Patient's Experience in Pediatric Primary Immunodeficiency Disorders: Computerized Classification of Questionnaires

Urs Mücke[1], Christian Klemann[2], Ulrich Baumann[3], Almut Meyer-Bahlburg[4], Xiaowei Kortum[5], Frank Klawonn[5,6], Werner M. Lechner[7] and Lorenz Grigull[1]*

[1] *Department of Pediatric Hematology and Oncology, Hannover Medical School, Hannover, Germany, [2] Department of Pediatric Surgery, Hannover Medical School, Hannover, Germany, [3] Department of Pediatric Pulmonology, Hannover Medical School, Hannover, Germany, [4] Department of Pediatrics, University Medicine Greifswald, Greifswald, Germany, [5] Helmholtz Centre for Infection Research, Braunschweig, Germany, [6] Ostfalia University of Applied Sciences, Wolfenbuettel, Germany, [7] Improved Medical Diagnostics, IMD GmbH, Hannover, Germany*

Introduction: Primary immunodeficiency disorders (PIDs) are a heterogeneous group of more than 200 rare diseases. Timely diagnosis is of uttermost importance. Therefore, we aimed to develop a diagnostic questionnaire with computerized pattern-recognition in order to support physicians to identify suspicious patient histories.

Materials and methods: Standardized interviews were conducted with guardians of children with PID. The questionnaire based on parental observations was developed using Colaizzis' framework for content analysis. Answers from 64 PID patients and 62 controls were analyzed by data mining methods in order to make a diagnostic prediction. Performance was evaluated by *k*-fold stratified cross-validation.

Results: The diagnostic support tool achieved a diagnostic sensitivity of up to 98%. The analysis of 12 interviews revealed 26 main phenomena observed by parents in the pre-diagnostic period. The questions were systematically phrased and selected resulting in a 36-item questionnaire. This was answered by 126 patients with or without PID to evaluate prediction. Item analysis revealed significant questions.

Discussion: Our approach proved suitable for recognizing patterns and thus differentiates between observations of PID patients and control groups. These findings provide the basis for developing a tool supporting physicians to consider a PID with a questionnaire. These data support the notion that patient's experience is a cornerstone in the diagnostic process.

Keywords: primary immunodeficiency disease, data mining, diagnostic support, questionnaire, Colaizzi

---

Abbreviations: PID, primary immunodeficiency disorder; SVMs, support vector machines, RFs, random forests; LR, logistic regression; NB, naive Bayes classifiers; LD, linear discriminant; NNs, analysis and nearest neighbor classifiers.

INTRODUCTION

Primary immunodeficiency disorders (PIDs) in children are a group of more than 200 rare diseases presenting a wide spectrum of symptoms (1, 2). Although progress in genetic definition has been made, clinical diagnostics remains a challenging task for the general practitioner (GP) or pediatrician. Early diagnosis is of paramount importance because the delay leads to increased mortality, morbidity, and reduced quality of life (3, 4). Although the time to diagnosis varies, diagnostic delay is common (5–8).

Hence, several efforts have been made to support patient–provider communication and early referral to specialists. Diagnostic warning signs, education campaigns, and guidelines have been introduced in order to raise awareness and to shorten diagnostic delay (9, 10). However, Subbarayan et al. (11) state that existing tools do not work sufficiently and new approaches are requested (12). In this regard, patient's medical histories offer clues for considering a PID. Previous studies have emphasized the importance of medical history taking as approximately 80% of the diagnoses could be established by careful history taking only (13, 14). However, these studies do not focus on PID. Due to the multitude of different immune defects and the highly variable clinical presentation, establishing the diagnosis of PID is particularly challenging. In some cases, the underlying rare condition mimics common diseases. In all cases, physician's experience and background determine whether a referral and further testing are ordered (15). Before establishing a diagnosis, parents often recognize peculiarities. Yet, without an immunological expert at hand, it is difficult to put these observations into the context of a particular disease. Therefore, the question of how to decide which patient should receive further investigation gains supreme significance.

Prior to establishing a correct diagnosis, diagnostic errors occur in all medical fields (16, 17) and might be caused by various factors such as unusual or silent presentation of the disease, unavailability of expertise, or inadequate knowledge (18). Key findings suggest that patients should be enabled to tell their story appropriately, and doctors have to compensate an unavoidable lack of experience concerning rare diseases (19, 20).

Therefore, we introduce computerized analysis of parental observations collected in a novel questionnaire to provide additional support. Our aim was not to define a correct and specific diagnosis but to identify the need for timely referral to a PID specialist. Some studies already indicate the potential use of patient-centered questionnaires and data mining in diseases with a rather narrow spectrum of symptoms (21–23). To aid the discovery of unknown correlations or to derive recommendations for further action, computer-assisted analysis of huge amounts of data is useful (24). This targeted pattern analysis has already been established in internet search engines, banking, insurance, and marketing. It is used to make a prediction or to serve as an immediate alert function. In medicine such algorithms have been applied successfully in different contexts (25, 26).

We hypothesized that pre-diagnostic experiences of PID patients could be used to develop a questionnaire. Such a questionnaire should distinguish different patient cohorts using data mining classifiers. This investigation could establish a basis to develop a computerized diagnostic support tool.

