KÜNSTLICHE INTELLIGENZ ZUR DIAGNOSTISCHEN UNTERSTÜTZUNG AUSGEWÄHLTER SELTENER LYSOSOMALER SPEICHERERKRANKUNGEN: ERGEBNISSE EINER PILOTSTUDIE

(ARTIFICIAL INTELLIGENCE FOR DIAGNOSTIC SUPPORT IN SELECTED RARE LYSOSOMAL STORAGE DISORDERS: RESULTS OF A PILOT STUDY)

**A.-L. SIEG, A.-M. DAS, N.M. MUSCHOL, A. KÖHN, C. LAMPE, X. KORTUM, S.
MEHMECKE, S. BLÖSS, W. LECHNER, F. KLAWONN, L. GRIGULL**

EINLEITUNG

Lysosomale Speicherkrankheiten (LSDs) sind seltene Stoffwechselerkrankungen, denen eine angeborene Enzymdefizienz oder ein Transporterdefekt zugrunde liegt. Dies führt zu einer Akkumulation von nicht abbaubaren Metaboliten in Körperzellen. Betroffene kommen in der Regel klinisch gesund zur Welt und entwickeln mit ganz unterschiedlicher Dynamik im Laufe der Zeit unspezifische Symptome [1, 7, 23]. Entsprechend entstehen heterogene Krankheitsbilder, die eine Diagnosestellung zur Herausforderung für Betroffene und behandelnde Ärzte machen [8, 14]. Die Symptome und Beschwerden der LSDs reichen von psychomotorischer Entwicklungsverzögerung, unspezifischen Gelenkproblemen und rezidivierenden Infektionen des HNO-Traktes bei Kindern mit Mucopolysaccharidose (MPS) über Störungen des blutbildenden Systems sowie Knochenschmerzen bei Patienten mit M. Gaucher bis hin zu Gliederschmerzen, kardialen Funktionsstörungen und Nierenversagen bei Betroffenen mit M. Fabry [1, 7, 24].

Gemeinsam ist den genannten Erkrankungen, neben dem zugrundeliegenden Pathomechanismus mit Speicherung endogener Substrate, dass sie zum Teil erst nach einer sehr großen Latenz erkannt werden. Ein möglicher Grund für diagnostische Odysseen von Betroffenen ist – neben heterogenen Symptomkonstellationen – auch fehlende Kenntnisse der Ärzte für LSDs [4, 5, 9, 13,

17]. Frühzeitiger Therapiebeginn kann die Lebensqualität der Patienten verbessern und Organschäden entgegen wirken [8, 9, 10, 12, 17,26, 27, 30].

In der vorliegenden Pilotstudie wurde daher der Einsatz von künstlicher Intelligenz (KI) für die diagnostische Unterstützung bei drei LSDs mit ganz unterschiedlichen klinischen Symptomen getestet, bei denen zum Teil eine Behandlungsoption besteht..

KI wird zunehmend für die diagnostische Unterstützung erprobt und hat in der Pathologie, Radiologie und Dermatologie schon gute Ergebnisse erzielt [29-31]. Die Kombination aus elterlichen Fragebögen und KI haben wir bereits für ausgewählte pädiatrisch-pulmonologische Krankheitsbilder und primäre Immundefekte bei Kindern erfolgreich erprobt [17, 22].

Ziel der Arbeit war es, aus elterlichen Erfahrungen einen diagnostischen Fragebogen für ausgewählte LSDs zu entwickeln. Dieser Fragebogen sollte dann, nachdem er von Menschen mit bekannter Diagnose beantwortet wurde, pilothaft zur diagnostischen Unterstützung erprobt werden. Hierfür wurden Verfahren aus dem Bereich der künstlichen Intelligenz (KI) eingesetzt, die Muster in den Antworten zur diagnostischen Klassifizierung nutzen.

METHODIK

Zur Erfassung der Daten wurden Interviews mit Betroffenen und Eltern von erkrankten Kindern durchgeführt. Ein Interview (Interview 8 – MPS I) aus einem Vorgängerprojekt lag bereits vor. Von den insgesamt 17 Interviews entfielen neun Interviews auf die Diagnosegruppe MPS, und je vier Interviews auf die Diagnosegruppen M. Fabry und M. Gaucher. Dabei wurden 12 der 17 Interviews mit Eltern betroffener Kinder geführt und drei Interviews mit erwachsenen Betroffenen. Der Studienablauf ist in der Abbildung 1 schematisch dargestellt. Zwei Interviewpartner waren sowohl Eltern eines betroffenen Kindes und auch selbst betroffen. Eine Übersicht zu den Interviewpartnern bzw. den Betroffenen gibt die Tabelle 1. Die Interviews wurden anschließend qualitativ analysiert und es wurden aus den Ergebnissen zwei Fragebögen entwickelt. In die Datenerhebung wurden nur

Patienten bzw. deren Eltern aufgenommen, bei denen eine (enzymatisch) gesicherte Diagnose vorlag.

