A diagnostic support system based on pain drawings: binary and k-disease classification of EDS, GBS, FSHD, PROMM, and a control group with Pain2D

Dorian Emmert
Universitätsklinikum Bonn: Universitatsklinikum Bonn
Natasza Szczypien
Ostfalia University of Applied Sciences: Ostfalia Hochschule fur angewandte Wissenschaften
Tim T.A. Bender
Universitätsklinikum Bonn: Universitatsklinikum Bonn
Lorenz Grigull
Universitätsklinikum Bonn: Universitatsklinikum Bonn
Annette Gass
Universitätsklinikum Bonn: Universitatsklinikum Bonn
Carolina Link
Universitätsklinikum Bonn: Universitatsklinikum Bonn
Frank Klawonn
Helmholtz Centre for Infection Research: Helmholtz-Zentrum fur Infektionsforschung GmbH
Rupert Conrad
Universitatsklinikum Bonn
Martin Mücke
University Hospital Aachen: Universitatsklinikum Aachen
Julia Sellin (jsellin@ukaachen.de)
University Hospital Aachen: Universitatsklinikum Aachen https://orcid.org/0000-0001-5354-2872

Research Article

Keywords: ORPHA: 269, ORPHA: 606, ORPHA: 2103, ORPHA: 287, diagnostic support, AI, rare diseases, k-disease classification, pain drawings, machine learning

Posted Date: March 28th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1357471/v1

A diagnostic support system based on pain drawings: binary and k-disease classification of EDS, GBS, FSHD, PROMM, and a control group with Pain2D

Emmert D[1*], Szczypien N[2,3*], Tim T.A. Bender[1], Grigull L[1], Gass A[4], Link C[4], , Klawonn F[2,3], Conrad R[5+], Mücke M[6,7+], Sellin J[6+]

[1]Center for Rare Diseases Bonn (ZSEB), University Hospital Bonn, Germany

[2]Institute for Information Engineering, Ostfalia University of Applied Sciences, Wolfenbüttel, Germany

[3]Biostatistics Group, Helmholtz Centre for Infection Research, Braunschweig, Germany

[4]Department of Pain therapy, University Hospital Bonn, Germany

[5]Department of Psychosomatic Medicine and Psychotherapy, University Hospital Bonn, Germany

[6]Institute for Digitalization and General Medicine, University Hospital RWTH Aachen, Germany

[7]Center for Rare Diseases Aachen (ZSEA), University Hospital RWTH Aachen, Germany

*These authors contributed equally to the study

[+]Correspondence to: jsellin@ukaachen.de; mamuecke@ukaachen.de; rupert.conrad@ukbonn.de;

Abstract

Background and objective: The diagnosis of rare diseases (RDs) is often challenging due to their rarity, variability and the high number of individual RDs, resulting in a delay in diagnosis with adverse effects for patients and healthcare systems. The development of computer assisted diagnostic decision support systems (DDSS) could help to improve these problems by supporting differential diagnosis and by prompting physicians to initiate the right diagnostic tests. Towards this end, we developed, trained and tested a machine learning model implemented as part of the software called Pain2D to classify four rare diseases (EDS, GBS, FSHD and PROMM), as well as a control group of unspecific chronic pain, from pen-and-paper pain drawings filled in by patients.

Methods: Pain drawings (PDs) were collected from patients suffering from one of the four RDs, or from unspecific chronic pain. The latter PDs were used as an outgroup in order to test how Pain2D handles more common pain causes. A total of 261 (59 EDS, 29 GBS, 35 FSHD, 89 PROMM, 50 unspecific chronic pain) PDs were collected and used to generate disease specific pain profiles (PP). PDs were then classified by Pain2D in a leave-one-out-cross-validation approach.

Results: Pain2D was able to classify the four rare diseases with an accuracy of 61-77% with its binary classifier. EDS, GS and FSHD were classified correctly by the Pain2D k-disease classifier with sensitivities between 63-83% and specificities between 83-90%. For PROMM, the k-disease classifier achieved a sensitivity of 51% and specificity of 90%.

Conclusions: Pain2D is a scalable, open-source tool that could potentially be trained for all diseases presenting with pain.

Keywords: ORPHA: 269, ORPHA: 606, ORPHA: 2103, ORPHA: 287, diagnostic support, AI, rare diseases, k-disease classification, pain drawings, machine learning

Introduction and Background

Rare diseases pose particular challenges for health care systems as a result of their infrequency, diversity and often complex symptomatology. In particular the diagnosis of rare diseases (RDs) is often difficult, with adverse consequences for affected individuals and health care systems. A disease is categorized as rare in the European Union if it affects less than 1 in 2000 people[1]. Similar definitions exist in other regions of the world (for example, the NIH defines a disease as rare if it affects less than 200 000 people in the US[2]. As there are more than 7000 known individual rare diseases, resulting in an estimated 30 mio affected in the EU and about 400 mio people worldwide, it is apparent that the sheer number of possibilities makes it impossible for individual physicians to know all of them. In addition, many rare diseases present with multifaceted clinical symptoms. As a result, affected individuals often wait for a long time until they receive the correct diagnosis (~7 years on average [3]). Long time to diagnosis contributes to mental, physical and social distress. In addition, due to many medical consultations and resulting redundant diagnostic procedures, the burden for health care systems is further increased[3].

In recent years, a number of studies and reviews focused on technical solutions to improve diagnosis of rare diseases ([4][5][6] and references therein). Clinical decision support systems (CDSSs), while used frequently in a number of clinical settings, are not yet widespread in the context of diagnosis (sometimes referred to as diagnosis decision support systems, DDSSs). Negative physician biases, insufficient accuracy and lacking integration with clinical information systems in use are discussed as underlying the lack of acceptance of DDSSs, in spite of promising results with regard to their effectiveness [7] [8].

With this study we investigate whether pain drawings (PD) can be used to build a machine learning based DDSS which allows the detection of rare diseases causing chronic pain in patients. PDs are used to communicate the experience of pain from patients to health care providers. Patients mark painful body regions in a simple line drawings of the human body and its parts to indicate where they experience pain. This often results in a more precise description of affected body parts than just via verbal description. Another advantage is that patients can fill out the PD for example in the waiting room during their visit at the physician, thereby saving time during the consultation. PDs were first established by Palmer in 1949 [9]. While there are a growing number of studies dealing with the topic of PDs [10] [11] [12][13][14][15], none of them focuses on their usefulness as a diagnostic tool for rare diseases.