MATERIALS AND METHODS

At first, patient-centered semi-structured interviews with guardians of children with a confirmed PID were conducted. The study protocol was approved by the ethics committee of Hannover Medical School and written informed consent was obtained from each guardian. An overview of the diseases included in the interview process is provided in **Table 1**.
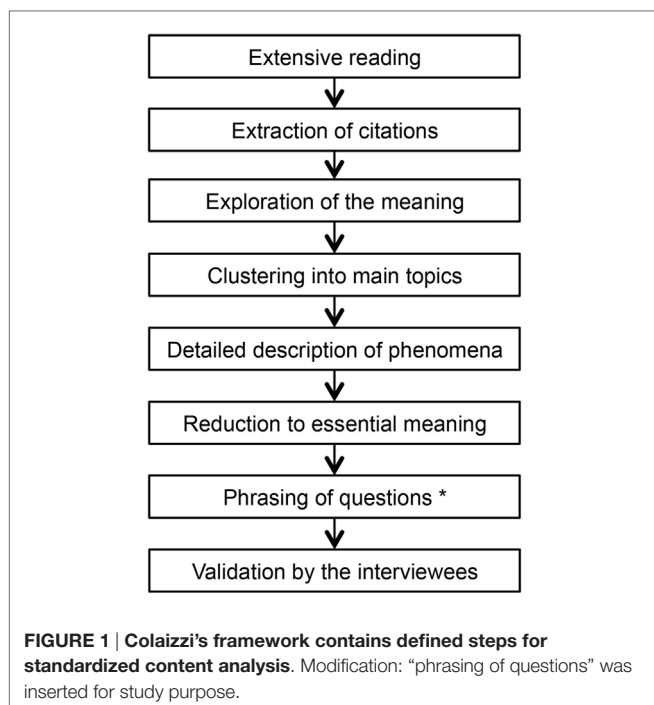
The selection of interview partners followed predefined principles: sufficient speech comprehension, written consent, confirmed PID, children's age 0–18 years, majority of the eight IUIS-categories (27) should be represented by at least one interview, inclusion of further interviewees until theoretical saturation (28). Interviewers used a guideline for standardized procedure and a uniform beginning (what did you observe regarding your child's health and general development before the doctors could finally tell you that your child suffers from a PID?). Each interview was analyzed using Colaizzi's framework with respect to phenomena experienced by parents in the pre-diagnostic period (see **Figure 1**). This standardized procedure contains seven defined steps and is well established in social sciences for content analysis (29). The results of the analysis were documented in a table containing a separate column for each step. The phenomena were sorted depending on occurrence in the interviews in order to integrate relevant observations in the questionnaire. For study purpose, we added a step for question generation (see **Figure 1**).

Members of the study group used the results to draft questions. The parental point of view and the choice of words were integrated in the development of the questionnaire. All questions were summarized in a question pool for further selection. Each question received an identification number for retraceability to its origin. The total number of questions was systematically

TABLE 1 | Spectrum of diseases in the interviews: 12 interviews with parents of children suffering from different diseases were conducted.

Interview	Category	Disease
A	Predominantly antibody deficiencies	CVID
B	Predominantly antibody deficiencies	CVID
C	Well-defined syndromes with immunodeficiency	Ataxia teleangiectasia
D	Predominantly antibody deficiencies	CVID
E	Complement deficiencies	C2 deficiency
F	Combined immunodeficiencies	Cernunnos
G	Undefined immunodeficiency	Combined immunodeficiency with different cytopenia
H	Well-defined syndromes with immunodeficiency	Nijmegen breakage syndrome
I	Congenital defects of phagocyte number, function, or both	X-linked chronic granulomatous disease
J	Undefined immunodeficiency	Undefined severe immunodeficiency
K	Predominantly antibody deficiencies	CVID
L	Well-defined syndromes with immunodeficiency	Nijmegen breakage syndrome

Most common diagnosis was common variable immunodeficiency (CVID, n = 4).

FIGURE 1 | Colaizzi's framework contains defined steps for standardized content analysis. Modification: "phrasing of questions" was inserted for study purpose.

reduced following predefined requirements: each phenomenon relevant for the pre-diagnostic period was represented by at least one item, and the most relevant phenomena were incorporated by additional questions. Duplicates were canceled, a pretest for comprehensibility, consensus in the study group, and expert opinion were integrated. The full resulting questionnaire consequently reflects the pre-diagnostic experience of parents with a child affected by PID (Table S1 in Supplementary Material).

Distribution of the Questionnaire

To generate a data set, the questionnaire was distributed to patients who visited the Immunological Outpatient Clinic of the Hannover Medical School in 2013 and 2014 for regular appointment. All patients had an established diagnosis of PID. An additional online version was accessible for members of the patient group of PID in Germany [Deutsche Selbsthilfe Angeborene Immundefekte (DSAI)] with a special access code. As a control group, we randomly collected questionnaires from guardians of healthy children and children who were hospitalized due to a disease other than PID.

