaning among physicians, the class of fine crackles and the class of crepitus are of equal quality and often confused with each other (when the correct response is fine crackles, the crepitus response is more often chosen (Fig. 6d). In turn, the class of coarse crackles (Fig. 6b) is slightly more confused with the class of rhonchi than with the other subtypes of crackles, which also indicates the difficulty in finding the acoustically distinctive features of these classes. Consequently, it can be concluded that acoustically and semantically they are often referred to the same class by physicians. The medium crackles seems unimportant, as they are confused with both crepitus and rhonchi, and the number of correct answers is half the incorrect answers (this class is classified as crepitus). In general, based on the analysis of the results presented in Fig. 8, it can be concluded that in the future with this kind of phenomena it would be good idea to use the classification according to Pasterkamp et al. (2016), which is comprised of only the following classes: fine crackles, crepitus and coarse crackles.

The most ambiguous is the class of rhonchi, which is at the boundary between the class of wheezes and the class of medium and coarse crackles. Rhonchi, like wheezes, are continuous and periodic sounds. However, their fundamental frequency is lower than that of wheezes. They are often mistaken for them (Fig. 8c). The formation of rhonchi is associated with the movements of secretions in the respiratory pathways, which often results in a stertorous, intermittent sound that may be confused with coarse and medium crackles (as shown in Figs. 8a and 8b). This is consistent with the results obtained by Willkins etal. (1990), who also showed that rhonchi are an ambiguous class, very often incorrectly identified.

As has already been mentioned, Fig. 8 also shows that wheezes are the most homogeneous class, and in the case of samples containing this phenomenon as a correct answer described by the standard they were mainly confused with respect to the respiratory cycle (inhale / exhale).

In conclusion, it can be considered that the effectiveness of physicians in the clear classification of auscultation sounds with the use of detailed sound classes is low and very heterogeneous. It seems that in the case of a screening test, which can be assumed for a respiratory examination with a stethoscope, the standardized nomenclature proposed by Pasterkamp et al. (2016) is sufficiently precise and should provide the basis for a standardized classification of sounds for every language and further detailed respiratory diagnostics. This kind of classification should be also standardized and taught during medical studies and scholarships as a globally unified and unequivocal theory with the same semantic meaning. This approach would eliminate misunderstanding between physicians which is very common even when they speak the same language.

It should also be emphasized that due to the ever-growing market of electronic stethoscopes, it is possible to record normal and pathological respiratory sounds.

Accordingly the results of the study prove that the ability of detection of different auscultation phenomena can be significantly improved by training (which is made in everyday practice by pulmonologists who obtained higher scores). Moreover the survey across the participants suggests that there is a strong need to have more training during study and further practice. It must be emphasized that this problem is a global one which was also mentioned by other authors. It seems reasonable to create a worldwide database with auscultation signals described by the best specialists that will be used during education. Additionally the possibility of recording and storage of signals during the examination may be a good verification of the quality of this subjective procedure and may lead to increase in objectivity. This is highlighted also by the results that suggest that physicians (except pulmonologists), in general, are not better than medical students. This confirms the need of practical education not only during studies but also during further medical practice.

Acknowledgements

This work was supported by the grant from The National Centre of Research and Development (NCRD) in Poland and European Union under Sub-Measure 1.1.1 of the Operational Programme Smart Growth (PO IR) "Industrial R&D carried out by enterprises" (Fast Track), agreement no. POIR.01.01.01-00- 0528/16-00.The Authors would like to thank the Management of K. Jonscher Clinical Hospital in Poznań, Poznań University of Medical Sciences and the physicians from Department of Pediatric Pneumonology, Allergology and Clinical Immunology of this hospital for their help in recording and description of the acoustic signals. We also thank Angelika Podoła for her help in survey collection.

References

Bunin, N.J., Loudon, R.G. "Lung sound terminology in case reports". Chest 1979; 76: 690–692.

Burrows, B., Pulmonary Terms and Symbols: A Report of the ACCP-ATS Joint Committee on Pulmonary Nomenclature, Chest, Volume 67, Issue 5, May 1975, Pages 583-593, American Thoracic Society Committee on Pulmonary Nomenclature. Pulmonary Nomenclature. American Thoracic Society News. 1977;3:6.

Cugell, D.W. Lung sound nomenclature. Am Rev Respir Dis 1987; 136: 1016.

Francis, N.A., Melbye, H., Kelly, M.J., Cals, J.W., Hopstaken, R.M., Coenen, S., Butler, C.C. "Variation in family physicians' recording of auscultation abnormalities in patients with acute cough is not explained by case mix. A study from 12 European networks. Eur J Gen Pract 2013; 19: 77–84.

Laënnec, R.T.H., De l'auscultation médiate ou traité du diagnostic des maladies des poumons et du coeur (on mediate auscultation or treatise on the diagnosis of the diseases of the lungs and heart). Paris: Brosson & Chaudé; 1819.

Laënnec, R. T. H., Forbes, J. T., Underwood, C., trans. 1821. Treatise on Diseases of the Chest in Which They Are Described According to Their Anatomical Characters and Their Diagnoses, Established on a New Principle by Means of Acoustic Instruments. C. Underwood, London. Golding Bird, "Advantages presented by the employment of a stethoscope with a flexible tube", London Medical Gazette, vol. 1, pp. 440–442, 11 December 1840.

Littmann, D. An approach to the ideal stethoscope. „JAMA". 178, s. 504-5, 1962. PMID: 14465914.

Littmann, D.: Stethoscope. Patent US 3152659. United States Patent Office, filing date: 16.06.1960; granted: 13.10.1964.

Mangione, S., Nieman, L. Z., Pulmonary Auscultatory Skills During Training in Internal Medicine and Family Practice,. American Journal of Respiratory and Critical Care Medicine 1999; 159:1119–1124.

Mikami, R., Murao, M., Cugell, D.W., Chretien, J., Cole, P., Meier-Sydow, J., Murphy, R.L., Loudon, R.G. International symposium on lung sounds. Synopsis of proceedings. Chest 1987;92:342–5.

Pasterkamp, H. , Brand, P.L.P., Everard, M., Garcia-Marcos, L.,Melbye, H.,Priftis, K. N., Towards the standardisation of lung sound nomenclature,. European Respiratory Journal 2016; 47: 724-732.

Pasterkamp, H., Montgomery. M., Wiebicke, W., “Nomenclature used by health care professionals to describe breath sounds in asthma”. Chest 1987; 92: 346–352.

Rowińska-Zakrzewska, E., Kus, J., Respiratory system diseases (in Polish: Choroby układu oddechowego), PZWL Warszawa 2004, s. 116.

Szczeklik, E., Szczeklik, A., General diagnostics of internal diseases (in Polish: Diagnostyka ogólna chorób wewnętrznych), Państwowy Zakład Wydawnictw Lekarskich, 1979.

Tatoń, J., Czech A., Internal diagnostics (in Polish: Diagnostyka internistyczna), PZWL, Warszawa 2005, s. 368 - 369.

Wilkins, R.L., Dexter, J.R., Murphy, R.L. Jr., et al. “Lung sound nomenclature survey.” Chest 1990; 98: 886–889.

Wilkins, R., Hodgkin, J., Lopez, B., Fundamentals of Lung and Heart Sounds, Third Edition (Book & CD-ROM) 3rd 2004.