Journal of Inherited Metabolic Disease

JIMD JOURNAL OF INHERITED METABOLIC DISEASE SSIEM

ORIGINAL ARTICLE OPEN ACCESS

Expert-Designed Fact Sheets and AI-Based Analysis of Patient Symptoms to Combat Diagnostic Delays in Inherited Metabolic Diseases

Aline Cano[1,2] | Xiaoyi Chen[3,4] | Azza Khemiri[1] | Anais Brassier[1,5] | Arnoux Jean-Baptiste[1,5] | Roseline Froissart[1,6] | Juliette Bouchereau[1,5] | Célia Hoebeke[1,2] | Karin Mazodier[1,2] | Bénédicte Héron[1,7] | Philippe Labrune[1,8] | Catherine Caillaud[1,9] | David Cheillan[1,6] | Yann Nadjar[1,10] | Samia Pichard[1,5] | Apolline Imbard[1,9] | Magali Pettazzoni[1,6] | Claire Douillard[1,11] | Belmatoug Nadia[1,12] | Anna-Line Calatayud[1] | Mounira Zerguini[1] | Nicolas Garcelon[3] | Jean-François Benoist[1,9] | Cécile Acquaviva[1,6] | Pascale De Lonlay[1,5] | the other members of the expert group consortium

[1]Filière nationale de santé maladies rares G2m- Maladies Héréditaires du Métabolisme: G2m French Rare Diseases Healthcare Network for Inherited Metabolic Diseases, Paris, France | [2]Reference Center for Inherited Metabolic Diseases, Marseille University Hospital, Assistance Publique—Hôpitaux de Marseille (AP-HM), Marseille, France | [3]Data Science Platform, Imagine Institute, Paris-Cité University, Paris, France | [4]Division of Computational Health Sciences, Department of Surgery, University of Minnesota, Minneapolis, Minnesota, USA | [5]Reference Center for Inherited Metabolic Diseases, Necker—Enfants Malades University Hospital, Assistance Publique—Hôpitaux de Paris (AP-HP), Paris-Cité University, Paris, France | [6]Metabolic Inborn Errors of Metabolism Unit, East Hospital Group, Biochemical and Molecular Biology Laboratory, Lyon University Hospital, Bron, France | [7]Department of Paediatric Neurology, Reference Center for Lysosomal Diseases, Armand Trousseau-La Roche Guyon University Hospital, and I2D2 Hospitalo-University Federation, Sorbonne University, Assistance Publique—Hôpitaux de Paris (AP-HP), Paris, France | [8]Reference Center for Inherited Metabolic Liver Diseases, Antoine Béclère Hospital, Assistance Publique—Hôpitaux de Paris (AP-HP), ClamartParis-Saclay University, France | [9]Metabolic Biochemistry Laboratory, Necker—Enfants Malades University Hospital, Assistance Publique—Hôpitaux de Paris (AP-HP), Paris-Cité University, Paris, France | [10]Neuro-Metabolism Unit, Neurology Department, Reference Center for Metabolic and Lysosomal Neurological Diseases, Pitié-Salpêtrière University Hospital, Assistance Publique—Hôpitaux de Paris (AP-HP), Paris, France | [11]Department of Endocrinology, Diabetology, Metabolism, Reference Center for Inherited Metabolic Diseases, Lille University Hospital, Lille, France | [12]Department of Internal Medicine and Reference Center for Lysosomal Disorders, University Hospital, Beaujon Hospital, Assistance Publique—Hôpitaux de Paris (AP-HP), Clichy, France

Correspondence: Aline Cano (alinecano@yahoo.fr)

Received: 22 September 2025 | **Revised:** 2 February 2026 | **Accepted:** 5 February 2026

Academic Editor: Sven F. Garbade

Keywords: delayed diagnosis | diagnostic errancy | diagnostic odyssey | fact sheets | inborn errors of metabolism (IEM) | inherited metabolic disorders (IMDs) | undiagnosed conditions

ABSTRACT

The importance of early diagnosis of inherited metabolic diseases (IMDs) is well known, as it allows early intervention to prevent or reduce complications and improve prognosis, since many of these disorders are treatable. However, diagnosis can still be delayed, and many patients remain undiagnosed. Reducing diagnosis delays is a primary goal of the French Ministry of Health and Prevention (Rare Disease Department). This article describes a national initiative coordinated by the French network for IMD, "Filière G2m." Sixty-seven IMD experts from various reference and competence centers in France drafted one-page summaries dedicated to specific diseases or groups of diseases in the field of IMDs, covering the full spectrum of IMDs. These documents include keywords summarizing clinical signs which, when considered alongside data from routine biological or imaging tests, should suggest the diagnosis of an IMD. A total of 48 summaries have been drafted and are available on the Filière G2m website.

Jean-François Benoist and Cécile Acquaviva contributed equally to this study.

To assess the accuracy and relevance of the diagnostic fact sheets, we selected 4 IMDs and compared their content with the clinical profiles of patients followed at Necker—Enfants Malades Hospital, using Natural Language Processing tools to automatically extract patient phenotypes from medical records (Dr Warehouse). We found a strong alignment between the fact sheets and the real-world clinical data from these patients. This tool will enable patients to recognize themselves in an IMD. General practitioners will use these documents alongside diagnostic aid software. It may also support new artificial intelligence-based technologies to identify undiagnosed patients in hospital databases.

1 | Introduction

Rare diseases affect approximately 25 million people in Europe and nearly 3 million in France, accounting for about 4.5% of the population [1]. Inherited Metabolic Disorders (IMDs) are individually rare but collectively encompass over 1000 genetic disorders, with a cumulative incidence of approximately 1 in 800 births [2]. In France, one in two individuals affected by a rare disease lacks a precise diagnosis, with an average diagnostic delay of 5 years [1]. Diagnostic errancy is defined as the period between the onset of initial symptoms and the establishment of a definitive diagnosis. This issue is critical in rare diseases such as IMDs, presenting a public health challenge with serious consequences for patients, particularly for treatable IMDs.

To address this, France has implemented a national policy aimed at reducing diagnostic errancy in rare diseases, as outlined in the successive National Plans for Rare Diseases (PNMRs). The objective of health authorities is to obtain a diagnosis within 1 year after the visit, at the latest 1 year afterwards; otherwise, it is referred to as a diagnosis impasse. The various successive PNMRs include measures such as further structuring of diagnostic pathways through reference and competence centers for rare diseases (CRMR and CCMR respectively), coordinated by rare disease networks like the French network for IMDs, "Filière G2m." These networks of specialized healthcare providers facilitate coordinated patient care and provide training for healthcare professionals in recognizing and managing IMDs, particularly for general practitioners. The latest PNMR has also facilitated the establishment of patient support mechanisms, dedicated information platforms, and enhanced access to advanced diagnostic technologies such as high-throughput genomic analysis platforms to overcome diagnostic impasses [3] (https://pfmg2025.aviesan.fr).

Our work represents a new tool to combat the diagnostic odyssey, assisting in the diagnosis of IMDs, and which can be used by patients, healthcare professionals, and hospitals when paired with artificial intelligence. This will lead to faster referral of patients to expert centers and earlier access to appropriate treatments.