Data Mining Methods

Statistical software libraries offer different computer-based methods to analyze and classify data. Most of these methods are variations of main statistical concepts like vector space methods or artificial neural networks. The tool under discussion uses support vector machines, random forests, logistic regression (LR), naive Bayes classifiers, linear discriminant analysis and nearest neighbor classifiers (30). Classifiers were trained and tested with questionnaires processed in numeric table format. A fusion algorithm combined the different predictions made by each single classifier to one final decision (31). Validation of the system was

performed in two steps. First, the diagnostic accuracy was challenged using two sets of questionnaires: children with PID and healthy children. In a second step, the system had to distinguish between questionnaires of PID children and a combination of randomly chosen healthy and sick children with different diseases, e.g., severe bronchitis, brain tumor, cystic fibrosis, and ulcerative colitis (full spectrum, see Table S2 in Supplementary Material). On both levels, validation was performed by k-fold stratified cross-validation. In addition, we analyzed questions concerning their contribution to the correct classification based on the p-Value for the coefficients in LR, i.e., with the null hypothesis that the coefficient for the question is 0, meaning that the corresponding question does not contribute significantly to the prediction based on LR.

RESULTS

Core Phenomena Were Revealed through Frequent Parental Observations in Standardized, Semi-Structured Interviews

The analysis of the interviews using Colaizzi's framework revealed major themes such as chronological characteristics of infections, parental perception of infections, susceptibility to infections in everyday life, infections of the respiratory tract, and effectiveness of antibiotic treatment (Table S3 in Supplementary Material). These were expressed in words and concepts by the target group.

Generation of a 36-Item Questionnaire by Standardized Content Analysis
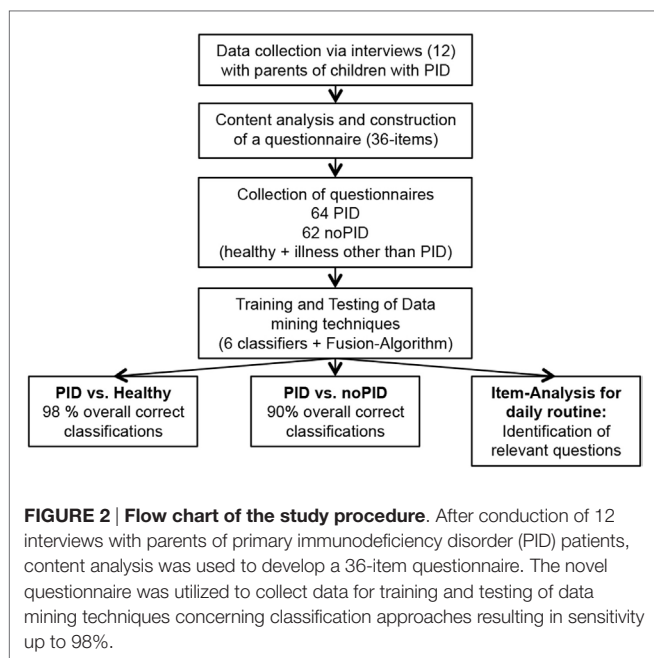
Based on major themes and quotations, 186 preliminary questions were phrased to represent the parental perspective (not shown). For the final version, the questions were systematically reduced (Table S1 in Supplementary Material). Reasons for exclusion were: duplications in form and content, irrelevance for the pre-diagnostic period, and incomprehensibility. Modifications were discussed with experienced immunologists and approved by consensus of the study group. The final questionnaire was subsequently completed by 126 parents. A flow chart of the process is shown in **Figure 2**.

Here, the final questionnaire was completed by parents from different patient collectives. The resulting raw data set consisted of 64 questionnaires from children with PID and 62 questionnaires without PID (35 healthy + 27 patients with illness other than PID, see **Figure 3**). The most common diagnoses in the PID group were common variable immunodeficiency disorders (CVID) ($n = 11$) and agammaglobulinemia ($n = 10$) (see **Table 2**).

Diagnoses in the illness-other-than-PID group were e.g., acute lymphatic leukemia, colitis ulcerosa, cystic fibrosis, and chronic renal failure (Table S2 in Supplementary Material).

Identification of Suspicious Answer Patterns by Data Mining Shows a Sensitivity of up to 98%

99 individuals answered the questionnaire for step 1. This group consisted of 64 individuals with PID and 35 healthy children serving as a control group (see **Figure 2**). A 11-fold stratified

FIGURE 2 | Flow chart of the study procedure. After conduction of 12 interviews with parents of primary immunodeficiency disorder (PID) patients, content analysis was used to develop a 36-item questionnaire. The novel questionnaire was utilized to collect data for training and testing of data mining techniques concerning classification approaches resulting in sensitivity up to 98%.

cross validation showed an overall sensitivity of 98%. Regarding the group of children with PID, 63 of 64 (98%) received the correct diagnosis, and 34/35 of the controls were classified correctly. In total, 97/99 individuals received the correct diagnosis in the cross-validation. Subsequently, 27 questionnaires of randomly chosen children with different diseases were added to the data set (see **Figure 3**). A 21-fold stratified cross-validation was performed to examine diagnostic sensitivity.

In this group (PID patients compared to healthy + illness other than PID), 113/126 (90%) of the questionnaires were classified correctly. In detail, 61/64 individuals with PID received the correct diagnosis, but only 52/62 individuals were correctly classified as "no-PID."