Als Kontrollgruppe dienten Patienten / Familien, die wegen Symptomen vorstellig wurden, die auch im Rahmen der Zielerkrankungen zu beobachten sind. Diese Kontrollpatienten mit z.B. unklaren Schmerzen, Entwicklungsstörung, Durchfällen, rezidivierende Infektionen im HNO-Bereich wurden aus den Ambulanzen der Kinderklinik rekrutiert.

Zur Datenerhebung wurde das in der empirischen qualitativen Sozialforschung genutzte narrative Interview gewählt. Dieses von Schütze entwickelte Verfahren legt den Fokus auf die subjektive Sichtweise und Erfahrung des zu befragenden Individuums [2]. Es wurden unter Berücksichtigung der ‚theoretischen Sättigung‘ [25] insgesamt 11 Interviews nach Colaizzi qualitativ analysiert. Dies bedeutet, dass Datenerhebung und Auswertung parallel erfolgten, sodass anhand der Ergebnisse der Bedarf an neuen Interviews festgemacht werden konnte. Zu dem Zeitpunkt, an dem es keinen relevanten Zuwachs an neuen Erkenntnissen für die Forschungsfrage mehr gab, wurde die Interviewanalyse abgeschlossen.

In Vorbereitung auf die Interviewtermine wurden einheitliche Eingangsfragen festgelegt, um die Vergleichbarkeit zwischen den Interviews zu wahren:

„Frau/Herr (Name), ich würde Sie nun bitten, mir über die Lebensphase Ihres Kindes zu berichten bevor Ihr Kind die Diagnose (Krankheit) bekommen hat. Erzählen Sie gerne von all den Phänomenen, die Sie vom ersten Moment an beobachtet haben.“

(Interview mit Eltern betroffener Kinder)

Die vorliegende Studie nutzte zur qualitativen Auswertung der Interviews das Verfahren nach Colaizzi. Colaizzi gilt als Vertreter der deskriptiven Phänomenologie, seine Methodik ist vielfach aufgegriffen und im gesundheitswissenschaftlichen Kontext angewandt worden [8, 10, 13]. Das Ziel dieser Analyseform besteht darin, die Erfahrung bzw. das Erleben der Interviewpartner in den Vordergrund zu stellen. Ziel der schrittweisen Analyse war dabei, prädiagnostische Phänomene aus Sicht der Betroffenen herauszuarbeiten, daraus ein System von Kategorien zu erstellen. Ergänzt wurde die von Colaizzi entwickelte Methodik in der vorliegenden Arbeit,

indem die Kernaussagen der Kategorien am Ende in Form eines Fragebogens abgebildet wurde.

Die entstandenen Fragebögen wurden von bereits diagnostizierten Betroffenen und einer Kontrollgruppe retrospektiv beantwortet, so als hätten sich die Teilnehmer in der Zeit vor der Diagnosestellung befunden.

Im nächsten Schritt sollten aus den beantworteten Fragebögen charakteristische Antwortmuster mit so genannten Data Mining Verfahren identifiziert werden. Data Mining bezeichnet als Überbegriff mathematische Klassifikationssysteme, die in der Lage sind, Muster in großen Datenmengen zu erkennen. In der vorliegenden Arbeit wurden vier verschiedenen Klassifikatoren (Support Vektor Maschine (SVM), Random Forest (RF), logistische Regression (LR) und lineare Diskriminanzanalyse (LD)) in Kombination verwendet.

Die genannten Klassifikatoren wurden mit 56 Fragebögen (27 Fragebögen mit der Diagnose ‚MPS‘ und 29 Fragebögen aus der Kontrollgruppe) trainiert, wobei mit Blick auf eine ganzzahlige Aufteilung eine 8-fache Kreuz-Validierung ($56/8 = 7$) gewählt wurde. Weitere 20 Fragebögen, die ganz bewusst nicht Teil der Trainingsmenge waren, dienten als erster prospektiver Test für die Data Mining Verfahren.

Das vorliegende Projekt wurde von den Ethikkommissionen der Medizinischen Hochschule Hannover geprüft und genehmigt. Alle Teilnehmer gaben ihr informiertes Einverständnis, im Fall von Minderjährigen erfolgte dies durch die Eltern. Die Studie wurde entsprechend den Prinzipien der WHO formuliert und nach den Empfehlungen der Deklaration von Helsinki durchgeführt.

ERGEBNISSE

Die Interview-Partner berichteten, dass die diagnostische Latenz, also die Zeit von den ersten Symptomen bis zur Diagnosestellung der entsprechenden LSD zwischen 10 Monaten (M. Gaucher) und 57 Jahren (M. Fabry) lag. Zum Teil wurden von den Interviewpartnern Angaben zu Fehldiagnosen gemacht, wobei sich unter diesen Angaben Diagnosen wie die Rheumatoide Arthritis bei M. Fabry, Wachstumsschmerzen bei M. Gaucher oder eine Kuhmilchintoleranz bei MPS III befanden (Tabelle 1).