2 | Methods

The objective of this project was to create concise and easy-to-understand fact sheets that summarize the essential clinical symptoms: when considered together and alongside standard biological or imaging data, fact sheets can assist clinicians and patients in navigating the diagnostic process toward specific diagnoses of IMDs and encourage requests for expert consultation.

The diseases were selected based on two criteria: those covered by emergency management protocols and/or those with specific treatments or management [4]. Additionally, we created symptom-based fact sheets, but instead of integrating them into these diagnostic fact sheets—since these symptoms are not yet associated with a specific disease—we incorporated them into emergency protocols by symptoms that feature decision trees for diagnosis (https://filiere-g2m.fr/urgences) [4].

An initial version of each "disease" document was created by a physician affiliated with the French G2m network. This version was supported by resources such as national diagnosis and treatment guidelines [5], and bibliographic searches. Subsequently, this draft was reviewed by a group of experts specializing in the specific disease or symptom. A total of 67 IMD experts including 43 clinicians (paediatricians or specialists for adult patients) and 24 biologists were involved in revising the documents based on their expertise. Sixteen specialists from other rare disease networks, such as Filnemus (neuromuscular diseases network), Cardiogen (hereditary or rare heart disease network), Filfoie (rare liver diseases network) and Firendo (rare endocrine diseases network) also contributed. Feedback from these experts led to multiple iterations before final approval by the expert group and patient associations. The documents were refined by a dedicated graphic design team prior to distribution. Any significant changes made were implemented to ensure consistency across all documents wherever feasible.

To evaluate the frequency of symptoms described in these fact sheets in real life based on a patient cohort, we selected four IMDs (Table 1) and conducted a detailed comparison between the content of the diagnostic fact sheets and the clinical profiles of patients diagnosed with these IMDs at Necker—Enfants Malades Hospital. For this analysis, we leveraged the Natural Language Processing (NLP) tools within the clinical data warehouse of Necker (Dr Warehouse) to automatically extract phenotypes from the clinical notes of patients followed at Necker Hospital. Dr. Warehouse refers to the clinical data warehouse infrastructure deployed at Necker Hospital, as well as the associated in-house software suite developed at Necker to ingest, structure and query electronic health record (EHR) data. The platform aggregates EHR-derived clinical narratives (e.g., medical reports and discharge summaries) and structured data for approximately 1 million patients treated at Necker Hospital [6]. The phenotype extraction process uses a hybrid strategy that combines a dictionary-based method with deep-learning approaches, and includes the ability to detect the context of the extracted phenotype, such as negation and the person experiencing it. We then determined whether these phenotypes were explicitly included in the corresponding diagnostic fact sheets. To facilitate interpretation, we grouped similar or related phenotypes

TABLE 1 | Details of diseases and patient distribution for which a comparative analysis assisted by AI was performed between our fact sheets and medical file data.

IMD	ORPHA code	OMIM code	Number of patients
VLCAD deficiency	ORPHA:26793	OMIM:201475	16
Lysinuric protein intolerance	ORPHA:470	OMIM:222700	23
Gaucher disease	ORPHA:355		19
Type 1	ORPHA:77259	OMIM:230800	8
Type 2	ORPHA:77260	OMIM:230900	1
Type 3	ORPHA:77261	OMIM:231000	10
Wolman disease	ORPHA:75233	OMIM:278000	7

Note: The table lists the OMIM and ORPHA codes for the four selected diseases: Long chain acyl-CoA dehydrogenase deficiency, lysinuric protein intolerance, Gaucher disease, and Wolman disease, along with the number of patients identified with each disease at Necker—Enfants Malades Hospital. The information contained in the fact sheets has been cross-referenced with phenotypic data extracted from the clinical data warehouse for these patients.

into broader clinical/biological phenotypes (e.g., grouping "anemia", "chronic anemia," and "symptomatic anemia" under a unified "anemia" category). In addition, each phenotype was also mapped to the organ or system category (e.g., nervous system, liver, digestive system), based on the physiological domain affected. We also evaluated whether each phenotype extracted from patients' medical records was associated with the disease in established knowledge bases, such as human phenotype ontology (HPO), OMIM, and Orphanet. In doing so, we considered not only the phenotypes directly linked to the disease but also their ancestral terms (more general phenotypes) within the HPO hierarchy. Since clinical terms in patients' medical records may vary in expressions, synonyms, or levels of granularity, an expert (A.C.) reviewed the output tables to ensure accuracy in these comparisons. The expert verified that when a phenotype was marked as absent from the diagnostic fact sheet, or present in the fact sheet but not in the knowledge bases, it was not simply expressed using a synonymous term or a variation in granularity. This analysis allowed us to assess how well the fact sheets capture clinically relevant information that may not be fully represented in existing knowledge bases, providing insight into their utility for guiding diagnosis and management in real-world clinical settings. A certain number of data points have been manually verified in the patients' files.

3 | Results

Forty-eight fact sheets have been created in French and were translated into English (Table 2). Forty-one fact sheets cover various categories of diseases based on the classification by Saudubray et al. [7], including 14 documents on disorders of protein or sugar intoxication leading to the acute or chronic accumulation of toxic compounds, 9 documents on energy deficiency disorders resulting from a deficiency in energy (ATP) production or utilization, and 18 on disorders of complex molecules caused by disturbances in the synthesis or catabolism of complex molecules. Seven fact sheets involve symptoms often indicative of a MHM, with a decision tree.

With regard to disease sheets, the initial document created focused on MPI-CDG (Figure 1). Apart from adjustments related to the specific pathologies addressed, the format of the disease fact sheets was standardised and they were designed as single-sided documents in A4 landscape format to be read from top to bottom. Each fact sheet begins with the title of the condition, followed by key general points aimed at facilitating comprehension of the document (e.g., phenotypic variability, various disease forms, etc.). Clinical manifestations are then presented in columns, typically categorized by the affected organ. Following the clinical aspects, details of abnormal findings from biological tests or radiological examinations used to investigate clinical anomalies are provided. Subsequently, a question format restates the title, emphasizing that a combination of these elements should prompt consideration of a diagnosis and referral to an expert center (with contact information provided). The following section presents, in general terms, the specialized diagnostic approach typically undertaken by expert centers to confirm the diagnostic hypothesis. While some aspects of specific management are included, they are intentionally succinct, as comprehensive details extend beyond diagnostic assistance and can be found in emergency protocols [4], PNDS [5], other documents, or specialized resources. Hyperlinks and contact information for references are indicated on the fact sheets. Regarding the symptom fact sheets, they include a diagnostic tree, as well as the main tests to be performed, and treatments to be implemented in a patient presenting with this symptom who has not yet been diagnosed. These documents are accessible on the G2M network website (https://www.filière-g2m.fr), providing free access for all visitors, including patients and non-specialist healthcare professionals.

Furthermore, the information contained in these resources (disease fact sheets) has been cross-referenced with phenotypic data extracted from clinical data warehouse [6] for patients followed at Necker—Enfants Malades Hospital with one of the four IMDs: very long chain acyl-CoA dehydrogenase (VLCAD) deficiency, lysinuric protein intolerance (LPI), Gaucher DISEASE, and Wolman disease. These four diseases were selected because they cover the three groups of IMDs, and because we believe that a delay in diagnosis may be particularly observed. The number of patients identified with each disease at Necker—Enfants Malades Hospital is provided in Table 1.