Analysis Discovered Questions Significant for Differentiation

The analysis of the 36 questions, using p-Value computation, revealed 32 questions which contributed to the correct classification significantly ($p < 0.05$). Questions with highest significance in step 2 (PID vs. healthy + illness other than PID) were

Q1   (Did your child suffer from ill health constantly?)
Q31   (Is it true that your child's infections lasted longer than the ones of other children?)
Q32   (Is it true that your child was treated with antibiotics regularly?)
Q35   (Is it true that the doctors could not tell what your child was suffering from?)

Three-dimensional Visualization of Answer Patterns Differentiates Healthy from Sick Children

In terms of a classification problem the answer pattern classification with 36 items consists of 36 dimensions as 36 questions

were answered. Based on Sammon mapping, a three-dimensional representation of the data was generated (see **Figure 4**) (32). It visualizes the questionnaires concerning the correct classification, and illustrates the general principle of distinguishing two groups of patients. Comparing step 1 and step 2 of the study (**Figures 4A,B**) more overlapping of the cohorts is shown in **Figure 4B**. Overlapping questionnaires can represent misdiagnosed cases.

DISCUSSION

The diagnosis of PID presents a challenge to pediatricians due to rare incidence and unspecific symptoms, resulting in delayed referral and negative outcome for patients and parents (3, 4). The pre-diagnostic experience of patients and their families is not sufficiently integrated into the diagnostic process. The study at hand indicates that a novel combination of a questionnaire and data mining techniques provides the means to identify suspicious answer patterns.

Our data show that patient interviews are a feasible tool to generate questionnaires. In order to improve the quality of the questions, parental interviews were obtained to collect observations from the pre-diagnostic time. Colaizzi's method is widely applied in nursing science and qualitative research and has proven useful to derive knowledge from interviews (33, 34). It appears to be useful to collect personal experience and to mine it for the generation of a questionnaire (35). The data generated in this study supports the impact of qualitative research in the diagnostic process (36). Our data give important insights into the parents' perceptions prior to diagnosis as opposed to clinical focus, and likewise reasons for the diagnostic delay could be exposed. An advantage of Colaizzi's framework is the standardized procedure, which makes the connection between citation and generated question retraceable (37).

In our approach, we combined six classifiers and added a fusion algorithm to increase sensitivity. Our combination showed a sensitivity of 84–98%. In comparison of step 1 and step 2, sensitivity decreased when adding sick children to healthy controls. The overall sensitivity of 90% in step 2 still underscores the promising performance of the techniques piloted here. The application of a fusion algorithm improved the diagnostic quality of the tool. Recently, the concept of questionnaires and data mining was successfully piloted by Rother et al. (23). They applied related techniques on pulmonary diseases and achieved encouraging results.

The combination of a questionnaire and data mining has already been tested in diagnosing gastroesophageal reflux disease (GERD) (21). The tool differentiates between complaints caused by GERD and other dyspeptic disorders. In contrast to our study, Horowitz et al. used a shorter questionnaire (15 vs. 36 items), did not use extensive interviews to accumulate observations, and applied less data mining techniques.

It is important to emphasize that the process of professional medical history taking is neither intended nor suitable to be replaced by questionnaires. In fact, history taking might be improved because the diagnostic tool highlights those questions,
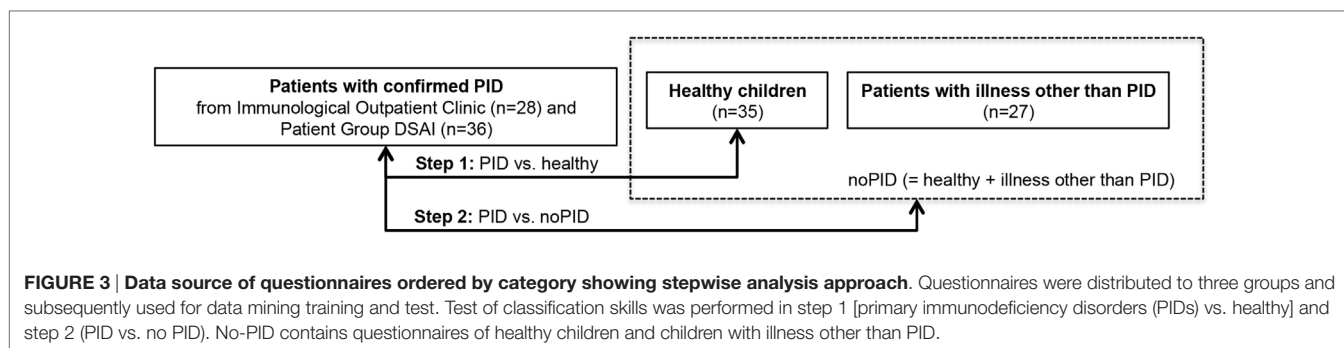
FIGURE 3 | Data source of questionnaires ordered by category showing stepwise analysis approach. Questionnaires were distributed to three groups and subsequently used for data mining training and test. Test of classification skills was performed in step 1 [primary immunodeficiency disorders (PIDs) vs. healthy] and step 2 (PID vs. no PID). No-PID contains questionnaires of healthy children and children with illness other than PID.

TABLE 2 | The four most frequent diagnoses of children with primary immunodeficiency disorders (PIDs): the four most frequent diagnoses of a total of 64 questionnaires received from PID patients.