Eine Übersicht der Interviewpartner bzw. deren Kinder / Betroffene ist in Tabelle 1 dargestellt.

Tabelle 1. Details zu Interviewpartnern bzw. Betroffenen

Interview-nummer	Diagnose	Diagnostische Latenz	Fehldiagnosen	Untersuchungen vor Diagnosestellung
1	MPS III (b)	6 Jahre 6 Monate	Milchunverträglichkeit	Mehrfache Blut- und Urinuntersuchungen, Sonografie, körperliche Untersuchungen
2	MPS III (a)	1 Jahr 7 Monate	k.A.	Magnetresonanztomografie, Hals-, Nasen-, Ohrenärztliche Untersuchung (Rachenmandelentfernung), Röntgenaufnahmen, Blut- und Urinuntersuchungen
3	MPS II (+I), Frühgeburt (29.SSW)	2 Jahre 10 Monate	k.A.	Hals-, Nasen-, Ohrenärztliche Untersuchung, Sonografien
4	MPS III (a)	12 Jahre	V.a. fragiles X-Syndrom	Intelligenz- und Entwicklungstests, Blut- und Urinuntersuchungen
5	MPS VI	2 Jahre 9 Monate	k.A.	Augenärztliche Untersuchung
6	MPS II	2 Jahre	k.A.	k.A.
7	MPS VI (a)	3 Jahre 10 Monate	k.A.	Orthopädische Untersuchung, körperliche Untersuchung, humangenetische Untersuchungen
8	MPS I	k.A.	k.A.	k.A.
9	M. Fabry	57 Jahre	k.A.	Nach Schlaganfall: Magnetresonanztomografie, Nieren- und Herzultraschall, Blutuntersuchung, neurologische Untersuchung
10	M. Gaucher	5 Jahre	k.A.	Blutuntersuchung, Sonografie
11	M. Fabry	51 Jahre	k.A.	Magnetresonanztomografie, Computertomografie, Sonografie der Halsgefäße, Schluckechokardiografie

12	M. Fabry	9 Jahre 6 Monate	Rheuma, Allergien, Wachstumsschmerzen, Borreliose	Lumbalpunktion, Knochenmarkspunktion, Blutuntersuchung, Magnetresonanztomografie, Sonografie, Kniegelenkspunktion
13	M. Gaucher	10 Monate	k.A.	Blutuntersuchung, körperliche Untersuchung, Sonografie
14	M. Fabry	4 Jahre	k.A.	Augenärztliche Untersuchung, Blutuntersuchung, Sonografie Abdomen
15	M. Gaucher	11 Jahre	Wachstumsschmerzen, Leukämie	Knochenmarkspunktion, Blutuntersuchung, Sonografie, Röntgenaufnahmen, Magnetresonanztomografie
16	M. Gaucher	23 Jahre	k.A.	k.A.
17	M. Gaucher	8 Jahre	k.A.	k.A.

Tab. 1. Legende: k.A. keine Angaben

Die Interview-Partner berichteten, dass die diagnostische Latenz, also die Zeit von den ersten Symptomen bis zur Diagnosestellung der entsprechenden LSD zwischen 10 Monaten (M. Gaucher) und 57 Jahren (M. Fabry) lag. Zum Teil wurden von den Interviewpartnern Angaben zu Fehldiagnosen gemacht, wobei sich unter diesen Angaben Diagnosen wie die Rheumatoide Arthritis bei M. Fabry, Wachstumsschmerzen bei M. Gaucher oder eine Kuhmilchintoleranz bei MPS III befanden (Tabelle 1).

Im Rahmen der qualitativen Analyse konnten aus den Interviews insgesamt 53 Kernaussagen extrahiert werden, die zur Fragengenerierung dienten. In der Tabelle 2 werden exemplarisch inhaltliche Aspekte dieser Kernaussagen dargestellt.

Tabelle 2. Zusammenstellung ausgewählter Kernaussagen aus den Interviews

Kernaussage	Zitate
Körperliche Belastung der Eltern	„Und ich hatte gerade, auch durch