Patients' symptoms were categorized by affected organs, and their frequencies were recorded in fact sheets created for the

TABLE 2 | List of diseases ($n = 41$) and symptoms ($n = 7$) for which these summary sheets have been produced.

Category	Disorders/symptoms	OMIM code
Disorders of intoxication ($N = 14$)	Urea cycle disorders	207900 (ASA), 237300 (CPS1), 311250 (OTC), 207800 (Arg), 215700 (Cit), 238970 (HHH), 237310 (NAGS)
	Organic acidurias = methylmalonic aciduria (MMA) and propionic acidemia (PA)	251000 (mutase), 609058 (mutase), 251100 (CblA), 251110 (CblB), 606054 (PA)
	Isovaleric acidemia (IVA)	243500
	Maple syrup urine disease (MSUD)	248600 and 615135
	Lysinuric protein intolerance (LPI)	222700
	Glutaric aciduria type 1 (GA1)	231670
	Homocystinuria due to CBS deficiency	236200
	Methylmalonic aciduria and homocystinuria, cblc type (MAHCC)	277400
	Tyrosinemia type I	276700
	Galactosemia	230200, 230350, 230400
	Hereditary fructose intolerance (HFI)	229600
	Glutathione synthetase deficiency (GSSD)	266130
	Alkaptonuria	203500
	Wilson disease	277900
Energy deficiency disorders ($N = 9$)	Glycogen storage disease type 1a (GSD1A) et type 1b (GSD1b)	232200, 232220
	Glycogen storage disease type 3 (GSD3)	232400
	Fructose 1,6-bisphosphatase deficiency (FBP1D)	229700
	Mitochondrial fatty acid ß-oxidation disorders: VLCAD, LCHAD, MCAD, CPT2, CACT or MTP deficiency, glutaric aciduria II (GA2), carnitine transporter deficiency	201475, 609016, 201450, 255110, 600649, 608836, 212138, 609015, 231680, 212140
	Ketogenesis disorders (HMG CoA lyase or HMG CoA synthase deficiencies)	246450, 605911
	Ketoacidosis from ketolysis disorders: SCOT, MAT, MCT1	245050, 203750, 616095
	Biotinidase and holocarboxylase synthetase deficiencies	253260, 253270
	MELAS and MELAS-related mitochondrial diseases	—
	Pyruvate dehydrogenase deficiency	312170 (PDHA), 608782 (PDHP), 608769 (PDHX), 300502 (PDHA1), 179060 (PDHB), 614111 (PDHBD)
Disorders of complex molecules ($N = 18$)	PMM2-CDG	212065
	MPI-CDG	602579
	Mucopolysaccharidosis (MPS)	607016 (MPS IS), 252900 (MPS IIIA), 252920 (MPS IIIB), 253000 (MPS IVA), 253010 (MPS IVB), 253200 (MPS VI), 309900 (MPS II), 607014 (MPS IH), 607015 (MPS I H/S)

(Continues)

TABLE 2 | (Continued)

Category	Disorders/symptoms	OMIM code
	Pompe disease	232300
	TANGO2 deficiency	616878
	Niemann Pick type C disease	257220
	GM2 gangliosidosis	272800
	Gaucher disease	230800 (type I), 230900 (type II), 231000 (type III)
	Acid sphingomyelinase deficiency	607616 (type B), 257200 (type A)
	Neuronal ceroïd lipofuscinosis	256730 (CLN1), 204500 (CLN2), 204200 (CLN3), 256731 (CLN5), 606725 (CLN6), 610951 (CLN7), 204300 (CLN8), 609055 (CLN9), 610127 (CLN10)
	Fabry disease	30500
	Wolman and CESD	620151 (Wolman disease), 278000 (CESD)
	Metachromatic leukodystrophy	607574
	Alpha-mannosidosis	609458
	Erythropoietic protoporphyria	177000
	Acute hepatic porphyria crisis (AIP, VP, HCP)	176000, 176200, 121300
	Rhabdomyolysis due to LPIN1 mutations	605518
	X-linked adrenoleukodystrophy	300100
Symptoms ($N=7$)	Hypoglycemia	—
	Coma	—
	Rhabdomyolysis	—
	Acute cardiac presentations of IMD	—
	Neonatal liver failure	—
	Metabolic acidosis	—
	Hyperammonaemia	—

Abbreviations: Arg, argininemia; ASA, argininosuccinic aciduria; CA5A, carbonic anhydrase 5A; CACT, carnitine-acylcarnitine translocase; Cbl, cobalamin; CBS, cystathionine beta synthase; CDG, congenital disorder of glycosylation; CESD, cholesterol ester storage disease; Cit, citrullinemia; CPS1, carbamoylphosphate synthetase I; HHH, hyperornithinemia–hyperammonaemia–homocitrullinemia; HMG-CoA, 3-hydroxy-3-methylglutaryl-CoA; LCHAD, long-chain 3-hydroxyacyl-coa dehydrogenase; MAT, mitochondrial acetoacetyl-CoA thiolase; MCAD, medium-chain acyl-CoA dehydrogenase; MCT1, monocarboxylate transporter 1; MPS, mucopolysaccharidosis; MTP, mitochondrial trifunctional protein; NAGS, *N*-acetylglutamate synthase; OMIM, Online Mendelian Inheritance in Man; OTC, ornithine transcarbamylase; PA, propionic acidemia; PDHA, pyruvate dehydrogenase E1-alpha; PDHA1, pyruvate dehydrogenase alpha-1; PDHB, pyruvate dehydrogenase beta polypeptide; PDHBD, pyruvate dehydrogenase E1-beta deficiency; PDHP, pyruvate dehydrogenase phosphatase; PDHX, pyruvate dehydrogenase complex component X; PMM2, phosphomannomutase 2; SCOT, succinyl-CoA:3-oxoacid CoA transferase; VLCAD, very long-chain acyl-coenzyme A dehydrogenase.

four diseases (Figure 2). It should be noted that symptom prevalence varies with age; for example, pulmonary and renal issues in LPI are more common later in life, and studies including older patients would likely report higher rates of these conditions.