Diagnosis	n
Common variable immune deficiency (CVID)	11
Agammaglobulinemia	10
Syndromes with recurrent fever (TRAPS + PFAPA)	6
Chronic granulomatous disease	5

TRAPS, TNF receptor-associated periodic syndrome; PFAPA, periodic fever with aphthous stomatitis, pharyngitis, and adenitis.
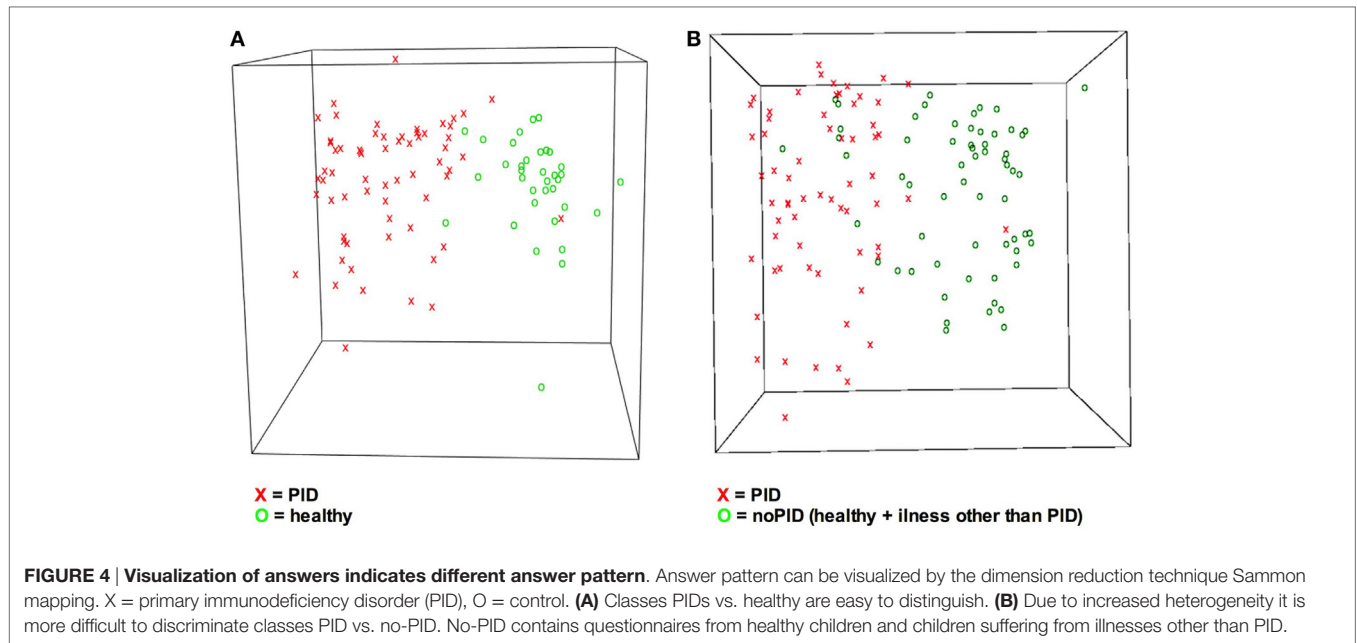
which data mining suggested to be of high diagnostic relevance. Physicians can clarify the past medical history and receive additional hints (38–40). The acceptance of diagnostic questionnaires has already been proven (41).

Some experts stress the notion that more awareness for PID is needed among physicians in tertiary hospitals (11). Yet, Lankisch et al. (42) suggest focusing on primary care physicians and pediatricians without profound experience in PID. Lankisch et al. provided a modified catalog of warning signs and achieved improved results for detecting children with PID compared to classic warning signs (12, 42). However, they underlined the need for improvements in the diagnostic pathway in order to detect preferably 9 out of 10 children suffering from PID. The tool piloted here might help to fill this gap despite the preliminary character of the study as well as its limitations. Different paper-based questionnaires are nowadays used on occasion in outpatient clinics to collect data about patients' medical history. They are based on expert opinion, most are not empirically verified, and they are interpreted manually deriving hints depending on examiners expertise. In contrast, our tool focuses on the patients or guardians perspective, can be statistically reviewed, is extendable, improved by adding more data, and it works independently from the doctor's experience.

Parts of our questionnaire reflect current guidelines (43). Additionally, the questions with high impact for the diagnosis PID share similarities to the classical "warning signs" but also differences. The usage of antibiotics is included in both tools. Simple questions like Q1 (did your child suffer from ill health constantly?) and Q36 (is it true that your child was absent from school/pre-school/kindergarten due to sickness more often than other children?) are solely derived from parental experience and

revealed an important contribution to correct classification. In contrast to an analysis by Subbarayan et al. (11), a positive family history was not significant for finding the correct diagnosis in our cohort. This result supports a presumption by Brodszki et al. (44) who states that the degree of consanguinity in the collective is essential. Kallus (45) underlines that a specific attribute has to be represented in a sample often enough to reach sufficient discriminatory power. For this reason, a questionnaire containing more than 10 items increases the probability to catch sufficient observations. Nevertheless, it is not a single question that makes the difference. It is the combination of several questions with different contributions to the correct classification that makes diagnostic hints possible.