	<p><i>diesen ganzen Aufwand auch für den S., für mich keine Energie mehr.“</i> (Interview 12 – Mutter eines Kindes mit M. Fabry)</p> <p><i>„Wobei es sehr schwer aushaltbar war, einfach die hat maximal zwei Stunden am Stück geschlafen. Und das hat sich hingezogen so ca. eineinhalb Jahre.“</i> (Interview 1 – Mutter eines Kindes mit MPS I)</p>
Emotionalität der Eltern	Einige Eltern berichten davon, dass sie mit der „Belastung dieses Unwahrscheinlichen“ (Interview 15 – Eltern eines Kindes mit M. Gaucher), der Ungewissheit über den gesundheitlichen Zustand ihres Kindes zu kämpfen hatten
Kontakt mit Ärzten	Eltern und Betroffene fühlten sich von Ärzten häufig nicht ernst genommen. Ihre Sorgen seien „ <i>abgetan</i> “ (Interview 12 – Kind leidet an M. Fabry) und „ <i>schön geredet</i> “ (Interview 7 – Mutter eines Kindes mit MPS IV) worden. Einige Eltern fühlten sich von Ärzten mit Aussagen wie „ <i>jetzt sind sie mal nicht so ungeduldig</i> “ (Interview 1 – Kind hat MPS III) oder „ <i>das wird sich schon wegwachsen</i> “ (Interview 8 – Kind wurde mit MPS I diagnostiziert) ohne „richtige Diagnose“ abgewiegelt.
Reaktionen des Umfelds	Manche Betroffene berichten, dass sie aufgrund ihrer geringen körperlichen Leistungsfähigkeit als Simulanten

	<p>abgestempelt wurden (z.B. Interview 14 – Betroffene mit M. Fabry). Auch Eltern erzählten, es nehme sie <i>„halt wirklich keiner ernst, wenn man diese ganzen komplexen Sachen sagt: Er hat mal da Schmerzen, er hat dies, er hat das“</i> (Interview 12 – Mutter eines Kindes mit M. Fabry). Außerdem wurde von Eltern mit an MPS erkrankten Kindern beschrieben, dass sie <i>„ganz oft so von, von der Außenwelt so taxiert worden“</i> (Interview 1 – Mutter eines Kindes mit MPS III) seien. Solche Situationen spitzten sich vereinzelt bis zur sozialen Ausgrenzung zu.</p>
--	---

Die aus den Kernaussagen entstandenen Fragebögen richten sich mit 33 Fragen an Eltern betroffener Kinder bzw. mit 40 Fragen an erwachsene Betroffene. Die unterschiedliche Fragenanzahl erklärt sich zum einen durch die Überarbeitung der Fragen bezüglich redundanten Inhalten und bezüglich der Aussagekraft einzelner Fragen für den Gesamtzusammenhang, zum anderen in der Tatsache, dass nicht alle Kernaussagen bzw. Fragen für beide Fokusgruppen (Eltern oder Erwachsene) geeignet waren.

Sammeln der Fragebögen

Die Fragebögen wurden anschließend über Selbsthilfegruppen und kooperierende Kliniken an Menschen mit bereits bekannter Diagnose ausgegeben. Betroffene bzw. deren Eltern wurden gebeten, den Fragebögen aus Sicht der Zeit **vor der Diagnosestellung** zu beantworten. Insgesamt konnten zum Zeitpunkt der Auswertung über diese Verteilungswege 37 Elternfragebögen mit der Diagnose ‚MPS‘, 2 mit der Diagnose ‚M. Gaucher‘ und kein Elternfragebögen mit der Diagnose ‚M. Fabry‘ gesammelt werden. Zudem lagen 11 Erwachsenenfragebögen mit der Diagnose ‚M. Gaucher‘, 5 mit der Diagnose ‚MPS‘ und 2 mit der Diagnose ‚M. Fabry‘ vor. 57 Elternfragebögen konnten in verschiedenen Ambulanzen der zugeordneten

Universitäts-Kinderklinik für die Kontrollgruppe gesammelt werden. Davon kamen für die Diagnose ‚MPS‘ differentialdiagnostisch 39 Kontrollfragebögen in Betracht, die z.B. Patienten mit allgemeiner Entwicklungsverzögerung, Erkrankungen des Skeletts (z.B. Hüftdysplasie), Hepatosplenomegalie und rezidivierenden Infektionen einschlossen.

Für die weitere Auswertung der gesammelten Fragebögen mithilfe von Data Mining Verfahren musste ein ausreichend großer Datensatz vorliegen, was nur für die Eltern-Fragebögen mit der Diagnose ‚MPS‘ der Fall war. Daher beschränkt sich die nachfolgende Darstellung der weiteren Ergebnisse auf die Diagnose ‚MPS‘ und somit auf 37 MPS- sowie 39 Kontroll-Fragebögen. Aus diesem Datensatz von $37+39=76$ Fragebögen wurden je 10 Fragebögen nach dem Zufallsprinzip für den prospektiven Datensatz abgespalten. Somit ergaben sich für die Auswertung zwei endgültige Datensätze, die aus $27+29=56$ retrospektiven und aus $10+10=20$ prospektiven Fragebögen bestanden.