Table 3 presents the most common symptoms extracted from patients' records for the four disorders. Our comparative analysis revealed that these frequent symptoms (or phenotypes) are generally listed in both existing knowledge bases and diagnostic fact sheets for the four diseases (Table 3). Phenotypes that were often missing from both sources (not associated with the disease in OMIM nor Orphanet) or from our fact sheets tended to be non-specific common symptoms, such as fever, vomiting, diarrhea, and cough. Although these symptoms are frequently noted in clinical records, they may not be directly relevant to the diagnosis of these IMDs. They probably represent intercurrent events, frequent in the general population, but possible triggers of metabolic imbalance in patients with underlying metabolic disease, and therefore particularly noted in the records. Overall, this indicates a strong alignment between the fact sheets and knowledge bases in capturing key diagnostic features from real-world clinical data. Our analysis also revealed that phenotypes extracted from real-world patient data may differ in granularity compared to how they are encoded in established knowledge bases like OMIM and Orphanet. To address these differences in granularity,

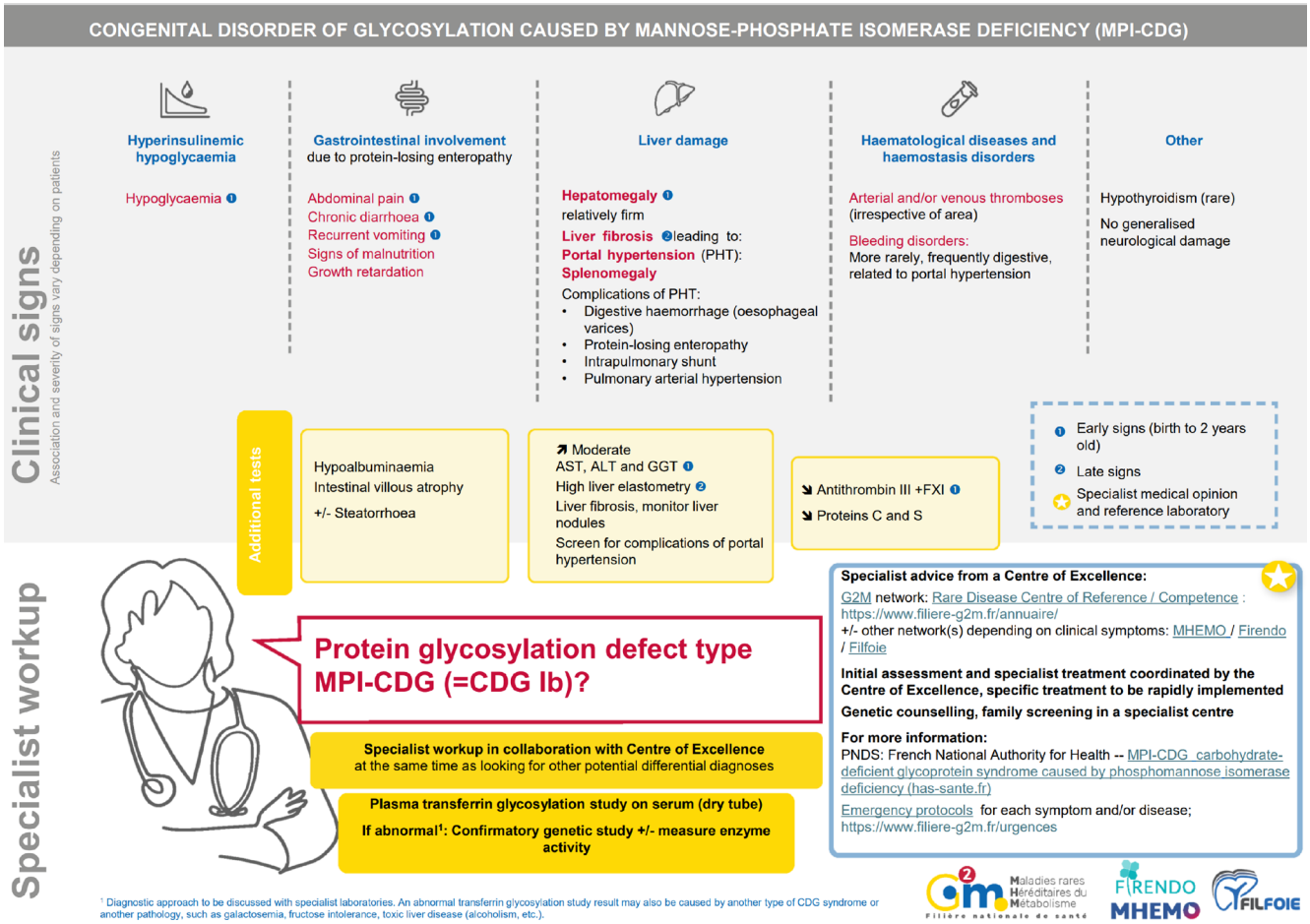


FIGURE 1 | Fact sheet for MPI-CDG. This initial template, developed by a multidisciplinary team of experts in inherited metabolic disorders from the G2m rare disease clinical network, served as the basis for creating the remaining documents.

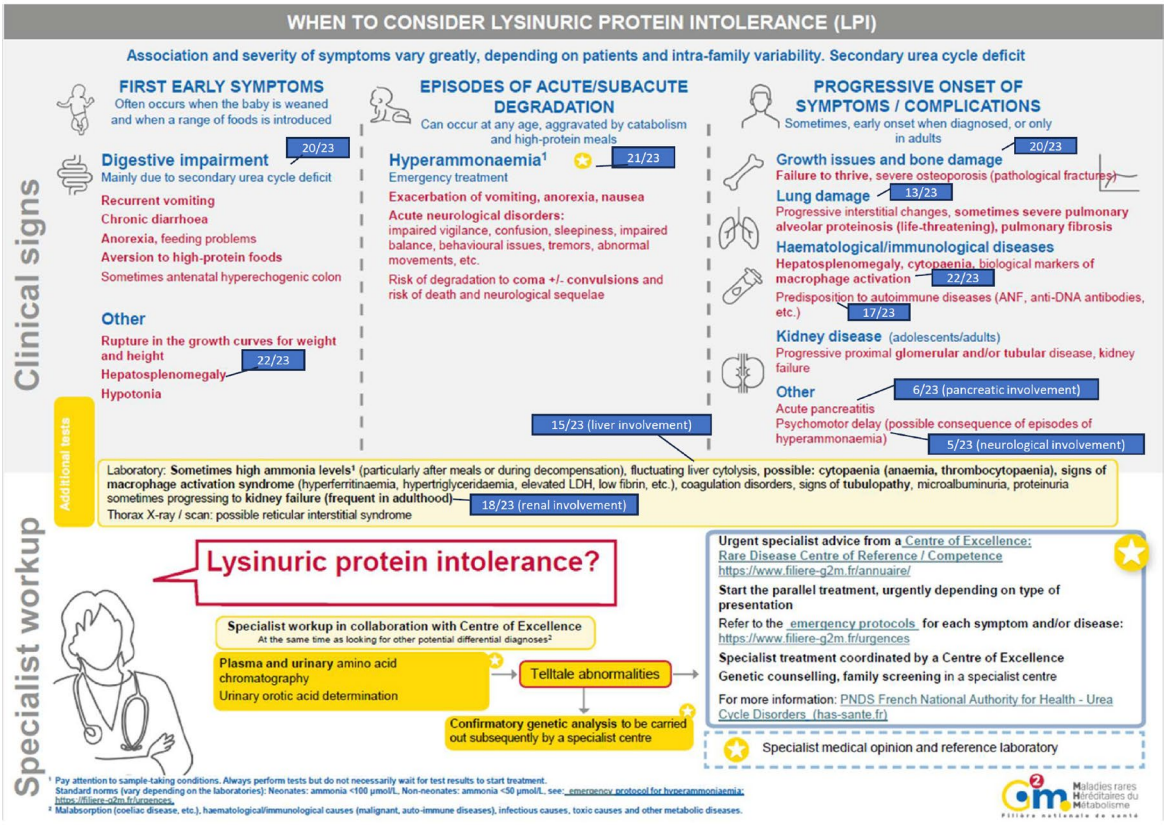
we utilized the HPO hierarchy to capture both specific phenotypes and their ancestral (more general) terms associated with each disease. Despite this, some sibling phenotypes, that is, phenotypes at the same hierarchical level but not directly linked, were present in patient's clinical notes but absent from the knowledge bases. For example, "anorexia," reported in 11/23 patients with LPI in Necker—Enfants Malades Hospital, is not associated with LPI in OMIM or Orphanet, which instead list its sibling phenotype "feeding difficulties," reported in only 2 of LPI patients, showing the more frequent use of "anorexia" in clinical notes. Similarly, "abdominal pain" noted in 5/7 patients with Wolman syndrome in Necker—Enfants Malades Hospital is absent from the knowledge bases. The sibling phenotype present in bases is "abdominal distension", an additional phenotype also observed in the patient records and among the patients. Finally, the term "asthenia" reported in 8/19 of Gaucher patients in Necker—Enfants Malades Hospital is missing from the knowledge bases. However, the HPO term "fatigue" is present in the databases for Gaucher disease, with the note that asthenia should be distinguished from fatigue, as asthenia refers to a sensation of exhaustion that occurs before any effort. Notably, the diagnostic fact sheets include these specific terms commonly used in practice to describe certain symptoms present in patients. This highlights the possible limitations of current ontologies in representing real-world clinical language for patient description and further emphasizes the practical utility of the diagnostic fact sheets in providing a comprehensive and clinically relevant view for diagnosis.

