Adv Ther (2022) 39:3361–3377
https://doi.org/10.1007/s12325-022-02186-2

ORIGINAL RESEARCH

Estimating the Prevalence of Rare Diseases: Long-Chain Fatty Acid Oxidation Disorders as an Illustrative Example

Eliza Kruger · Paul McNiven · Deborah Marsden

Received: February 10, 2022 / Accepted: May 10, 2022 / Published online: June 8, 2022

ABSTRACT

Introduction: Determining the epidemiology of disease is critical for multiple reasons, including to perform risk assessment, compare disease rates in varying populations, support diagnostic decisions, evaluate health care needs and disease burden, and determine the economic benefit of treatment. However, establishing epidemiological measures for rare diseases can be difficult owing to small patient populations, variable diagnostic techniques, and potential disease and diagnostic heterogeneity. To determine the epidemiology of rare diseases, investigators often develop estimation models to account for missing or unobtainable data, and to ensure that their findings are representative of their desired patient population.

Methods: A modeling methodology to estimate the prevalence of rare diseases in one such population—patients with long-chain fatty acid oxidation disorders (LC-FAOD)—as an illustrative example of its applicability.

Results: The proposed model begins with reliable source data from newborn screening reports and applies to them key modifiers. These modifiers include changes in population growth over time and variable standardization rates of LC-FAOD screening that lead to (1) a confirmed diagnosis and (2) improvements in standards of care and survival estimates relative to the general population. The model also makes necessary assumptions to allow the broad applicability of the estimation of LC-FAOD prevalence, including rates of diagnosed versus undiagnosed patients in the USA over time.

Conclusions: Although each rare disease is unique, the approach described here and demonstrated in the estimation of LC-FAOD prevalence provides the necessary tools to calculate key epidemiological estimates useful in performing risk assessment analyses; comparing disease rates between different subgroups of people; supporting diagnostic decisions; planning health care needs; comparing disease burden, including cost; and determining the economic benefit of treatment.

Keywords: Prevalence model; Rare diseases; Long-chain fatty acid oxidation disorders

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12325-022-02186-2.

E. Kruger (✉) · D. Marsden
Ultragenyx Pharmaceutical, Inc., 60 Leveroni Ct, Novato, CA 94949, USA
e-mail: EKruger@ultragenyx.com

P. McNiven
Humanumeric LLC, Austin, TX, USA

Key Summary Points

Why carry out this study?

Modeling the epidemiology of rare diseases, including long-chain fatty acid oxidation disorders (LC-FAOD), is challenging because of limited data availability and heterogeneous disease presentation.

This study presents a prevalence estimation model of rare diseases using LC-FAOD as an illustrative example of its applicability.

What was learned from the study?

This study demonstrates a model for estimating the prevalence of rare diseases supported by an illustrative example of LC-FAOD that demonstrates increases in prevalence and diagnostic rate associated with the gradual implementation of newborn screening practices in the USA.

The findings from the present study highlight an unmet need for research into LC-FAOD, particularly with regards to mortality and the impact of treatment on outcomes in these patients.

INTRODUCTION

Epidemiological assessments hold tremendous value in the evaluation and understanding of disease. The prevalence of a disease identifies the total number of people affected at a given time and place, accounting for rates of mortality or patient survival and cures, if applicable, whereas the incidence rate of a disease indicates the number of new cases in a location or over a period of time [1]. With these data, researchers and clinicians can compare disease rates between different subgroups of people to perform risk assessment analyses, support diagnostic decisions, plan health care needs, evaluate disease burden, and determine the economic benefit of treatment [1–3]. However, for epidemiological metrics to be useful in these applications, they must be both reliable and applicable to different subgroups of a population of interest. One way to ensure this is to design epidemiological studies such that they encompass a large, diverse sample of patients, representing various geographic and sociological subpopulations [1]. This becomes a challenge when examining rare diseases for which the patient population is sparse, and where identifying a representative sample population to accurately estimate disease prevalence is challenging.

Rare diseases are characterized as those affecting fewer than five people per 10,000 [4] or less than 200,000 with the disease in the USA [5]. In the USA, it is estimated that approximately 30 million individuals are affected by rare diseases [6]. With more than 7000 rare diseases identified, some more rare than others, these data highlight the limited availability of patient data for any one disease [6]. Deriving meaningful epidemiological data from such a limited pool poses several challenges [7]. Existing epidemiological reports are often scarce and may not be standardized. Similarly, diagnostic criteria and methods may vary, potentially biasing the apparent incidence or prevalence of a disease of interest [7]. Disease heterogeneity may alter patient classification, further limiting the patient pool from which to assess epidemiological data for a specific disease [6].