Retrospektive Auswertung der Fragebögen mittels Data Mining Verfahren

In der Ergebnisanalyse zeigte sich, dass die Kombination der eingesetzten Klassifikatoren (‚Fusion‘) mit 91% in der Kreuz-Validierung den Fragebögen eines MPS Patienten von einem Fragebögen der Kontrollgruppe unterscheiden konnte (Abb. 2). Hierbei fanden sich zum Teil erhebliche Unterschiede in der Ergebnisqualität der eingesetzten Klassifikatoren. Extreme Schwankungen und ein deutlich schlechteres Klassifikationsergebnis von 75% korrekter Ergebnisse wies die LD auf. Das solideste Ergebnis der einzelnen Klassifikatoren zeigten der SVM- und der RF-Klassifikator mit jeweils einem Wert für die erzielte Anzahl an korrekten Diagnosen von 91% für die Diagnose ‚MPS‘. Die dreidimensionale grafische Darstellung der beantworteten Fragebögen in Abb. 3 illustriert unterschiedliche Antwortmuster, die mittels Data Mining Verfahren räumlich getrennt werden. Die Abbildung 3 zeigt, wie sich die gesammelten Antwortmuster zwischen Eltern mit an MPS erkrankten Kindern und den Eltern der Kontrollgruppe unterscheiden. Die Qualität der Data Mining Verfahren lässt sich mittels ‚Receiver Operating Characteristic‘-Kurven (ROC-Kurven) und der dazu gehörenden AUC-Flächen (‚Area Under the Curve‘) visualisieren (Abb. 4). Die ROC-Kurve des ‚Fusion‘-Klassifikators, bewegt sich deutlich oberhalb der Diagonalen und weist somit auf korrekt

interpretierte Fragebögen hin. Der hohe AUC-Wert von 95% stellt ein Qualitätsmaß für den ‚Fusion‘-Klassifikator dar und bestätigt die guten Ergebnisse der Klassifizierung.

Erster prospektiver Test der Data Mining Verfahren

Ein erster vorläufiger prospektiver Test mit 20 Fragebögen, die nie Teil der Trainingsmenge waren, lieferte als Ergebnis 18 korrekte Diagnosen, was einer Sensitivität von 90% entspricht. Obwohl sich das vorliegende Ergebnis nur auf die relativ kleine Stichprobe von 20 Fragebögen bezieht, weist es doch auf das Potenzial hin, das in Data Mining Verfahren zur unterstützenden Diagnostik von MPS liegt.

Bestimmung relevanter Einzelfragen

Ergänzend erfolgte eine Sub-Analyse dahingehend, welche Fragen besonders starken Einfluss auf die diagnostische Zuordnung des Computer-Systems besitzen. Der ‚Exakte Fisher-Test‘ [28] lieferte dabei die 5 Fragen mit den kleinsten p-Werten hinsichtlich der Zuordnung zur Diagnose ‚MPS‘:

1. Wurde bei Ihrem Kind eine vergrößerte Leber und/oder Milz festgestellt?
2. Haben Sie den Eindruck, dass Ihr Kind sich in manchen Bereichen langsamer entwickelt als andere Kinder (z.B. sprachliche Entwicklung, Bewegungsentwicklung, Gang zur Toilette etc.)?
3. Haben Sie das Gefühl, dass Ihr Kind „anders“ ist als andere Kinder?
4. Trifft es zu, dass Sie viel Engagement aufbringen müssen, um die normale Entwicklung (z.B. Laufen, Treppen steigen, einfache Koordination, Sprache etc.) Ihres Kindes voran zu bringen?
5. Bedrückt Sie starke Ungewissheit bezüglich der Gesundheit Ihres Kindes?

DISKUSSION

Es wurde in dieser Studie untersucht, ob ein auf Fragebögen basierendes und computergestütztes Werkzeug, welches aus Interviews mit Eltern und Betroffenen

generiert wurde und unterschiedliche Verfahren der künstlichen Intelligenz (KI) kombiniert, potenziell zur Diagnoseunterstützung ausgewählter seltener lysosomaler Speichererkrankungen (M. Gaucher, M. Fabry, MPS) anwendbar ist. Aufgrund mangelnden Fragebögen-Rücklauf beschränkte sich die Analyse auf MPS.

Die Erstellung eines Fragebogens zur Diagnoseunterstützung ausgewählter LSDs erfolgte anhand qualitativer Forschungsmethoden. Die Ergebnisse der vorliegenden Studie zeigen, dass durch die Inhaltsanalyse prädiagnostische Phänomene auffindbar sind, die in dieser Form auch auf andere Personen und Situationen übertragen werden können [3, 11, 25]. Gemeinsamkeiten betreffen unter anderem die Wahrnehmung von Symptomen, die Schwierigkeiten, denen Betroffene im Alltag ausgesetzt sind und den Umgang mit ihrer gesundheitlichen Situation.