Our study has several limitations: first of all, as it is a study for proof of concept, only a narrow spectrum of diagnoses was included in the interview phase. Likewise, only a small part of all known PID was represented. Second, the number of patients and controls is limited. Nevertheless, cross-validation supports the results which need re-evaluation in a prospective trial. Besides, diagnostic support was also possible for those diagnoses not included in the interviews indicating the plasticity of the system. Third, using a written questionnaire limits the usage of a tool. For example, it requires enough language comprehension. The questionnaire was generated and evaluated in German. Today, the transfer into different languages and cultures and reproduction of the excellent results is unproven. In the era of multicultural societies, a universal comprehensibility independent of language ability would be desirable (46). Tablet computers, which have already been positively evaluated in other areas could be used to reduce the effort for completion of the questionnaire and its interpretation (47, 48). Furthermore, possible online adaptions of paper versions already proofed usability in other areas (49). A further limitation is related to the selection of controls. The cohorts do not reflect the regular spectrum of patients seen by GP or pediatricians. Ideally, this study should have incorporated a control group representative of the day-to-day population of GP and/or general pediatricians. Controls presenting with symptoms suspicious for a PID would have answered a questionnaire prior to the referral to an immunologist. Due to limited resources, this evaluation will be part of a future trial based on this pilot proof of concept. Thus, further prospective studies with patients presenting PID-like symptoms are needed to verify our approach.

FIGURE 4 | **Visualization of answers indicates different answer pattern**. Answer pattern can be visualized by the dimension reduction technique Sammon mapping. X = primary immunodeficiency disorder (PID), O = control. **(A)** Classes PIDs vs. healthy are easy to distinguish. **(B)** Due to increased heterogeneity it is more difficult to discriminate classes PID vs. no-PID. No-PID contains questionnaires from healthy children and children suffering from illnesses other than PID.

Taken together, we successfully piloted a questionnaire-based tool intended to classify different patient cohorts.

A diagnostic tool for detection of children with PID could be used in different settings: the GP or the pediatrician could hand it to parents with children showing recurrent infections to reassure the need for further diagnostics. The questionnaire could be answered on a tablet computer while sitting in the waiting area. A diagnostic suggestion would only be presented to the physician who could include the computerized diagnostic suggestion into the diagnostic workflow and use eye-catching answers for clarifying.

A series of five clinical cases as an example for the prospective classification process is given in Table S4 in Supplementary Material. Extensive investigation of user friendliness is part of a prospective evaluation. The principles applied in this study might also be extended to other groups of rare diseases beyond PID.

AUTHOR CONTRIBUTIONS

UM and LG conceptualized the study, conducted the interview, analyzed transcripts, and wrote the manuscript. WL, FK, and XK performed all data mining. CK, UB, and AM-B supervised the clinical process. All authors discussed the results and commented on the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://journal.frontiersin.org/article/10.3389/fimmu.2017.00384/full#supplementary-material.

REFERENCES

1. Picard C, Al-Herz W, Bousfiha A, Casanova JL, Chatila T, Conley ME, et al. Primary immunodeficiency diseases: an update on the classification from the international union of immunological societies expert committee for primary immunodeficiency 2015. *J Clin Immunol* (2015) 35:696–726. doi:10.1007/s10875-015-0201-1
2. Boyle JM, Buckley RH. Population prevalence of diagnosed primary immunodeficiency diseases in the United States. *J Clin Immunol* (2007) 27:497–502. doi:10.1007/s10875-007-9103-1
3. Joshi AY, Iyer VN, Hagan JB, St Sauver JL, Boyce TG. Incidence and temporal trends of primary immunodeficiency: a population-based cohort study. *Mayo Clin Proc* (2009) 84:16–22. doi:10.1016/S0025-6196(11)60802-1
4. Tabolli S, Giannantoni P, Pulvirenti F, La Marra F, Granata G, Milito C, et al. Longitudinal study on health-related quality of life in a cohort of 96 patients with common variable immune deficiencies. *Front Immunol* (2014) 5:1–9. doi:10.3389/fimmu.2014.00605
5. Michos A, Raptaki M, Tantou S, Tzanoudaki M, Spanou K, Liatsis M, et al. Primary immunodeficiency diseases: a 30-year patient registry from the referral center for primary immunodeficiencies in Greece. *J Clin Immunol* (2014) 34(7):836–43. doi:10.1007/s10875-014-0066-8
6. Aghamohammadi A, Bahrami A, Mamishi S, Mohammadi B, Abolhassani H, Parvaneh N, et al. Impact of delayed diagnosis in children with primary antibody deficiencies. *J Microbiol Immunol Infect* (2011) 44:229–34. doi:10.1016/j.jmii.2011.01.026

7. Maarschalk-Ellerbroek LJ, Hoepelman AIM, Van Montfrans JM, Ellerbroek PM. The spectrum of disease manifestations in patients with common variable immunodeficiency disorders and partial antibody deficiency in a university hospital. *J Clin Immunol* (2012) 32:907–21. doi:10.1007/s10875-012-9671-6

8. Quinti I, Soresina A, Spadaro G, Martino S, Donnanno S, Agostini C, et al. Long-term follow-up and outcome of a large cohort of patients with common variable immunodeficiency. *J Clin Immunol* (2007) 27:308–16. doi:10.1007/s10875-007-9075-1