In order to make appropriate estimates of prevalence, mortality data are also needed. Assessing the mortality rates of patients with rare diseases is also challenging, particularly in diseases for which diagnosis is made on the basis of symptom presentation, as death can occur prior to a clear diagnosis [8]. Finally, there may be differences across geographic locations, with certain regions presenting higher incidence and prevalence rates owing to genetic susceptibility or the availability of diagnostic tools and health care [7]. To account for these potential pitfalls, epidemiological studies of rare diseases must make certain assumptions and estimations to develop the best possible model from limited available data [7].

Long-chain fatty acid oxidation disorders (LC-FAOD) comprise one group of rare genetic metabolic disorders, affecting the transport of fatty acids into the mitochondrion via the carnitine shuttle system or mitochondrial β-oxidation pathway [9, 10]. In healthy individuals, this pathway is responsible for the conversion of fatty acids to energy as ATP, or ketones (for the body during periods of prolonged fasting or during other forms of metabolic stress, e.g., acute illness, when other energy sources are exhausted) [11]. Because this process is impaired in patients with LC-FAOD, it can lead to metabolic crises, particularly in organ systems high in energy demand [12]. Within the group of LC-FAOD, there are several variants with differing symptomology, classified by the specific underlying enzyme deficiency: very-long-chain acyl-CoA dehydrogenase deficiency (VLCAD), long-chain 3-hydroxy acyl-CoA dehydrogenase deficiency (LCHAD), trifunctional protein deficiency (TFP), carnitine palmitoyltransferase type 1 (CPTI) deficiency, carnitine-acylcarnitine translocase (CACT) deficiency, CPT type 2 (CPTII) deficiency, and carnitine transport deficiency (CTD) [11].

Historically, patients with LC-FAOD were identified by an initial presentation of acute metabolic decompensation that includes cardiomyopathy, hypoketotic hypoglycemia, liver dysfunction, lethargy, muscle weakness, myalgia, rhabdomyolysis, renal damage, or sudden death [11]. Presentation can occur at any age; however, the most serious, life-threatening episodes most frequently occur shortly after birth. Thus, it is critical that patients be identified early in life to preempt or minimize serious, life-threatening events through prompt disease management [11].

The introduction of newborn screening (NBS) programs for LC-FAOD in the late 1990s allowed for earlier detection of patients, including those with asymptomatic disease [13]. Observed disease prevalence increased owing to the identification of this subset of previously undiagnosed patients and to longer survival through earlier initiation of disease management [13, 14]. NBS for LC-FAOD introduces a particular challenge for assessing the epidemiology of this disease, as countries or states have incorporated LC-FAOD on NBS panels at different rates or have included some, but not all, variants of the disease [13, 15].

The current article explores a modeling methodology in proposing an approach to calculate epidemiological metrics for estimating disease prevalence using LC-FAOD as an example of its applicability. As with all rare diseases, a scarcity of reliable data on LC-FAOD from which to derive epidemiological outcomes poses a considerable challenge. Therefore, the present model includes a number of critical assumptions and estimations necessary to provide a best possible estimate of the prevalence of LC-FAOD in the USA.

METHODS

Literature Search

Epidemiological data were assessed using a targeted literature search. Databases were searched using PubMed, Medline, and Google, with no date restrictions on results. Search terms included, but were not restricted to, the following: LC-FAOD terms ("long-chain fatty acid oxidation disorders," "long-chain 3-hydroxy-acyl-CoA dehydrogenase deficiency," "very long-chain acyl-CoA dehydrogenase deficiency," "trifunctional protein defects," "carnitine palmitoyl transferase deficiency," "carnitine-acylcarnitine translocase deficiency," "LC-FAOD," "LCHAD," "VLCAD," "TFP," "MTP," "CPTI," "CPTII," "CACT") combined with a specific outcome term ("incidence," "prevalence," "mortality," "death").

LC-FAOD Disease Prevalence Model

Overview

A model was programmed in Microsoft Excel 2010 (Microsoft Corporation, Redmond, WA, USA) to forecast the prevalence of LC-FAOD from incidence-derived prevalence data. Incidence-derived prevalence was determined as a combination of calculated incidence based on known sources, population, mortality, and rate of diagnosis at a given time point based on the

Table 1 Summary of epidemiology and genetic inputs

Input	Parameter	Source
Population		
Global birth rate (historical and forecast)	Year (1950–2040) Data beyond 2019 forecast Projection accounts for mortality, migration, and fertility rates. Low, medium, and high variants available	Derived from United Nations Population Division (2019) [25]
Incidence of LC-FAOD (%)		
Netherlands	0.0025	Derived from Diekman et al. (2016) [17], adjusted for subtype incidence
Germany	0.0020	Derived from DGNS National Screening Reports (2005–2016) [23]
USA		
US NBS+	0.0056	Derived from: National Newborn Screening & Global Resource Center (NNSGRC) (2006) [22] Martin et al. (2011) [20] 2010 US census data
Non-Alaska US NBS+	0.0047	
Alaska NBS+	0.2661	
Native Alaskan NBS+	1.4216	Derived from: 2010 US census data Martin et al. (2009) [20]
US true positive rate	0.0022	Derived from 2006 NBS data by state adjusted for false positive rate
Global true positive rate	0.0020	Triangulation of: Diekman et al. (2016) [17] Bhattacharya et al. (2016) [16] Merritt 2nd et al. (2014) [21] National Newborn Screening & Global Resource Center (2006) [22]
Variant distribution		
VLCAD	59.3	Derived from DGNS National Screening Reports (2005–2016) [23]
LCHAD	25.1	
TFP	6.6	Derived from Sander et al. (2005) [24]