We have previously shown that a binary classifier from Pain2D-Tool of the PD software Pain2D is able to distinguish between two rare diseases, Ehlers-Danlos syndrome (EDS) and Guillain-Barré syndrome (GBS) [16]. Here, we test the performance of the binary classifier, as well as a new k-disease classifier implemented into Pain2D, on four different diseases (EDS, GBS, facioscapulohumeral muscular dystrophy (FSHD) and proximal myotonic myopathy (PROMM)) and a control group of patients with non-specific chronic pain.

EDS is a group of inherited disorders affecting the connective tissue with a prevalence between 1:150 000 and 1:5 000, depending on the population [17]. EDS can present clinically with variable symptomatology, from mild skin hyperextensibility, joint hypermobility, and tissue fragility, to severe physical disability and life-threatening vascular complications [18].

GBS is caused by autoantibodies attacking peripheral nerve components triggered by an infection, resulting in a polyradiculoneuropathy with variable clinical presentation [19]. 1.1 to 1.8 per 100,000 persons suffer from GBS each year [20]. Symptoms can include a range from ascending bilateral limb weakness to decreased reflexes and severe back or extremity pain [21] [22]

FSHD is an autosomal-dominantly inherited muscular dystrophy which characteristically affects facial muscles, shoulder girdles, and upper arms [23]. The prevalence of FSHD is estimated to range between 2.03 and 6.8 per 100,000 individuals [24]. Pain in the affected regions is a common symptom of patients suffering from FSHD [25]. The diagnosis of FSHD can be challenging, especially in milder forms, as typical symptoms of FSHD may not be present [26].

PROMM is a subtype of myotonic dystrophies also referred to as myotonic dystrophy type 2 (DM2). Myotonic dystrophies are autosomal-dominantly inherited diseases that have in common muscular involvement (myotonia, muscle weakness, muscular dystrophy), eye manifestations (early onset cataracts), cardiac conduction defects, and endocrine disorders [27]. Only a few studies deal with the prevalence of PROMM, with estimates for Europe ranging between 9:100,000 [28][29] and 1:1830 in Finland [30]. As the name implies, PROMM in contrast to DM1 typically affects proximal muscles [31]. 50-80 % of PROMM patients suffer from pain, which can be exercise-related, musculoskeletal, or abdominal [32]. While myotonic dystrophies are the most common forms of adult-onset dystrophies, PROMM is likely underdiagnosed due to its heterogeneous phenotype and unclear age of onset [31].

The four diseases and the control group were chosen to cover a range from very different causes of pain (e.g., GBS as an inflammatory disease vs. FSHD as an inherited

neuromuscular disease) to more similar causes (e.g., FSHD and PROMM as two autosomal-dominantly inherited neuromuscular diseases) to test the ability of the Pain2D classifiers to distinguish between more or less similar rare diseases. The control group was added to elucidate if more common and unspecific causes of pain can be separated from rare diseases with Pain2D-Tool.

Material and Methods

Study design and data collection

Between 2017 and 2019, a total of 35 patients with FSHD (10 male, 25 female) and 90 patients with PROMM participated in this study. Of the latter, one PD was empty and had to be excluded from further analysis, resulting in the inclusion of 89 PROMM PDs (29 male, 60 female participants). Patients were recruited at the neuromuscular out-patient clinic of University Hospital Bonn, Germany. In addition, we contacted the German association for neuromuscular diseases to support us in finding patients willing to participate in our study. Inclusion criteria were a confirmed diagnosis, age above 18 years, and written informed consent.

Between 2019 and 2020, a total of 50 participants with unspecific chronic pain (19 male, 31 female) were recruited for this study. Inclusion criteria for this group were chronic pain due to a common disease (e.g., post-zoster neuralgia) above six months duration and age above 18 years. Exclusion criteria for this group were rare comorbidities. These patients were recruited at the out-patient pain clinic of University Hospital Bonn and at general medicine practices in Bonn, Germany.

In this study, 35 PDs of Facioscapulohumeral muscular dystrophy (FSHD), 89 PDs of proximal myotonic dystrophy (PROMM), and 50 PDs from a control group with common

causes of chronic pain (CP) were included. Based on genetic findings, two subtypes of FSHD (FSHD1 and FSHD2) can be differentiated, but both have a similar clinical phenotype [23]. We therefore did not distinguish between the two subtypes in our study population. In addition, we used 88 PDs from two different rare diseases that were recruited as part of a previous study: 59 EDS from Ehlers-Danlos syndrome (EDS) and 29 GBS from Guillain Barre syndrome (GBS) [16].

Recruited participants filled in pain drawings as previously described [16].

In accordance with German privacy laws and the declaration of Helsinki, PDs were pseudonymized before sending them from the *Center for Rare Diseases Bonn (ZSEB)* to the analysis server.

This study was registered at the German register for clinical studies DRKS (DRKS-ID: DRKS00014776 (participants recruited for this study) and DRKS00014777 (previously recruited participants, [16]).

Pain2D - Software package for pain drawing analysis

Pain2D is a software package previously developed by our group [16] for the automated processing and analysis of pain drawings based on a template, which can be printed and filled out on paper. As part of the package, an application for tablets (Pain2D-Tablet, paperless) is available, which was however not used for this study. The application Pain2D is open-source and holds a GPL v3.0 license for researchers who want to participate in its future development or test their own pain drawings with the classifiers we have implemented. The application was developed with the open-source statistics and graphics software R and RShiny. For more detailed information about Pain2D, please visit www.pain2d.com.

Pain2D generates pain profiles (PPs, also known as pain frequency maps[12]) by overlapping all PDs which belong to one diagnostic group. They are depicted as color coded heatmaps, with lower to higher percentages of marked pixels in the summarized PDs labeled in blue to red, with yellow indicating 50% of PDs had marked that pixel. Pixels that were empty in all PDs are depicted in white.