Furthermore, this study highlighted a particular cardiac involvement in LPI. Indeed, by grouping the symptoms into categories of involvement, it appears that 12 of the 23 patients with LPI followed at the Necker—Enfants Malades Hospital have at least one symptom of cardiac involvement including valvulopathy, arrhythmia, cardiomegaly, ventricular hypertrophy, and ischemic cardiomyopathy.

4 | Discussion

Here, we report on an initiative by the Filière G2m to combat diagnostic odyssey: the creation of 48 fact sheets: 41 dedicated to specific diseases or groups of diseases in the field of inherited metabolic disorders, and 7 focused on symptoms often indicative of an IMD. These fact sheets facilitate diagnosis by including keywords summarizing clinical signs, as well as routine biological and/or imaging tests. To assess the accuracy and relevance of the diagnostic fact sheets, we selected 4 IMDs and compared their content with the clinical profiles of patients followed at Necker—Enfants Malades Hospital, using NLP tools to automatically extract patient phenotypes from medical records.

a:



b:

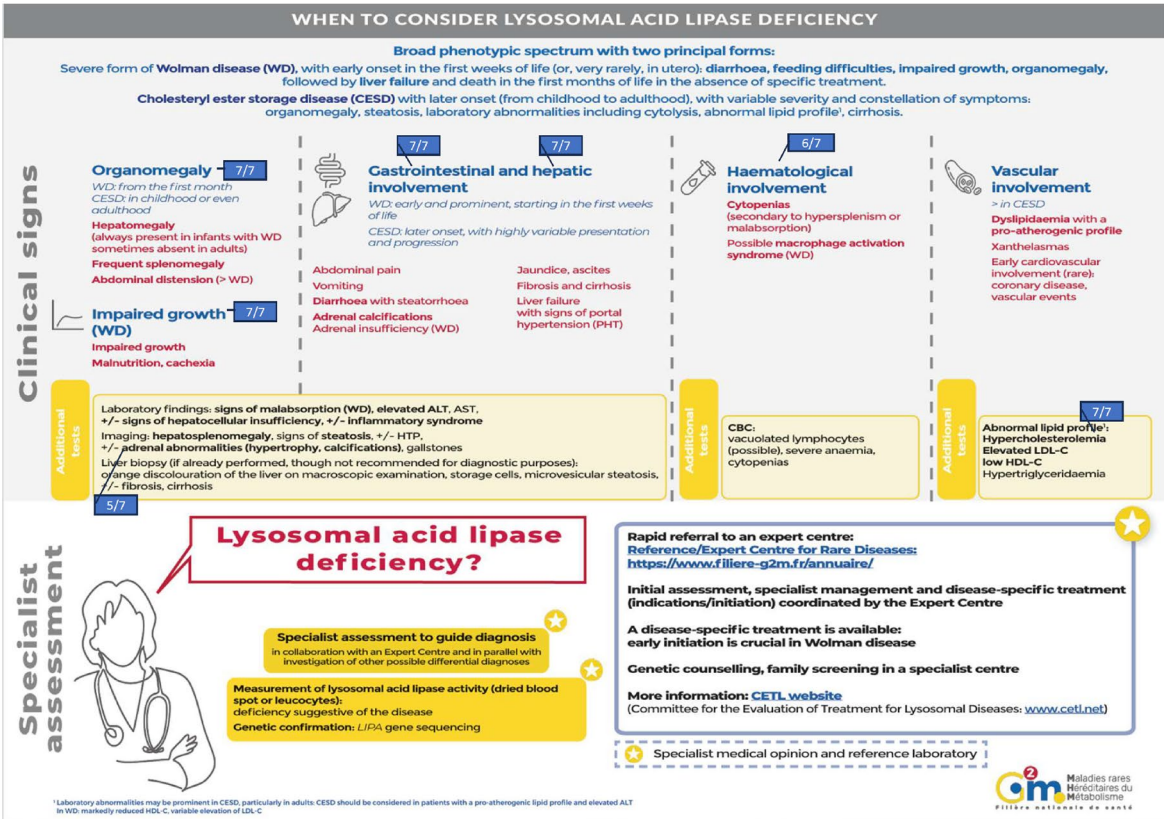
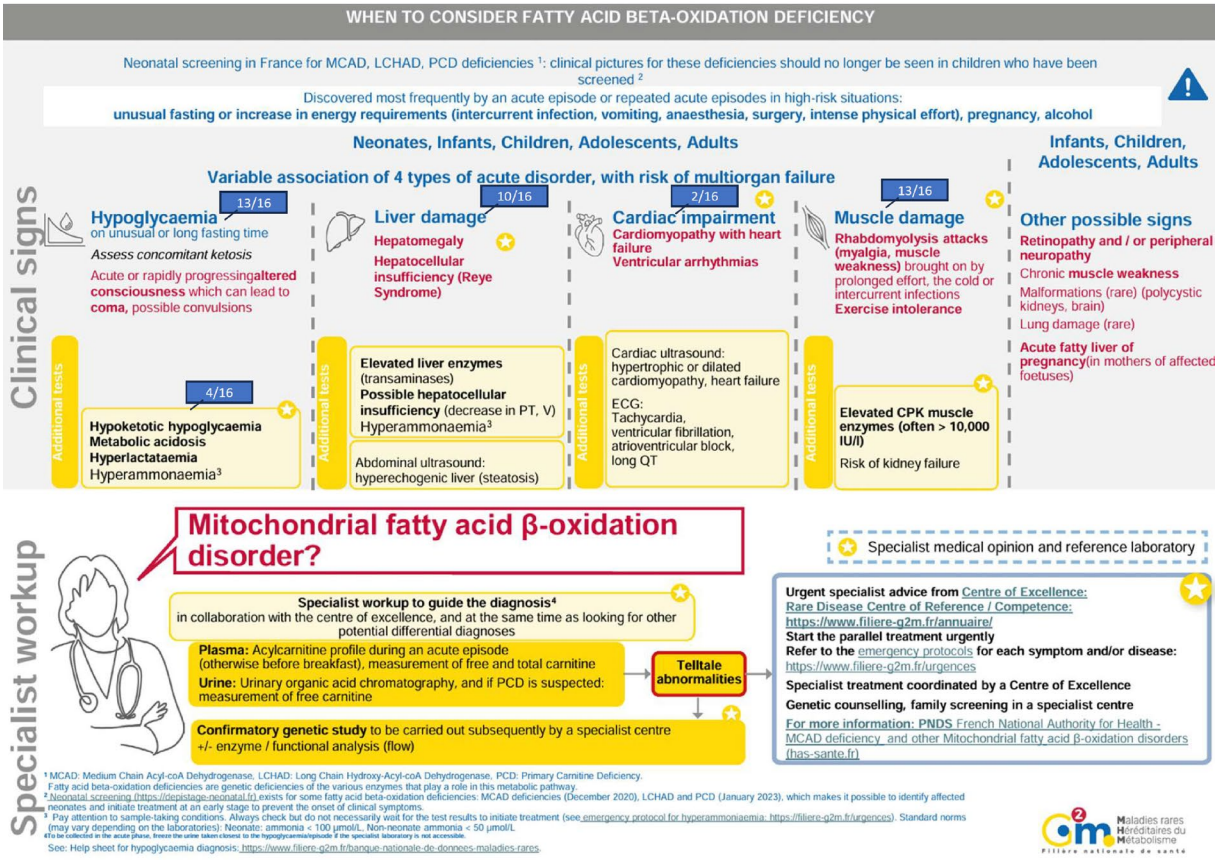