9. Modell V. The impact of physician education and public awareness on early diagnosis of primary immunodeficiencies: Robert A. Good immunology symposium. *Immunol Res* (2007) 38:43–7. doi:10.1007/s12026-007-0048-5

10. Farmand S, Baumann U, Borte M, Franke K, Habermehl P, Kapaun P, et al. Interdisziplinäre AWMF-Leitlinie zur Diagnostik von primären Immundefekten (S2k) interdisciplinary AWMF guideline for the diagnostics of primary immunodeficiency. *Klin Padiatr* (2011) 223:378–85. doi:10.1055/s-0031-1287835

11. Subbarayan A, Colarusso G, Hughes MS, Gennery RA, Slatter M, Cant JA, et al. Clinical features that identify children with primary immunodeficiency diseases. *Pediatrics* (2011) 127:810–6. doi:10.1542/peds.2010-3680

12. Arkwright PD, Gennery AR. Ten warning signs of primary immunodeficiency: a new paradigm is needed for the 21st century. *Ann N Y Acad Sci* (2011) 1238:7–14. doi:10.1111/j.1749-6632.2011.06206.x

13. Hampton JR, Harrison MJ, Mitchell JR, Prichard JS, Seymour C. Relative contributions of history-taking, physical examination, and laboratory investigation to diagnosis and management of medical outpatients. *Br Med J* (1975) 2:486–9. doi:10.1136/bmj.2.5969.486

14. Peterson MC, Holbrook JH, Von Hales D, Smith NL, Staker LV. Contributions of the history, physical examination, and laboratory investigation in making medical diagnoses. *West J Med* (1992) 156:163–5. doi:10.1097/00006254-199210000-00013

15. Budych K, Helms TM, Schultz C. How do patients with rare diseases experience the medical encounter? Exploring role behavior and its impact on patient-physician interaction. *Health Policy* (2012) 105:154–64. doi:10.1016/j.healthpol.2012.02.018

16. Singh H, Thomas EJ, Wilson L, Kelly PA, Pietz K, Elkeeb D, et al. Errors of diagnosis in pediatric practice: a multisite survey. *Pediatrics* (2010) 126:70–9. doi:10.1542/peds.2009-3218

17. Singh H, Petersen LA, Thomas EJ. Understanding diagnostic errors in medicine: a lesson from aviation. *Qual Saf Health Care* (2006) 15:159–64. doi:10.1136/qshc.2005.016444

18. Graber M, Gordon R, Franklin N. Reducing diagnostic errors in medicine: what's the goal? *Acad Med* (2002) 77:981–92. doi:10.1097/00001888-200210000-00009

19. Graber ML. Taking steps towards a safer future: measures to promote timely and accurate medical diagnosis. *Am J Med* (2008) 121:S43–6. doi:10.1016/j.amjmed.2008.02.006

20. Sarkar U, Bonacum D, Strull W, Spitzmueller C, Jin N, López A, et al. Challenges of making a diagnosis in the outpatient setting: a multi-site survey of primary care physicians. *BMJ Qual Saf* (2012) 21:641–8. doi:10.1136/bmjqs-2011-000541

21. Horowitz N, Moshkowitz M, Halpern Z, Leshno M. Applying data mining techniques in the development of a diagnostics questionnaire for GERD. *Dig Dis Sci* (2007) 52:1871–8. doi:10.1007/s10620-006-9202-5

22. Grigull L, Lechner WM. Supporting diagnostic decisions using hybrid and complementary data mining applications: a pilot study in the pediatric emergency department. *Pediatr Res* (2012) 71:725–31. doi:10.1038/pr.2012.34

23. Rother AK, Schwerk N, Brinkmann F, Klawonn F, Lechner W, Grigull L. Diagnostic support for selected paediatric pulmonary diseases using answer-pattern recognition in questionnaires based on combined data mining applications-a monocentric observational pilot study. *PLoS One* (2015) 10:1–17. doi:10.1371/journal.pone.0135180

24. Yoo I, Alafaireet P, Marinov M, Pena-Hernandez K, Gopidi R, Chang J-F, et al. Data mining in healthcare and biomedicine: a survey of the literature. *J Med Syst* (2012) 36:2431–48. doi:10.1007/s10916-011-9710-5

25. Yang ST, Lee JD, Chang TC, Huang CH, Wang JJ, Hsu WC, et al. Discrimination between Alzheimer's disease and mild cognitive impairment using SOM and PSO-SVM. *Comput Math Methods Med* (2013) 2013:253670. doi:10.1155/2013/253670

26. Baxt WG, Shofer FS, Sites FD, Hollander JE. A neural network aid for the early diagnosis of cardiac ischemia in patients presenting to the emergency department with chest pain. *Ann Emerg Med* (2002) 40:575–83. doi:10.1067/mem.2002.129171

27. Al-Herz W, Bousfiha A, Casanova JL, Chapel H, Conley ME, Cunningham-Rundles C, et al. Primary immunodeficiency diseases: an update on the classification from the International Union of Immunological Societies Expert Committee for Primary Immunodeficiency. *Front Immunol* (2011) 5:162. doi:10.3389/fimmu.2011.00054

28. Flick U. *An Introduction to Qualitative Research.* 4th ed. SAGE (2009). 529 p. Available from: http://books.google.cz/books?id=sFv1oWX2DoEC

29. Colaizzi PF. Psychological Research as the Phenomenologist's View It. In: Valle RS, King M, editors. *Existential-Phenomenological Alternatives for Psychology.* New York: Oxford University Press (1978). p. 48–71.