Pain2D offers a binary and a k-disease classifier. The function of the binary classifier has been previously described [16]. In short, it classifies PDs by calculating the Ružička similarity of a given PD to two pain profiles, in a leave-one-out cross-validation (LOOCV) approach. Classification occurs according to the highest probability, with the cut-off set to 0.5. The k-disease classifier works in a similar manner, but calculates Ružička similarities of a given PD to k PPs, where k is the number of considered diseases. In our case, k was 5, as 5 diagnostic groups were part of the test data set.

Statistical evaluation

For the classification results, receiver operating characteristic (ROC) curves were plotted and AUC values calculated using the pROC package from R. Confidence intervals were calculated as indicated. The leave-one-out-cross-validated confusion matrices were tested with Fisher's exact test (binary classification) or χ^2 test (k-disease classification) for better than random classification, as indicated.

Results

Pain profiles of the five diagnostic groups (EDS, GBS, PROMM, FSHD, CP)

Pain2D generates pain profiles, which provide a visual result of the sum of all PDs of one diagnostic group and serve as the basis for similarity measurement of an individual PD for classification by Pain2D.

EDS. The pain profile of EDS reveals that most patients experience pain along the vertebral column with the neck and the tailbone, and the knee joints. These regions were marked by approximately 70% of the participating patients. In addition, nearly 50% of the participants marked the shoulder region, the elbows and the thumb saddle joint (Figure 1A; compare [16].

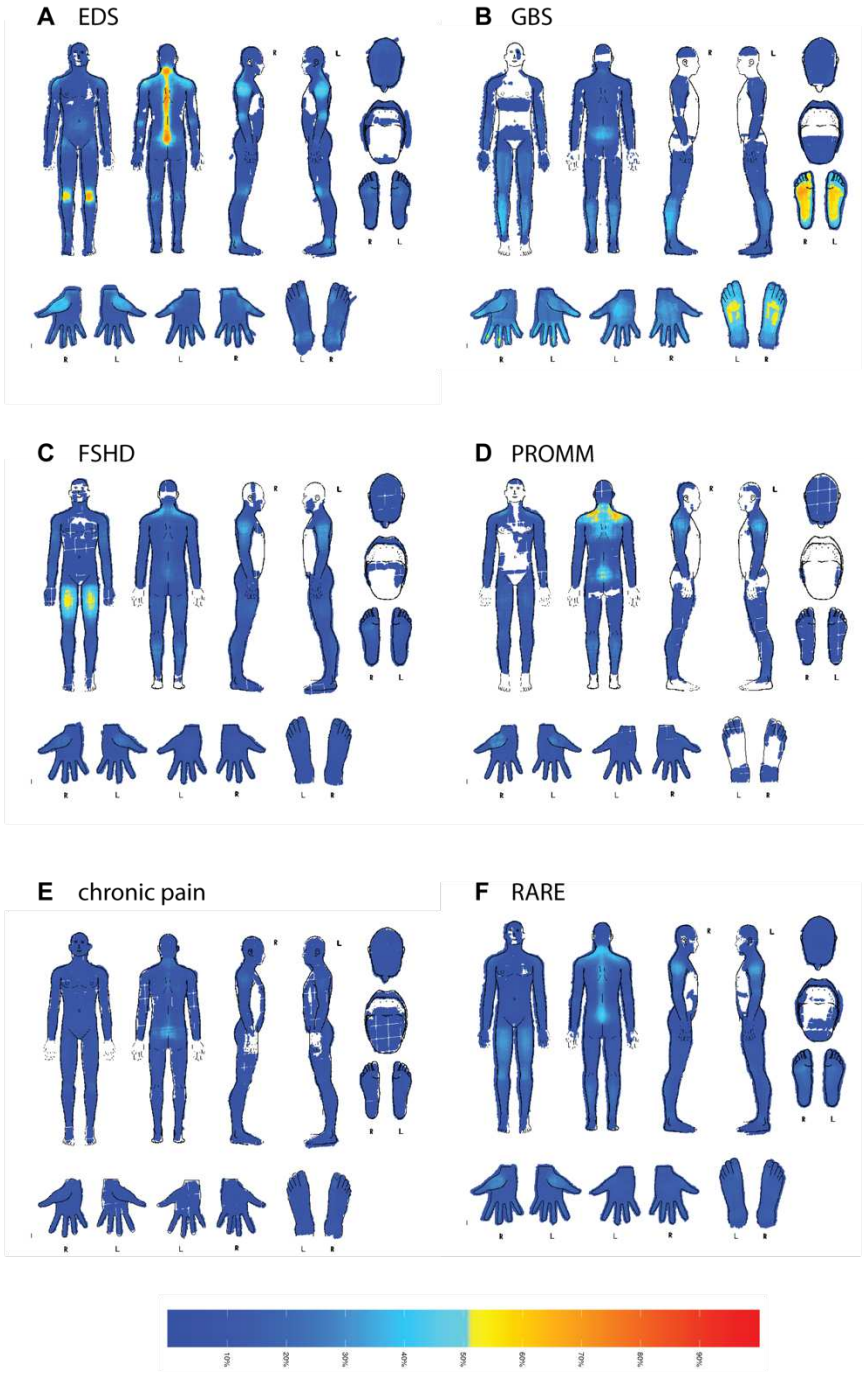
GBS. The most prominent regions marked by patients were the dorsal and plantar side of the feet (~70 % of patients). In addition, about 50 % marked the palmar side of the finger tips, the dorsal side of the left palm and the tailbone (Figure 1B; compare [16].

PROMM. The pain profile of PROMM shows that around 50-60% of patients marked the upper legs as a painful region. Other less frequent regions marked were the shoulders, the lower legs and the lower back (Figure 1C).

FSHD. As shown in Figure 1D, the most frequent body regions marked by patients with FSHD are the shoulders and the lower back with percentages of 50 to 60 %. In addition, the upper arms were marked by ca. 40% (Figure 1D).

Chronic pain (CP). The disease pattern of our control group (Figure 1E) shows that pain patterns are more equally distributed between patients, with percentages marked well below 50%. The localization with the highest percentage is the lower back with ca. 40 %. This result is consistent with the expectation of a more unspecific pain pattern as this group was suffering from various common causes of chronic pain.