FIGURE 2 | Example of four diagnostic fact sheets showing the number of patients at Necker—Enfants Malades Hospital with symptoms categorized by organ/system involvement for each disease. (a) Lysinuric protein intolerance. (b) Wolman disease. (c) VLCAD deficiency (beta-oxidation disorders fact sheet). (d) Gaucher disease.

c:



d:

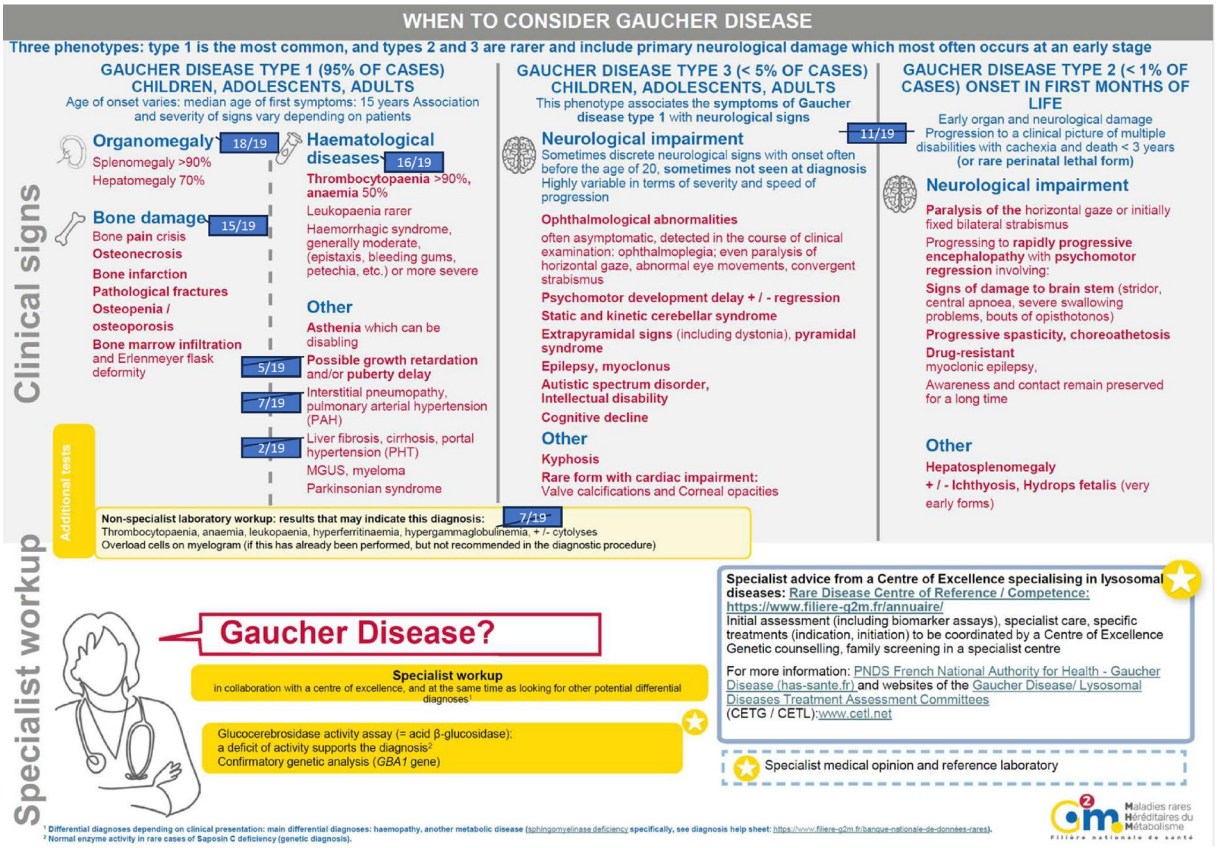


FIGURE 2 | (Continued)

TABLE 3 | Phenotypes extracted from patients' medical records of patients followed at Necker—Enfants Malades Hospital with lysinuric protein intolerance, Wolman disease, VLCAD deficiency, and Gaucher disease.

Symptoms/biological or imaging abnormalities	Number of patients (*n*)	Number of patients < 18 years	Included in fact sheet	Included in knowledge bases
Disease: LPI (n = 23)				
Hepatomegaly	22	10	✓	✓
Hyperammonaemia	21	9	✓	✓
Splenomegaly	21	10	✓	✓
Macrophage activation syndrome	20	8	✓	✓
Anemia	17	8	✓	✓
Growth delay	17	7	✓	✓
Vomiting	17	6	✓	✓
Hyperferritinemia	16	6	✓	✓
Hepatosplenomegaly	16	6	✓	✓
Thrombocytopenia	15	6	✓	✓
Cough	15	8	✗	✗
Renal insufficiency	14	2	✓	✓
Alveolar proteinosis	13	4	✓	✓
Asthenia	12	1	✗	✗
Hepatic cytolysis	12	6	✓	✓
Hypertriglyceridemia	12	3	✓	✓
Hematoma	12	4	✗	✗
Disease: Wolman (n = 7)				
Hepato/splenomegaly	6	6	✓	✓
Denutrition	5	5	✓	✓
Malabsorption	5	5	✓	✓
Abdominal pain	5	5	✓	✗
Diarrhea	5	5	✓	✓
Adenopathy	5	5	✓	✗
Cough	5	5	✗	✗
Steatosis	5	5	✓	✓
Vomiting	5	5	✓	✓
Adrenal calcifications	4	4	✓	✓
Vitamin D deficiency	4	4	✗	✗
Abdominal distention	4	4	✓	✓
Anemia	4	4	✓	✓
Abdominal bloating	4	4	✓	✓
Portal fibrosis	4	4	✓	✓
Hypotonia	4	4	✗	✗
Nausea	4	4	✗	✗
Constipation	4	4	✗	✗