Zusätzlich zu den hier genannten Überschneidungen zu anderen Studien konnten im vorliegenden Projekt weitere detaillierte Erkenntnisse über die emotionale Situation von Eltern und Betroffenen aus der prädiagnostischen Phase gesammelt werden. Zudem konnte das Verhältnis zwischen Eltern bzw. Patienten und Ärzten genauer dargestellt und die Berührungspunkte mit dem Gesundheitssystem beleuchtet werden. Ebenfalls wurden körperliche Phänomene analysiert und herausgearbeitet. In bisherigen Studien zu LSDs oder SEs lag der Schwerpunkt entweder auf körperlichen Symptomkomplexen [1, 7, 24] oder auf dem Erleben Betroffener und deren Angehörigen [3, 11, 25]. Durch die offene Interviewstruktur berichteten die Interviewpartner sowohl von körperlichen Symptomen als auch von deren Erleben der prädiagnostischen Situation. Dies bewirkte eine Auflösung der inhaltlichen Trennung beider Komponenten, sodass die entstandenen Fragebögen sowohl körperliche Phänomene, als auch Phänomene zur Wahrnehmung und zum Erleben der Situation beinhalten.

Zum Zwecke der Diagnoseunterstützung wurden die gesammelten Fragebögen mit gesicherter Diagnose durch KI - Verfahren ausgewertet. Die genannten Verfahren sind ursprünglich für die Analyse großer Datenmengen entwickelt worden [20], können jedoch auch für kleinere Datensätze angewandt werden, deren komplexe Struktur (z.B. große Anzahl an Fragen mit mehreren Antwortmöglichkeiten in einem Fragebogen) eine computergestützte Auswertung notwendig machen [13, 18, 23]. Horowitz et al. entwarfen ein auf Fragebögen basierendes Werkzeug, welches in der Lage ist, Patienten mit einer gastroösophagealen Refluxerkrankung (GERD) in einer

Gruppe von Betroffenen mit gastrointestinalen Oberbauchbeschwerden zu finden [12]. Die Kontrollgruppe bestand in der zitierten Studie aus Menschen mit zur Zielerkrankung ähnlichen Symptomen. Entsprechend der höheren diagnostischen Herausforderung liegt die Sensitivität des Werkzeugs von Horowitz et al. – je nach verwendetem Klassifikator – bei maximal 78% [13]. Für das vorliegende Verfahren bewegt sich die ROC-Kurve deutlich oberhalb der Diagonalen und weist somit auf die große Anzahl der korrekt interpretierten Fragebogen hin. Außerdem stellt der AUC-Wert von 95% ein hohes Qualitätsmaß für den ‚Fusion‘-Klassifikator dar und bestätigt die Ergebnisse der Kreuz-Validierung und des ersten prospektiven Tests mit 20 Fragebogen, von denen 18 korrekt klassifiziert werden konnten. Dieses gute Ergebnis kann möglicherweise in der Verwendung des ‚Fusion‘-Algorithmus begründet sein, also in der Kombination unterschiedlicher, aber sich ergänzender Klassifikationsverfahren in der vorliegenden Arbeit [13, 15, 20]. Einschränkend bleibt, dass in der vorliegenden Arbeit keine ‚echte‘ Kontrollgruppe vorlag und die prospektive Feldstudie, also die Testung in einer Spezialambulanz für die Diagnostik von Stoffwechselerkrankungen bzw. für Kinder ohne Diagnose erst Gegenstand einer zukünftigen Feldstudie sein wird.

Müller et al. vom ‚Zentrum für unerkannte und seltene Erkrankungen (ZSE)‘ am Universitätsklinikum Gießen und Marburg unterstützen prinzipiell den Ansatz, durch computerunterstützte Diagnosefindung den diagnostischen Hürden, die SEs aufgrund ihrer Heterogenität und Vielzahl mit sich bringen, entgegen zu wirken [19]. Das hier vorgestellte Werkzeug soll daher zukünftig auf einen Einsatz z.B. an Zentren für SEs oder in Kinderarztpraxen geprüft werden.

Einen Beitrag zur computergestützten Diagnosefindung leisten ‚Clinical Decision Support Systems‘ (CDSS). Müller et al. weisen darauf hin, dass neuere CDSS wie ‚Isabel Healthcare‘ und ‚FindZebra‘ hilfreiche Ergebnisse im Bereich der Diagnoseunterstützung von SEs hervorbringen [19]. Beide Programme stellen auf der Basis manuell eingegebener Symptomkonstellationen eine Liste in Frage kommender Differentialdiagnosen auf, die anschließend vom Arzt geprüft werden können [19, 22]. Im Unterschied zum vorliegenden System verfolgen die genannten CDSS einen rein symptomorientierten Ansatz und stellen weniger die Anamnese in den Mittelpunkt. Der anamnesezentrierte Ansatz des vorliegenden auf Fragebögen basierenden Systems beruht auf der Annahme, dass Patienten (oder deren

Angehörige) über implizites wie explizites Wissen verfügen, welches für die individuelle Diagnosestellung von Bedeutung sein kann [19].