RARE. We also generated a pain profile of all four rare diseases (EDS, GBS, PROMM and FSHD) in order to test if PDs from a group of rare diseases show similarities that allow distinction from other causes of pain. Accordingly, the resulting pain profile shows the typical pain areas from all four rare diseases, but at lower percentages (as a sum projection of the four RD pain profiles; Fig. 1F). This “consensus” pain profile is however not equally informed by the four rare diseases, as different numbers of individual PDs for each disease were included. For example, the data set contains 89 PDs of PROMM, but only 29 of GBS.



210

211 Figure 1: Pain profiles of the five diagnostic groups used in this study, EDS (A), GBS (B), PROMM

212 (C), FSHD (D), chronic pain (CP, E) and RARE (F). The depicted pain profiles were constructed by

Pain2D from 29 EDS (A), 59 GBS (B), 89 PROMM (C), 35 FSHD (D) and 50 CP (E) PDs. RARE is based on 29 EDS, 59 GBS, 89 PROMM and 35 FSHD P'D's (F).

The binary classifier of Pain2D can differentiate between a group of rare diseases (comprised of EDS, GBS, FSHD and PROMM) and unspecific chronic pain (CP) with high sensitivity but lower specificity.

		Predicted		
		RARE	CP	Sum_{true}
TRUE	RARE	200 (94%)	12 (6%)	212 (100%)
	CP	29 (58%)	21 (42%)	50 (100%)
	$Sum_{predicted}$	228 (87%)	33 (13%)	261 (100%)

Table 1: Confusion matrix RARE vs. CP. Percentages are relative to Sum_{true}.

We were interested if PDs can be used to predict the presence of a rare disease as opposed to a common cause for chronic pain. As a test, we grouped all four rare diseases into the group RARE and classified all PDs (EDS, GBS, PROMM, FSHD, CP) with the binary classifier of Pain2D into RARE or CP. With the standard threshold of 0.5 for the binary classifier we reached a very good sensitivity of 94 % and a notably lower specificity of 42%.

Fisher's exact test was applied to the confusion matrix (Tab. 1) and resulted in a p-value < 0.001, indicating that the binary classifier performs better than random guessing and can indeed distinguish between rare diseases and common causes for chronic pain in the test setting. A receiver operating characteristic (ROC) curve was plotted and the R package pROC [33] was used to calculate the best threshold for classification of the given data set (Figure 2, blue crosshair). This resulted in an optimal threshold for classification of 0.41, which led to a slightly lower sensitivity of 82%, but considerably increased specificity of 70%. The calculated 95% confidence band (light blue area) for sensitivity shows low variance for all thresholds. The area under the curve (AUC) of 0.82 indicates good separability between the four rare diseases and chronic pain.

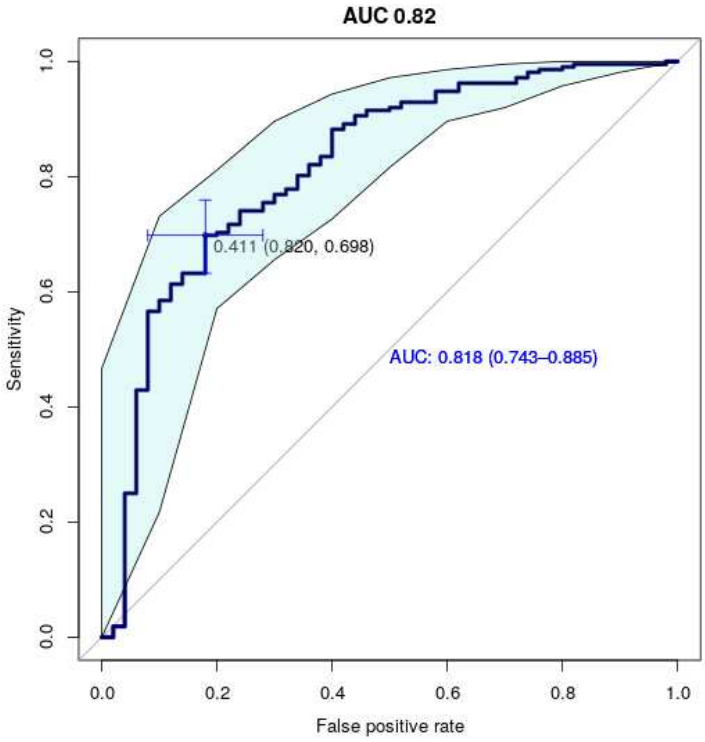


Figure 2: ROC curve for classification of PDs into RARE and CG with the binary classifier of Pain2D. The light blue area indicates the 95% confidence interval. Blue crosshairs indicate optimal classification threshold of 0.41.

The binary classifier of Pain2D can separate each of the four rare diseases from chronic pain with high sensitivity but low values for specificity.

In addition to a general prediction of the presence of a rare disease vs. a common cause of chronic pain (Table 1, Fig. 2), we wanted also to test if the binary classifier of Pain2D is able to separate PDs of each of the four tested rare diseases (DES, GBS, PROMM, FSHD) from PDs of more common causes of chronic pain (CP).

In these four cases, Pain2D classified PDs with an accuracy of ≥61% (Table 2). The most accurate classifier was for GBS vs. CP (77%) and the most inaccurate one for FSHD vs. CP (61%). Overall, the sensitivity achieved by the binary classifier of Pain2D for these four cases was ≥90%, with the best result for EDS vs. CP at 98%. The values for specificity were relatively low at ≥30% (best result for GBS vs. CP at 66%).

	TP	FP	FN	TN	p-value	Acc	Sens	Spec	AUC_{ROC}
PROMM vs.CP	80	24	9	26	< 0.001	0.76	0.90	0.52	0,846 (CI:0.774-0.908)
EDS vs.CP	58	35	1	15	< 0.001	0.67	0.98	0.30	0,899(CI:0.892-0.954)
FSHD vs.CP	34	32	1	18	< 0.001	0.61	0.97	0.36	0,854 (CI:0.77-0.93)
GBS vs.CP	28	17	1	33	< 0.001	0.77	0.96	0.66	0.921(CI:0.853-0.973)

Table 2: Classification results with the binary classifier of Pain2D for each rare disease vs. CP, listing values for true positives (TP), false positives (FP), true negatives (TN), false negatives (FN), p-value (Fisher's exact test), accuracy (Acc), sensitivity (Sens), specificity (Spec), AUC of the ROC curve (AUC_{ROC}). In all cases, the p-values (Fisher's exact test) suggest strongly that the classifier performs much better than random guessing.