(Continues)

TABLE 3 | (Continued)

Symptoms/biological or imaging abnormalities	Number of patients (*n*)	Number of patients < 18 years	Included in fact sheet	Included in knowledge bases
Disease: VLCAD (n = 16)				
Hypoglycaemia	12	8	✓	✓
Rhabdomyolysis	12	7	✓	✓
Pain	9	5	✓	✓
Vomiting	9	7	✗	✓
Fever	8	5	✗	✗
Infection	7	5	✗	✗
Liver cytolysis	6	5	✓	✓
Abdominal pain	6	4	✗	✗
Hypotonia	6	6	✓	✓
Hepatomegaly	6	5	✓	✓
Disease: Gaucher (n = 19)			✓	✓
Hepatomegaly	17	14	✓	✓
Splenomegaly	15	13	✓	✓
Anemia	14	12	✓	✓
Thrombocytopenia	12	9	✓	✓
Hepatosplenomegaly	10	7	✓	✓
Epilepsia	9	8	✓	✓
Cough	9	9	✗	✓
Impaired oculomotricity	8	8	✓	✓
Asthenia	8	6	✓	✗
Hypergammaglobulinemia	7	5	✓	✓

Note: This table presents the 10 most frequent phenotypes or those found in ≥ 50% of patients. ✓ = yes/ ✗ = no.
Abbreviation: LPI, lysinuric protein intolerance.

We found a strong alignment between the fact sheets and the real-world clinical data from these patients, although some patients did not have all of these symptoms; we indicate in Figure 2 the real-world frequency of these symptoms.

Diagnostic odyssey, or diagnostic errancy, occurs when a patient undergoes multiple medical consultations and diagnostic tests without receiving a conclusive diagnosis. This often results from the rarity or limited understanding of their condition, as well as the perception that their symptoms are insignificant or subjective. This situation may arise when a general practitioner struggles to connect various symptoms or fails to consider the possibility of a rare disease that warrants referral to a specialized center for further evaluation, diagnosis confirmation, and tailored treatment, including innovative therapies. By providing clear and concise information on a single page, including essential clinical symptoms and common radiological and/or biological indicators that should raise suspicion of various IMDs, these fact sheets aim to serve as a valuable resource for non-specialist physicians. Now that these fact sheets are freely available on the Filière G2m website, our plans are to present them at conferences for general practitioners and doctors who are not specialised in metabolic diseases, and link them to diagnostic support software. An English translation has been done, and these documents may also be useful in other countries. For instance, they could be shared by European Reference Networks (ERNs), particularly MetabERN [8]. Finally, although developed by a multidisciplinary group of highly specialized professionals in collaboration with colleagues from other specialties, no formal methodology (e.g., Delphi) was employed to validate them. Gathering feedback from users (physicians, patient associations) is planned 1 year after distribution, with updates as needed. Several educational resources are available to inform non-specialist physicians about metabolic diseases, including national guidelines [5], metabolic disease literature, books, scientific publications, videos, and webinars conducted by experts and patient associations. However, these resources typically focus on providing information post-diagnosis. The unique benefit of the "diagnostic fact sheets" introduced here is their rapid and efficient overview of a particular disease, starting from specific symptoms, along with their quick and easily digestible format,

making them accessible for both patients and healthcare professionals on the internet.

Diagnostic support software is also being developed for physicians, primarily in community settings but also for hospital doctors, and is freely accessible to patients. Users can input multiple clinical or biological symptoms, and the software aggregates the data to suggest potential diagnoses based on symptom correlations. One such tool is "IEMbase," a specialized search engine for IMDs. This software allows users to input key clinical or biological symptoms and identifies potential IMDs, providing links to expert information on numerous IMDs [9]. However, it does not cover possible differential diagnoses among other rare diseases, and in practice, this tool seems to cater more to specialized professionals who are already suspecting an IMD or seeking expert insights. The second tool is "RDK" by Tekkare [10], a reliable diagnostic aid that utilizes symptoms to help doctors in identifying pathologies, whether metabolic or not, while also pinpointing relevant expert centers and rare disease networks (https://www.aswek now.com/press). Orphanet, a leading resource for information on rare diseases and orphan drugs [11], provides details on rare diseases, including IMDs, and features a search function for diagnosing rare diseases based on symptoms. Since our fact sheets were formulated in 2023–2024 with guidance from national IMD experts, they will be used to enhance various software tools aimed at reducing diagnosis delay, such as the Orphanet platform and HPO codes, integrating this data into the RDK application.

As mentioned above, the results of our comparative analysis between the medical records of patients followed at Necker—Enfants Malades Hospital and our diagnostic fact sheets validate the diagnostic value of our sheets, showing a strong alignment between the fact sheets and knowledge bases in capturing key symptoms. The usefulness of our fact sheets in collecting clinically relevant information is also emphasized. We know in real-world data the percentage of patients who have certain symptoms. However, precise quantification and statistical analysis were not feasible due to the absence of standardized coding for the signs in the diagnostic fact sheets and differences in granularity. Consequently, the analysis primarily emphasizes qualitative comparisons rather than quantitative metrics, limiting the ability to statistically validate the clinical significance of the captured phenotypes.

Interestingly, in our study, for one disease, an analysis of symptoms recorded in patient files revealed a condition not typically associated with it, suggesting a potential underlying cardiac involvement in LPI patients.

Although cardiac symptoms have been reported in two patients in the literature [12], cardiac involvement as a specific manifestation of LPI is not clearly established in the medical community. In our records, cardiac presentations were heterogeneous. The most frequent finding was valvular disease, observed in four patients. Other findings included ischemic changes (one patient) and a range of conditions such as arrhythmia, cardiomegaly, and ventricular hypertrophy. Thus, the presence of diverse cardiac abnormalities in 12 patients, although heterogeneous and sometimes nonspecific, raises questions about a possible association between LPI and cardiac involvement. Artificial intelligence may reveal symptoms not described in initial published cases or new age-related symptoms; in LPI, this finding merits further study and may encourage clinicians to screen for cardiac involvement. The mechanisms of cardiac damage, if related to the disease, remain unknown, although proposed hypotheses include immune deficiency and abnormalities in NO metabolism [13, 14]. Because cardiac involvement in LPI is not formally established and our fact sheets are intended as diagnostic aids for non-specialists focusing on the core recognized features of the disorder, we did not include it in the fact sheet.

Thus, further updates could be considered after systematically comparing all fact sheets with phenotypic data extracted from hospital medical records, particularly by integrating clinical and biological phenotypes with patient age and aging cohorts. These tools may provide insights into the outcomes of elderly patients.

In conclusion, the creation of these "disease" and "symptom" fact sheets is the result of collaboration among experts from various centers and disciplines, coordinated by the French IMD network, "Filière G2m." These fact sheets provide a concise resource on rare diseases within IMDs, aiming to address diagnostic delays by highlighting critical symptoms for referral to specialized centers. They are freely accessible on the Filière G2m website and will soon be distributed more widely. Our sheets will be used as a database for an artificial intelligent tool currently under development, accessible directly on the website of our French IMD network, to facilitate IMD diagnoses. They could also serve to update existing diagnostic support software for rare diseases.