Die Einschränkungen des entwickelten Projektes ergeben sich aus den begrenzten Trainingsdaten, die für das System gesammelt werden konnten, der fehlenden Evaluation im klinischen Alltag und der Reduktion auf MPS. Zwar erreichte das System im prospektiven Test eine Sensitivität von 90%, es kann jedoch noch keine Aussage darüber getroffen werden, ob das System tatsächlich zwischen unterschiedlichen seltenen Stoffwechselerkrankungen, insb. den LSDs, und SEs anderer Genese unterscheiden kann bzw. ob im klinischen Alltag vergleichbar gute Ergebnisse erzielt werden. Zudem handelt es sich derzeit lediglich um Ergebnisse einer Pilotstudie, die noch nicht für den Praxiseinsatz geeignet ist. Die starke Wertigkeit der Frage nach Organomegalie und Entwicklungsauffälligkeiten illustriert, dass das vorgestellte System in der aktuellen Trainingsversion stark von den Antworten der MPS-Familien geprägt ist und die Erkennung anderer Symptome mangels Trainingsdaten derzeit noch defizitär sein muss. Es ist daher notwendig, das vorgestellte fragebögenbasierte System zu erweitern und dann mit neuen Trainingsdatensätzen zu testen.

Zusammengefasst wurde gezeigt, dass die strukturierte, qualitative Analyse von Interviews wichtige Erkenntnisse zum prädiagnostischen Erleben bei Menschen mit LSDs ergibt. Diese Erfahrungs-Muster von betroffenen Familien dienten in diesem Projekt als Trainingsmenge für ein KI-gestütztes Verfahren. In dem Pilotprojekt zeigte sich eine diagnostische Sicherheit für MPS von 90%. Das illustriert zwar den Wert der Anamnese für die Diagnose ausgewählter LSDs und gibt Raum für Hoffnung, dass zukünftig diagnoseunterstützende Verfahren die Diagnostik von SEs unterstützen können. Es erscheint jedoch verfrüht, anhand der vorliegenden Ergebnisse zu erwarten, dass vergleichbare Systeme rasch in der Praxis einsetzbar sind.

Literaturverzeichnis

1. Baris H. N., Cohen I. J., Mistry P. K.: Gaucher Disease: The Metabolic Defect, Pathophysiology, Phenotypes And Natural History. In: *Pediatr Endocrinol Rev.* 2014; 12(01):72–81
2. Bortz J., Döring N.: Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. 3. Auflage. Springer Verlag; Berlin, Heidelberg; 2002; S.297
3. Bouwman M. G., de Ru M. H., Linthorst G. E., Hollak C. E., Wijburg F. A., van Zweiten M. C.: Fabry patients' experiences with the timing of diagnosis relevant for the discussion on newborn screening. In: *Mol. Genet. Metab.* 2013; 109(2):201–207
4. Cimaz R., Guillaume S., Hilz M. J., Horneff G., Manger B., Thorne J. C., et al.: Awareness of Fabry disease among rheumatologists – current status and perspectives. In: *Clin Rheumatol* 2011; 30:467–475
5. Cimaz R., Vijay S., Haase C., Coppa G.V., Bruni S., Wraith E., et al.: Attenuated type I mucopolysaccharidosis in the differential diagnosis of juvenile idiopathic arthritis: a series of 13 Patients with Scheie syndrome. In: *Clin Exp Rheumatol* 2006; 24(2):196–202
6. Das A., Lagler M., Beck M. et al.: Lysosomal Storage Disease: Challenges in Multiprofessional Patient Care with Enzyme Replacement Therapy. In: *Klin. Pädiatr.* 2017; 229(3):168-174
7. Deegan P. B., Baehner A. F., Barba Romero M.-Á., Hughes D. A., Kampmann C., Beck M.: Natural history of Fabry disease in females in the Fabry Outcome Survey. In: *J Med Genet* 2006; 43:347–352
8. Ellaway C.: Diagnostic dilemma and delay in Fabry disease: Insights from a case series of young female patients. In: *Journal of Paediatrics and Child Health* 2015; 51:369–372
9. Edward K. L., Welch T.: The extension of Colaizzi's method of phenomenological enquiry. In: *Contemp Nurse* 2011; 39(2):163–171
10. Gabrielli O., Clarke L. A., Ficcadenti A., Santoro L., Zampini L., Volpi N., et al.: 12-year follow up of enzyme-replacement therapy in two siblings with attenuated mucopolysaccharidosis I: the important role of early treatment. In: *BMC Medical Genetics* 2016; 17:19
11. Garrino L., Picco E., Finiguerra I., Rossi D., Simone P., Roccatello D.: Living With and Treating Rare Diseases: Experiences of Patients and Professional Health Care Providers. In: *Qual Health Res.* 2015; 25(5):636–651
12. Harmatz P., Guigliani R., Schwartz I., Guffon N., Teles E. L., Miranda M. C., et al.: Enzyme replacement therapy for mucopolysaccharidosis VI: A phase 3, randomized, double-blind, placebo-controlled, multinational study of

recombinant human N-acetylgalactosamine 4-sulfatase (recombinant human arylsulfatase B or rhASB) and follow-on, open-label extension study. In: *J. Pediatr.* 2006; 148(4):533–539