AUCs of ROC curves were ≥0.845, with best results for GBS vs. CP at 0.921 (Table 2, Fig. 3). Taken together, these results suggest that the binary classifier can distinguish between the control group and each of the four rare diseases investigated with good sensitivity, but considerably lower specificity.

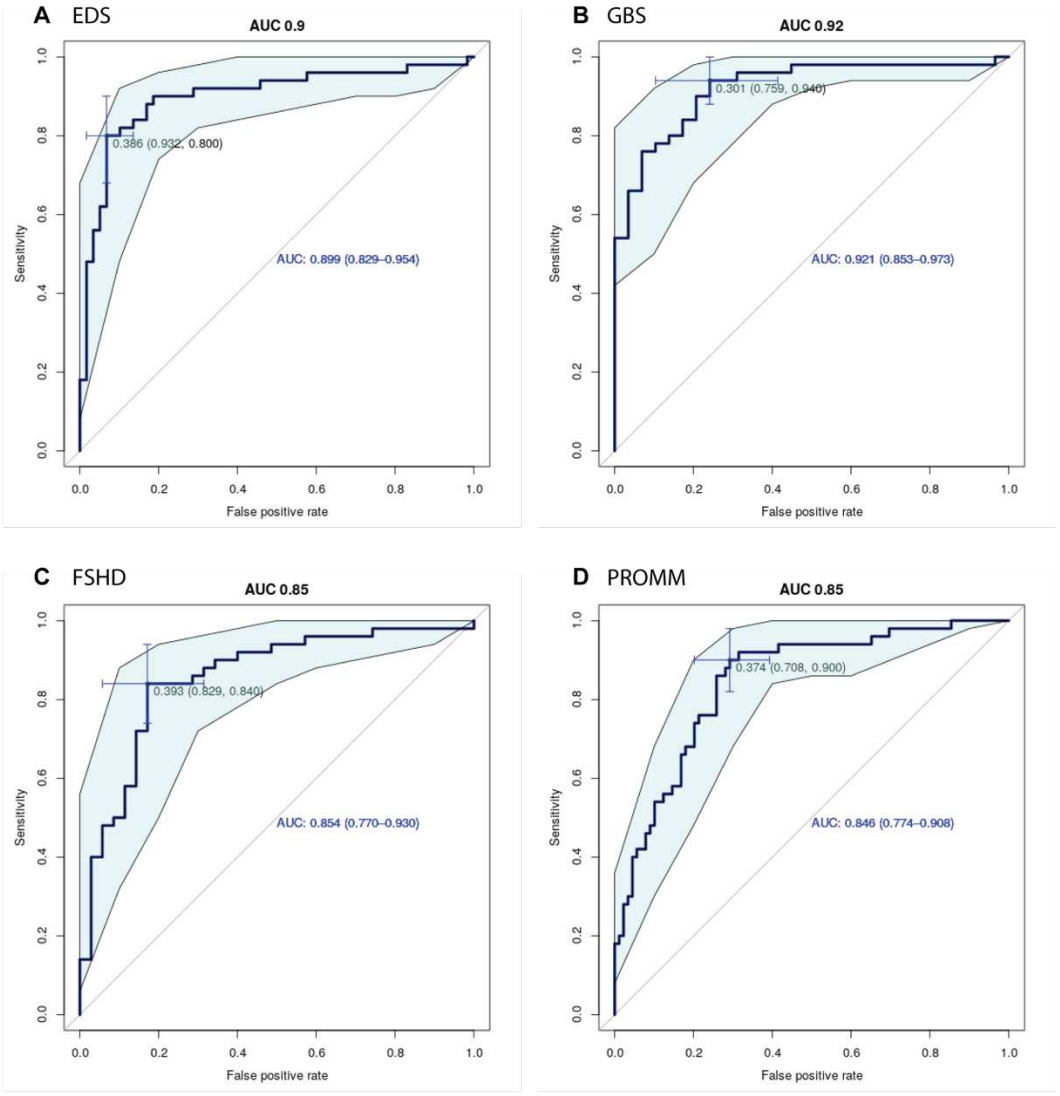


Figure 3: Receiver operating characteristics (ROC) curves for binary classification of each RD vs. CP. A) ROC curve binary classification of EDS and CP. AUC=0.899 (CI:0.829-0.954), B) ROC curve binary classification of GBS and CP. AUC=0.921 (CI:0.853-0.973), C) ROC curve binary classification of FSHD or CP. AUC=0.854 (CI:0.770-0.930), D) ROC curve binary classification of PROMM and CP. AUC=0.846 (CI: 0.774-0.908). Confidence intervals are depicted as light blue band. Blue crosshairs indicate optimal threshold for classification.

The k-disease classifier of Pain2D can classify PDs as PROMM, EDS, FSHD, GBS and chronic pain with varying sensitivity and overall high specificity.

The classification of all PDs with the k-disease classifier of Pain2D gave results with varying sensitivities for the five groups, ranging from 0.51 to 0.83 for the four diseases analyzed. The control group of unspecific chronic pain (CP) was classified with a low sensitivity of only 0.14. Specificity was overall high with values between 0.83 and 0.99 for the five diagnostic groups. In all cases, low p-values of < 0.001 (χ^2-Test) indicated that the k-diseases classifier performed better than random assignment of the diseases.

		True					
		PROMM	EDS	FSHD	GBS	CP	$Sum_{predicted}$
Predicted	PROMM	46	5	17	20	1	89
	EDS	1	38	8	11	1	59
	FSHD	5	4	22	3	1	35
	GBS	3	1	0	25	0	29
	CP	7	13	13	10	7	50
	Sum_{true}	62	61	60	69	10	262
Sensitivity		0.517	0.644	0.629	0.862	0.140	
Specificity		0.908	0.887	0.833	0.811	0.986	

Table 3: Confusion matrix resulting from classification of all PDs with the k-disease classifier of Pain2D.