30. Fernández-Delgado M, Cernadas E, Barro S, Amorim D. Do we need hundreds of classifiers to solve real world classification problems? *J Mach Learn Res* (2014) 15:3133–81.

31. Grigull L, Lechner W, Petri S, Kollewe K, Dengler R, Mehmecke S, et al. Diagnostic support for selected neuromuscular diseases using answer-pattern recognition and data mining techniques: a proof of concept multicenter prospective trial. *BMC Med Inform Decis Mak* (2016) 16:31. doi:10.1186/s12911-016-0268-5

32. Sammon JW. A nonlinear mapping for data structure analysis. *IEEE Trans Comput* (1969) C-18:401–9. doi:10.1109/T-C.1969.222678

33. Keaton KA, Pierce LL. Cardiac therapy for men with coronary artery disease: the lived experience. *J Holist Nurs* (2000) 18:63–85. doi:10.1177/089801010001800107

34. Hilliard C, O'Neill M. Nurses' emotional experience of caring for children with burns. *J Clin Nurs* (2010) 19:2907–15. doi:10.1111/j.1365-2702.2009.03177.x

35. Nadarajah SR, Buchholz SW, Wiegand DL, Berger A. The lived experience of individuals in cardiac rehabilitation who have a positive outlook on their cardiac recovery: a phenomenological inquiry. *Eur J Cardiovasc Nurs* (2016) 16(3):230–9. doi:10.1177/1474515116651977

36. Hinton D, Kirk S. Paediatric multiple sclerosis: a qualitative study of families' diagnosis experiences. *Arch Dis Child* (2015) 100:623–9. doi:10.1136/archdischild-2014-306523

37. Sanders C. Application of Colaizzi's method: interpretation of an auditable decision trail by a novice researcher. *Contemp Nurse* (2014) 14:292–302. doi:10.5172/conu.14.3.292

38. Arora S, Goldberg AD, Menchine M. Patient impression and satisfaction of a self-administered, automated medical history-taking device in the emergency department. *West J Emerg Med* (2014) 15:35–40. doi:10.5811/westjem.2013.2.11498

39. Koeneman MM, Kruse A-J, Sep SJS, Gubbels CS, Slangen BFM, van Gorp T, et al. A family history questionnaire improves detection of women at risk for hereditary gynecologic cancer: a pilot study. *Fam Cancer* (2014) 13:469–75. doi:10.1007/s10689-014-9711-x

40. DeAngelis AF, Chambers IG, Hall GM. The accuracy of medical history information in referral letters. *Aust Dent J* (2010) 55:188–92. doi:10.1111/j.1834-7819.2010.01221.x

41. Hershey CO, Grant BJB. Controlled trial of a patient-completed history questionnaire: effects on quality of documentation and patient and physician satisfaction. *Am J Med Qual* (2016) 17:126–35. doi:10.1177/106286060201700402

42. Lankisch P, Schiffner J, Ghosh S, Babor F, Borkhardt A, Laws HJ. The Duesseldorf warning signs for primary immunodeficiency: is it time to change the rules? *J Clin Immunol* (2015) 35:273–9. doi:10.1007/s10875-015-0149-1

43. Farmand S, Baumann U, von Bernuth H, Borte M, Foerster-Waldl E, Franke K, et al. [Interdisciplinary AWMF guideline for the diagnostics of primary immunodeficiency]. *Klin Padiatr* (2011) 223:378–85. doi:10.1055/s-0031-1287835

44. Brodszki N, Jönsson G, Skattum L, Truedsson L. Primary immunodeficiency in infection-prone children in southern Sweden: occurrence, clinical characteristics and immunological findings. *BMC Immunol* (2014) 15:31. doi:10.1186/s12865-014-0031-6

45. Kallus KW. *Erstellung von Frabeogen. 1. Aufl.* Wien: Facultas.wuv (2010).

46. Gogtay NJ, Thatte UM, Dasgupta B, Deshpande S. Use of the WOMAC questionnaire in Mumbai and the challenges of translation and cross cultural adaptation. *Indian J Med Ethics* (2016) 10:33–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23439195

47. Mickan S, Atherton H, Roberts NW, Heneghan C, Tilson JK. Use of handheld computers in clinical practice: a systematic review. *BMC Med Inform Decis Mak* (2014) 14:56. doi:10.1186/1472-6947-14-56

48. Fleischmann R, Duhm J, Hupperts H, Brandt SA. Tablet computers with mobile electronic medical records enhance clinical routine and promote bedside time: a controlled prospective crossover study. *J Neurol* (2015) 262:532–40. doi:10.1007/s00415-014-7581-7

49. Bland JDP, Rudolfer S, Weller P. Prospective analysis of the accuracy of diagnosis of carpal tunnel syndrome using a web-based questionnaire. *BMJ Open* (2014) 4:e005141. doi:10.1136/bmjopen-2014-005141