13. Horowitz N., Moshkowitz M., Halpern Z., Leshno M.: Applying data mining techniques in the development of a diagnostics questionnaire for GERD. In: *Dig. Dis. Sci.* 2007; 52(8):1871–1878
14. Lehman T. J. A., Miller N., Norquist B., Underhill L., Keutzer J.: Diagnosis of the mucopolysaccharidoses. In: *Rheumatology* 2011; 50:v41–v48
15. Maghooli K., Langarizaden M., Shahmoradi L., Habibi-Koolaei M., Jebraeily M., Bouraghi H.: Differential Diagnosis of Erythmato-Squamous Diseases Using Classification and Regression Tree. In: *Acta Inform. Med.* 2016; 24(5):338–342
16. Mehta A., Ricci R., Widmer U., Dehout F., Garcia de Lorenzo A., Kampmann C., et al.: Fabry Disease defined: baseline clinical manifestations of 366 patients in the Fabry Outcome Survey. In: *Eur J Clin Invest* 2004; 34(3):236–242
17. Mistry P. K., Sadan S., Yang R., Yee J., Yang M.: Consequences of diagnostic delays in type 1 Gaucher disease: The need for greater awareness among Hematologists-Oncologists and an opportunity for early diagnosis and intervention. In: *Am. J. Hematol.* 2007; 82:697–701
18. Mücke U., Klemann C., Baumann U., Meyer-Bahlburg A., Kortum X., Klawonn F., et al.: Patient's Experience in Pediatric Primary Immunodeficiency Disorders: Computerized Classification of Questionnaires. In: *Front Immunol.* 2017; doi: 10.3389/fimmu.2017.00384
19. Müller T., Jerrentrup A., Schäfer J. R.: Computerunterstützte Diagnosefindung bei seltenen Erkrankungen. In: *Internist* 2017; doi:10.1007/s00108-017-0218-z
20. Naydenova E., Tsanas A., Howie S., Casals-Pascual C., De Vos M.: The power of data mining in diagnosis of childhood pneumonia. In: *J. R. Soc. Interface* 2016; doi: 10.1098/rsif.2016.0266
21. Obermeyer Z., Emanuel E. J.: Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. In: *N Engl J Med.* 2016; 375(13):1216–1219
22. Ramnarayan P., Tomlinson A., Rao A., Coren M., Winrow A., Britto J.: ISABEL: a web-based differential diagnostic aid for paediatrics: results from an initial performance evaluation. In: *Arch. Dis. Child.* 2003; 88(5):408–413
23. Rother A-K., Schwert 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. In: *PLOS ONE* 2015; 10(8):e0135180

24. Shapiro E. G., Nestrasil I., Delaney K. A., Rudser K., Kovac V., Nair N., et al.: A prospective natural history study of Mucopolysaccharidosis Type IIIA. In: *J Pediatr*. 2016; 170:278–287
25. Soni-Jaiswal A., Mercer J., Jones S. A., Bruce I. A., Callery P.: Mucopolysaccharidosis I; Parental beliefs about the impact of disease on the quality of life of their children. In: *Orphanet J Rare Dis*. 2016; 11(1):96
26. Strauss, A. L., Corbin, J. M.: *Grounded theory: Grundlagen qualitativer Sozialforschung*. Beltz Psychologie Verlags Union; 1996; S.148ff.
27. Wraith J. E., Beck M., Lane R., van der Loeg A., Shapiro E., Xue Y., et al.: Enzyme Replacement Therapy in Patients Who Have Mucopolysaccharidosis I and Are Younger Than 5 Years: Results of a Multinational Study of Recombinant Human α -L-Iduronidase (Laronidase). In: *Pediatrics* 2007; 120(1):37–46
28. Zhang X., Klawonn F., Grigull L., Lechner W.: *VoQs : A Web Application for Visualization of Questionnaire Surveys*.
29. Olsen TG, Jackson BH, Feeser TA, Kent MN, Moad JC, Krishnamurthy S, Lunsford DD, Soans RE. Diagnostic Performance of Deep Learning Algorithms Applied to Three Common Diagnoses in Dermatopathology. *J Pathol Inform*. 2018;9:32. doi: 10.4103/jpi.jpi_31_18. eCollection 2018
30. Zhang X, Wang S, Liu J, Tao C. Towards improving diagnosis of skin diseases by combining deep neural network and human knowledge. *BMC Med Inform Decis Mak*. 2018 Jul 23;18(Suppl 2):59. doi: 10.1186/s12911-018-0631-9
31. Jung H, Kim B, Lee I, Yoo M, Lee J, Ham S, Woo O, Kang J. Detection of masses in mammograms using a one-stage object detector based on a deep convolutional neural network. *PLoS One*. 2018 Sep 18;13(9):e0203355. doi: 10.1371/journal.pone.0203355. eCollection 2018