Pain2D k-disease classification works well for EDS, GBS and FSHD

The classification for GBS, EDS and FSHD gave good results with sensitivities of 86%, 64% and 62%, respectively. Of note, EDS and GBS were previously efficiently classified with the binary classifier of Pain2D as well (sensitivity 86%, specificity 96%; [16]). As a possible explanation, the pain profiles of EDS and GBS each show painful regions that are not present in the other pain profiles, like the knees in EDS and the feet in GBS. These regions are only rarely marked in the other disease patterns, which could contribute to efficient classification. FSHD typically affects the shoulder and the lower back. These regions are also sometimes marked in PDs of other diseases, as well as in the unspecific chronic pain group. Nevertheless, the k-disease classifier of Pain2D was still able to classify FSHD with a sensitivity of 63% and specificity of 83%.

Pain2D k-disease classification of PROMM PDs is less efficient

PROMM PDs were classified correctly only in 52%; the most common mis-classifications were as GBS (23%) and FSHD (19%). Less efficient classification of PROMM could be the result of a number of reasons. The most likely explanation is based on a PD sample that is more heterogeneous than for the other diagnostic groups, for example as a result of the presence of different stages of the disease in the sample, with earlier stages differing with regard to the pain pattern from later stages. To address this, we indeed plan to perform a longitudinal analysis of the pain drawings in different disease stages as a follow up study.

The k-disease classifier of Pain2D achieved a high specificity for PROMM: only 17 non-PROMM PDs were classified as PROMM (false positives), resulting in a specificity of 90 %. It is however impossible to say if this could be attributed to the higher sensitivities for classification of the other rare diseases as a result of their unique features in the pain profile (e.g., painful feet for GBS, etc.), or vice versa, or (most likely) both. Of note, PROMM was

the biggest sample with 88PDs in total, as opposed to only between 29-59 for the other diagnostic groups.

PDs from people with chronic pain are preferentially classified as any rare disease (EDS, GBS, FSHD, PROMM)

One reason why we included PDs from people with unspecific chronic pain in our study was to test if additional, non-specific pain patterns in patients with rare diseases (a “background pain noise”) changed the pain profile enough to interfere with classification. It turned out that the separation between CP and RD PDs was not an issue in our sample, as the number of false positives for the CP group was quite low with only 3 out of 211 rare disease PDs classified as CP (Table 3). In addition, the PDs of the CP diagnostic group were not preferentially classified as one specific rare disease, but were more or less randomly distributed into the five diagnostic groups (classification as PROMM in 14 %, EDS in 26 %, FSHD in 26 % and GBS in 20 %), which is reflected in the low sensitivity of 14% for CP (Table 3). Taken together, PD classification by Pain2D was not hampered in a relevant manner in our sample group by the putative presence of a background pattern in RD pain profiles.

Discussion

Comparison of Pain2D generated pain profiles (PPs) to pain patterns described in the literature

We were able to show that Pain2D is a useful tool to generate disease specific pain patterns and utilize them for automated diagnostic support. The pain pattern of FSHD shows a similar

distribution as described by Morís et al. [34]. The percentages they observed are slightly lower, but the main localizations are the same. The Pain2D generated EDS PP correlates with the description of EDS in the literature, as joint and spinal pain are known typical manifestations of EDS [35]. Pain in distal extremities is a common symptom in GBS described in the literature [36], which fits to the Pain2D generated PP for GBS. Of note, the Pain2D PP shows that the plantar sides of the feet are marked by far more patients than the dorsal side or the distal leg, and thereby adds details to the currently described observations with regard to pain in GBS.

Classification of PROMM PDs was less efficient than for the other diagnostic groups, which seems to be related to more heterogeneous PDs. It is worth mentioning that the PROMM PDs differed with regard to the presence or absence of pain in the upper legs, which might be one factor that hampers classification. Indeed, descriptions of the pain pattern of PROMM vary a lot in the literature. The Pain2D generated PROMM pain profile generally fits to the findings of Eger et al. [37]. In contrast, Peric et al. observed higher pain frequencies in the lower legs than in the upper legs compared to the Pain2D PP [38]. Our results show a higher difference in the frequencies between upper and lower legs compared to prior published results, and we obtained overall lower pain frequencies in nearly all body regions [24]. Taken together, the pain pattern in PROMM seems to differ among patients, which is consistent with more difficult classification. Indeed, its variable clinical manifestation is often discussed as a reason for PROMM being an underdiagnosed disease [31].

Potential Usefulness of Pain2D as a DDSS

The binary classifier of Pain2D achieved a sensitivity of over 90 % for all the diseases we investigated vs. chronic pain, suggesting that Pain2D might be a useful tool to ensure that rare diseases are taken into consideration for the differential diagnosis of unclear pain

manifestations. Furthermore, since the binary classifier performed generally with high accuracy, it might be useful in the differential diagnosis of two similar diseases.

Generally speaking, many rare diseases present with pain as one of the first symptoms prompting patients to see a doctor. For example, pain is described as one of the first symptoms of FSHD [25]. For PROMM, a study of 2013 has shown that leg pain is the first symptom for the disease in 5.2 % of cases, in addition to a similar fraction that presents with general pain at first [39]. As both pain types could potentially be detected by PD, we hope that Pain2D can contribute to the process of diagnosis for these patients in the future. In addition, all Pain2D applications are published as open-source and can be trained to generate specific pain profiles for many, if not most, rare diseases manifesting with pain as a symptom.

The sensitivity of the k-disease classifier of Pain2D ranged between 51% and 83% for our test sample. For example, FSHD classification achieved a sensitivity of 63% and a specificity of 83%, compared to molecular genetic testing for FSHD with a sensitivity of 93% and a specificity of ca. 98% [40]. Given the simplicity and non-invasive nature of generating a pain drawing, a fully trained version of Pain2D with a comprehensive library of pain profiles could therefore indeed efficiently guide physicians in the search for rare diseases by pointing out the more likely candidates in order to initiate specific diagnostic procedures, especially when pain drawings are enhanced with additional patient information, for example from questionnaires. It thus could contribute to shortening the “diagnostic odyssey” that patients with rare diseases often have to endure. In specific cases, it might even develop as a diagnostic indicator in its own right, conceivably in the differentiation of two similar diseases with a clear distinction in pain pattern.