Author Contributions

A.C., X.C., A.K., M.Z., and P.D.L. contributed to the design, planning, conduct, and reporting of the work described in this article. A.C. developed the first version of the diagnostic fact sheets, and all authors and the other members of the expert group consortium contributed to the review and finalization of the fact sheets. A.-L.C. and A.K. were responsible for distributing and communicating these protocols through the G2M website, meetings, newsletters, webinars, and other platforms. A.C., X.C., A.K., and P.D.L. contributed to drafting and critically reviewing the manuscript. All authors reviewed and approved the final version.

Acknowledgements

The authors would like to express their gratitude to Anne-Sophie Lapointe, project manager at the DGOS (Direction Générale de l'Offre de Soins) of the French Ministry of Health, and to the DGOS for its actions and funding in support of rare disease networks through the national plan for rare diseases. The authors would also like to acknowledge Marielle Romet and Francoise Nourrit-Poirette (Synergy Pharm—Santé Active Edition) for conception of the first fact sheets, Callimedia for the formatting, and Soltengroup for translating the fact sheets into English. The authors would like to thank Pamela Dushime for the publication and presentation of the data sheets on the G2M network website.
Experts in metabolic diseases: Marie-Thérèse Abi-Warde, Cécile Acquaviva, Jean-Baptiste Arnoux, Stéphanie Badiou, Magalie Barth, Nadia Belmatoug, Jean-François Benoist, Juliette Bouchereau, Anais Brassier, Arnaud Bruneel, Catherine Caillaud, Aline Cano, Brigitte Chabrol, David Cheillan, Emmanuelle Corbe-Guillard, Christelle Corne, Lena Damaj, Myriam Dao, Pascale De Lonlay, Anne-Frédérique

Dessein, Dries Dobbelaere, Claire Douillard, Thierry Dupré, François Feillet, Roseline Froissart, Margaux Gaschignard, Magali Gorce, Laurent Gouya, Anne-Sophie Guemann, Bénédicte Héron, Célia Hoebeke, Apolline Imbard, Elsa Kaphan, François Labarthe, Philippe Labrune, Pascal Laforet, Thierry Levade, Elise Lebigot, Edouard Le Guillou, Olivier Lidove, Julien Maquet, Wladimir Mauhin, Clothilde Marbach, Karin Mazodier, Karine Mention, Fanny Mochel, Caroline Moreau, Yann Nadjar, Esther Noel, Mickael Obadia, Cécile Pagan, Magali Pettazzoni, Samia Pichard, Clement Pontoizeau, Aurélia Poujois, Isabelle Redonnet-Vernhet, Frédérique Sabourdy, Manuel Schiff, Christine Serratrice, Aude Servais, Caroline Sevin, Anne Spraul, Bénédicte Sudrié, Marine Tardieu, Sandrine Vuillaumier, Camille Wicker, Arnaud Wiedemann-Fode. **Other medical experts:** Vincent Barlogis (pediatric hematologist), Nathalie Boddaert (radiologist), Kanetee Busiah (pediatric endocrinologist), Annabelle Chaussenot (neurologist), Dominique Debray (pediatric hepatologist), Céline Falaise (pediatric hematologist), Muriel Girard (pediatric hepatologist), Dalila Habes (pediatric hepatologist and gastroenterologist), Annie Harroche (pediatric hematologist), Florence Lacaille (pediatric hepatologist), Mehdi Oualha (pediatric intensive care specialist), Caroline Ovaert (pediatric cardiologist), Rachel Reynaud (pediatric endocrinologist), Caroline Rousset-Rouvière (pediatric nephrologist), Cécile Rouzier (geneticist), Karim Wahbi (cardiologist). Open access publication funding provided by COUPERIN CY26.

Funding

The authors have nothing to report.

Ethics Statement

Ethics approval was not required for this study.

Consent

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. DGOS, "Les maladies rares," accessed July 29, 2025, https://sante.gouv.fr/soins-et-maladies/prises-en-charge-specialisees/maladies-rares/article/les-maladies-rares.

2. T. Pampols, "Inherited Metabolic Rare Disease," *Advances in Experimental Medicine and Biology* 686 (2010): 397–431.

3. PFMG2025 Contributors, "*PFMG2025*-Integrating Genomic Medicine Into the National Healthcare System in France," *Lancet Regional Health—Europe* 50 (2025): 101183.

4. J. Bouchereau, C. Wicker, K. Mention, et al., "Standardized Emergency Protocols to Improve the Management of Patients With Suspected or Confirmed Inherited Metabolic Disorders (IMDs): An Initiative of the French IMDs Healthcare Network for Rare Diseases," *Molecular Genetics and Metabolism* 143, no. 1–2 (2024): 108579.

5. HAS, "Protocoles Nationaux de Diagnostic et de Soins (PNDS)," accessed July 29, 2025, https://www.has-sante.fr/jcms/c_1340879/fr/protocoles-nationaux-de-diagnostic-et-de-soins-pnds.

6. N. Garcelon, A. Neuraz, V. Benoit, et al., "Finding Patients Using Similarity Measures in a Rare Diseases-Oriented Clinical Data Warehouse: Dr. Warehouse and the Needle in the Needle Stack," *Journal of Biomedical Informatics* 73 (2017): 51–61.

7. J.-M. Saudubray, F. Mochel, F. Lamari, and A. Garcia-Cazorla, "Proposal for a Simplified Classification of IMD Based on a Pathophysiological Approach: A Practical Guide for Clinicians," *Journal of Inherited Metabolic Disease* 42, no. 4 (2019): 706–727.

8. MetabERN, "MetabERN: European Reference Network for Metabolic Disorders," accessed July 29, 2025, https://metab.ern-net.eu.

9. J. J. Y. Lee, W. W. Wasserman, G. F. Hoffmann, C. D. M. van Karnebeek, and N. Blau, "Knowledge Base and Mini-Expert Platform for the Diagnosis of Inborn Errors of Metabolism," *Genetics in Medicine* 20, no. 1 (2018): 151–158.

10. RDK, accessed February 15, 2026, https://rdk.asweknow.com/assistant.

11. Orphanet, "Orphanet: Diseases," accessed July 29, 2025, https://www.orpha.net/fr/disease.

12. W. Mauhin, F. Habarou, S. Gobin, et al., "Update on Lysinuric Protein Intolerance, a Multi-Faceted Disease Retrospective Cohort Analysis From Birth to Adulthood," *Orphanet Journal of Rare Diseases* 12, no. 1 (2017): 3.

13. Y. Kamada, H. Nagaretani, S. Tamura, et al., "Vascular Endothelial Dysfunction Resulting From l-Arginine Deficiency in a Patient With Lysinuric Protein Intolerance," *Journal of Clinical Investigation* 108, no. 5 (2001): 717–724.

14. M. Lukkarinen, K. Parto, O. Ruuskanen, et al., "B and T Cell Immunity in Patients With Lysinuric Protein Intolerance," *Clinical and Experimental Immunology* 116, no. 3 (1999): 430–434.