Challenges and future directions

Although Pain2D achieved overall good results with classification of the five tested diagnostic groups, there are some limitations to the tool as of now. Classification of PROMM PDs turned out to be less efficient, in part due to the presence of at least two subgroups of pain patterns in the PP (thighs marked vs. thighs not marked). Since similarity is measured against the PP, which could be understood as an “average PD” of the disease, diseases with more than one “typical” manifestation of pain, or the presence of more complex pain patterns, could potentially be problematic for Pain2D, as classification based on similarity to a PP masks the presence of more than one pattern per disease. However, this approach achieved superior results compared to a nearest neighbor classifier, that was tested during Pain2D development.

As differences in pain patterns might occur over the course of a disease, a longitudinal study to follow disease progression is currently in preparation. In this case, as well as other cases of “more than one pain pattern” per disease, training Pain2D with subgroups could overcome the problem, as long as enough PDs per subgroup are available for training.

Acquiring enough training data for a rare disease classifier is generally a challenge. Not only are the number of people suffering from a specific rare disease limited, but, as of now, ~7000 rare diseases are known, and the number is still growing. Pain2D-Tool therefore needs to be trained for many RDs presenting with pain before it can become a generally useful DDSS: PDs of many RDs need to be sampled and compiled into PPs, which is almost impossible for some ultra-rare diseases with only a few known cases. However, Pain2D can potentially grow over time by integrating data from many researchers/physicians, as it is available as a free and open-source tool.

While classification of PDs with Pain2D-Tool worked overall well with our test data set, one has to take into account that the latter is not a representative sample of patients presenting with pain without a diagnosis. Prevalences of rare diseases are by definition low, with consequences for classification results in a clinical setting. For example, the k-disease classifier preferentially mis-classified PDs from the CP group as rare (43 of 50), although the latter are far more often encountered in practice. As a result, practically all PDs in a realistic setting will be classified as rare (of which only a small percentage is truly rare), and the positive predictive value will be rather low. In its current form Pain2D is therefore not useful to distinguish between rare and non-rare causes of pain in general. It is furthermore unclear how the introduction of further rare diseases into Pain2D (i.e., increasing k) will influence classification efficiency. However, since Pain2D is a non-invasive and cost-efficient software, its strength is the suggestion of possible diagnoses for further testing in order to abbreviate the diagnostic odyssey patients with rare diseases often have to endure, and we consider higher numbers of false positives from Pain2D to be an acceptable intermediate step towards this goal.

Compared to our initial results with Pain2D [16], this study has two important innovations. Firstly, we were able to show that Pain2D is indeed able to distinguish between two related diseases of the same category (neuromuscular disorders; PROMM and FSHD). Secondly, we successfully used the new k-disease classifier of Pain2D for the classification of five diagnostic groups, including four rare diseases. Taken together, our study could show that Pain2D has the potential to develop into a full DDSS for pain-associated diseases with a focus on rare diseases, and opens up the route towards further exploitation of pain symptoms for AI-assisted diagnosis of rare diseases.

Declarations

Ethics approval and consent to participate

The study was approved by the Ethics Committee of the Rheinische Friedrich-Wilhelm-Universität, Medizinische Fakultät, on May 30th, 2017. All patients gave informed, written consent to participate.

Consent for publication

Not applicable.

Funding

No third-party funding.

Availability of data and materials

All data discussed are included with the published article.

Competing interests

The authors report no conflicts of interest in this work.

Authors' contributions

Conceptualization: LG, FK, RC, MM, JS; Methodology: FK, RC, MM, JS; Software: NS; Formal analysis: NS, FK; Investigation: DE, NS; Resources: DE, LG, RC, MM, JS; Writing – original draft preparation: DE, NS, RC, MM, JS; Writing – review and editing: DE, NS, FK, RC, MM, JS; Visualization: NS, JS; Supervision: FK, RC, MM, JS; Project administration: FK, RC, MM, JS.

Acknowledgments

We are indebted to all participants for their consent and co-operation.

References

1. About Rare Diseases | www.eurordis.org. https://www.eurordis.org/about-rare-diseases. Accessed 9 Feb 2022

2. FAQs About Rare Diseases | Genetic and Rare Diseases Information Center (GARD) – an NCATS Program. https://rarediseases.info.nih.gov/diseases/pages/31/faqs-about-rare-diseases. Accessed 9 Feb 2022

3. Stieber C, Mücke M, Windheuser IC, Grigull L, Klawonn F, Tunc S, Münchau A, Klockgether T (2017) [On the fast track to diagnosis : Recommendations for patients without a diagnosis]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 60:517–522

4. Schaefer J, Lehne M, Schepers J, Prasser F, Thun S (2020) The use of machine learning in rare diseases: a scoping review. Orphanet J Rare Dis 15:145

5. Faviez C, Chen X, Garcelon N, Neuraz A, Knebelmann B, Salomon R, Lyonnet S, Saunier S, Burgun A (2020) Diagnosis support systems for rare diseases: a scoping review. Orphanet J Rare Dis 15:94

6. Schaaf J, Sedlmayr M, Schaefer J, Storf H (2020) Diagnosis of Rare Diseases: a scoping review of clinical decision support systems. Orphanet Journal of Rare Diseases 15:263

7. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI (2020) An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digit Med 3:17

8. Berner ES (2006) Diagnostic Decision Support Systems: Why Aren't They Used More And What Can We Do About It? AMIA Annu Symp Proc 2006:1167–1168

9. Palmer H (1949) Pain charts; a description of a technique whereby functional pain may be diagnosed from organic pain. N Z Med J 48:187–213

10. Shaballout N, Neubert T-A, Boudreau S, Beissner F (2019) From Paper to Digital Applications of the Pain Drawing: Systematic Review of Methodological Milestones. JMIR mHealth and uHealth 7:e14569

11. Egloff N, Gander M, Cámara R, Klingler N, Wegmann B, Marti E, von Känel R (2011) Pain Drawings Help to Distinguish Between Somatic and Somatoform Pain. PPmP - Psychotherapie · Psychosomatik · Medizinische Psychologie. https://doi.org/10.1055/s-0031-1272370

12. Mann NH, Brown MD (1991) Artificial intelligence in the diagnosis of low back pain. Orthop Clin North Am 22:303–314

13. Hüllemann P, Keller T, Kabelitz M, Freynhagen R, Tölle T, Baron R (2017) Pain Drawings Improve Subgrouping of Low Back Pain Patients. Pain Pract 17:293–304

14. Tachibana T, Maruo K, Inoue S, Arizumi F, Kusuyama K, Yoshiya S (2016) Use of pain drawing as an assessment tool of sciatica for patients with single level lumbar disc herniation. SpringerPlus 5:1312

15. Rennerfelt K, Zhang Q, Karlsson J, Styf J (2018) Patient pain drawing is a valuable instrument in assessing the causes of exercise-induced leg pain. BMJ Open Sport Exerc Med 4:e000262

16. Wester L, Mücke M, Bender TTA, Sellin J, Klawonn F, Conrad R, Szczypien N (2020) Pain drawings as a diagnostic tool for the differentiation between two pain-associated rare diseases (Ehlers-Danlos-Syndrome, Guillain-Barré-Syndrome). Orphanet Journal of Rare Diseases 15:323

17. Zhou Z, Rewari A, Shanthanna H (2018) Management of chronic pain in Ehlers-Danlos syndrome: Two case reports and a review of literature. Medicine (Baltimore) 97:e13115

18. Germain D-P (2017) [Ehlers-Danlos syndromes]. Ann Dermatol Venereol 144:744–758

19. Peña L, Moreno CB, Gutierrez-Alvarez AM (2015) Pain management in Guillain–Barre syndrome: A systematic review. Neurología (English Edition) 30:433–438

20. Fujimura H (2013) The Guillain-Barré syndrome. Handb Clin Neurol 115:383–402

21. Leonhard SE, Mandarakas MR, Gondim FAA, et al (2019) Diagnosis and management of Guillain–Barré syndrome in ten steps. Nature Reviews Neurology 15:671–683

22. Liu S, Dong C, Ubogu EE (2018) Immunotherapy of Guillain-Barré syndrome. Hum Vaccin Immunother 14:2568–2579

23. Statland J, Tawil R (2014) Facioscapulohumeral Muscular Dystrophy. Neurologic Clinics 32:721–728

24. Deenen JCW, Arnts H, van der Maarel SM, Padberg GW, Verschuuren JJGM, Bakker E, Weinreich SS, Verbeek ALM, van Engelen BGM (2014) Population-based incidence and prevalence of facioscapulohumeral dystrophy. Neurology 83:1056–1059

25. Bushby KMD, Pollitt C, Johnson MA, Rogers MT, Chinnery PF (1998) Muscle pain as a prominent feature of facioscapulohumeral muscular dystrophy (FSHD): four illustrative case reports. Neuromuscular Disorders 8:574–579

26. Hamel J, Tawil R (2018) Facioscapulohumeral Muscular Dystrophy: Update on Pathogenesis and Future Treatments. Neurotherapeutics 15:863–871

27. Turner C, Hilton-Jones D (2010) The myotonic dystrophies: diagnosis and management. Journal of Neurology, Neurosurgery & Psychiatry 81:358–367

28. Mahyera AS, Schneider T, Halliger-Keller B, Schrooten K, Hörner E-M, Rost S, Kress W (2018) Distribution and Structure of DM2 Repeat Tract Alleles in the German Population. Front Neurol. https://doi.org/10.3389/fneur.2018.00463

29. Vanacore N, Rastelli E, Antonini G, et al (2016) An Age-Standardized Prevalence Estimate and a Sex and Age Distribution of Myotonic Dystrophy Types 1 and 2 in the Rome Province, Italy. Neuroepidemiology 46:191–197

30. Suominen T, Bachinski LL, Auvinen S, Hackman P, Baggerly KA, Angelini C, Peltonen L, Krahe R, Udd B (2011) Population frequency of myotonic dystrophy: higher than expected frequency of myotonic dystrophy type 2 (DM2) mutation in Finland. European Journal of Human Genetics 19:776–782

31. Meola G, Cardani R (2017) Myotonic dystrophy type 2 and modifier genes: an update on clinical and pathomolecular aspects. Neurological Sciences 38:535–546

32. Suokas KI, Haanpää M, Kautiainen H, Udd B, Hietaharju AJ (2012) Pain in patients with myotonic dystrophy type 2: A postal survey in finland. Muscle Nerve 45:70–74

33. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, Müller M (2011) pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 12:77

34. Morís G, Wood L, FernáNdez-Torrón R, et al (2018) Chronic pain has a strong impact on quality of life in facioscapulohumeral muscular dystrophy. Muscle Nerve 57:380–387

35. Gazit Y, Jacob G, Grahame R (2016) Ehlers-Danlos Syndrome-Hypermobility Type: A Much Neglected Multisystemic Disorder. Rambam Maimonides Med J. https://doi.org/10.5041/RMMJ.10261

36. Moulin DE, Hagen N, Feasby TE, Amireh R, Hahn A (1997) Pain in Guillain-Barré syndrome. Neurology 48:328–331

37. Eger K, Schulte-Mattler WJ, Zierz S (1997) Proximale myotone Myopathie (PROMM)Klinische Variabilität innerhalb einer Familie. Nervenarzt 68:839–844

38. Peric M, Peric S, Rapajic N, Dobricic V, Savic-Pavicevic D, Nesic I, Radojicic S, Novakovic I, Lavrnic D, Rakocevic-Stojanovic V (2015) Multidimensional aspects of pain in myotonic dystrophies. Acta Myol 34:126–132

39. Hilbert JE, Ashizawa T, Day JW, Luebbe EA, Martens WB, McDermott MP, Tawil R, Thornton CA, Moxley RT (2013) Diagnostic odyssey of patients with myotonic dystrophy. J Neurol 260:2497–2504

40. Tawil R, Kissel JT, Heatwole C, Pandya S, Gronseth G, Benatar M (2015) Evidence-based guideline summary: Evaluation, diagnosis, and management of facioscapulohumeral muscular dystrophy. Neurology 85:357–364