Table 1 continued

Input	Parameter	Source
CPT-1	4.8	Derived from DGNS National Screening Reports (2005–2016) [23]
CPT-2	3.0	
CACT	1.2	
Newborn screening of LC-FAOD		
Proportion of infants with locally screened variants (%)	98.3	Derived from NewSTEPs (2021) [15]
Screening rate (%)	99.9	National Institutes of Health (2017) [18]
Overall proportion captured by NBS (%)	98.2	Derived from the proportion of infants with screenable variants and the national screening rate
Mortality		
% proportion of newborns in the general population surviving to age:		Derived from Heron (2009) [19]
5 years	99.28	
10 years	99.21	
15 years	99.12	
20 years	98.81	
25 years	98.39	
30 years	97.97	
35 years	97.47	
40 years	96.74	
Life expectancy (years)	78.1	
Excess mortality due to LC-FAOD (%)		

Table 1 continued

Input	Parameter	Source
Pre NBS		Assumption derived from clinical expertise and based on current standard of care and intensity of disease presentation
Undiagnosed	0.24–18.00	
Diagnosed symptomatically	0.80–18.00	
Diagnosed by NBS	0.80–18.00	
Post NBS		
Undiagnosed	0.24–10.00	
Diagnosed symptomatically	0.00–5.00	
Diagnosed by NBS	0.00–4.00	

CACT carnitine-acylcarnitine translocase deficiency, *CPT-1* carnitine palmitoyltransferase 1, *CPT-2* carnitine palmitoyltransferase 2, *DGNS* German Society for Newborn Screening, *LC-FAOD* long-chain fatty acid oxidation disorders, *LCHAD* long-chain 3-hydroxy acyl-CoA dehydrogenase deficiency, *NBS* newborn screening, *NBS+* newborn screening positive, *TFP* trifunctional protein, *VLCAD* very-long-chain acyl-CoA dehydrogenase deficiency

availability and implementation of diagnostic technology, as well as disease management strategies. These inputs are summarized in Table 1 and detailed below [15–24]. This article is based on previously conducted studies and does not contain any new studies with human participants or animals performed by any of the authors.

Population Inputs

Population estimates were captured, beginning in 1950 to ensure that relevant population swings due to historical events would be included. There are few patients in the world with known LC-FAOD older than 70 years, as LC-FAOD were only first recognized in the 1970s. In this population, affected patients from this time may have not survived or remain undiagnosed.

To project live births for a given year in the current model, population growth was derived from United Nations (UN) birth rate estimates, which account for multiple parameters, including age, sex, mortality, migration, and fertility factors [25]. Several scenarios for fertility rates are generated by the UN that in turn dictate the number of projected live births. The three fertility rates incorporated in the model (low/medium/high) represent the boundaries of the likely forecast range. The medium variant was utilized for calculations of LC-FAOD, but the other variants are available for sensitivity analyses.

Incidence Inputs

Disease incidence was broadly defined as the number of individuals born within a given year with confirmation of LC-FAOD types per the annual number of live births. The baseline global incidence rate (technically, prevalence at birth) was calculated as the sum of all confirmed positive LC-FAOD cases divided by the total number of live births during that interval.

Incidence inputs were derived from available literature (Supplementary Material, Table S1a, b). Germany has had complete NBS in place for decades, with a specific focus on fatty acid oxidation disorders since at least 2005 [23]. German NBS incidence data from the German Society for Neonatal Screening were compared with other country populations, incorporating modifications for the USA and Canada to account for known population differences due to specific disease variants (e.g., Arctic variant)

[23]. Coverage for German NBS data is essentially 100%, and only confirmed cases of LC-FAOD are reported. This provides a cumulative population-representative sample of over 8 million live births across all years analyzed. Incidence rates from Germany were then triangulated with incidence data from the Netherlands [17] and the USA [21], and confirmed by expert opinion [26] to determine the overall incidence rate for the model.

Diagnostic Rate Estimates

Diagnostic rates differed on the basis of several factors: availability of NBS, proportion of the population to receive NBS, age, and symptomology screened. To estimate the diagnostic rate of LC-FAOD over time, rates of symptomatic diagnosis derived from clinical expertise and qualitative expert interviews were used and the gradual standardization of LC-FAOD on NBS panels and diagnostic rate with NBS were incorporated into the model. This was because, globally, the standardization of LC-FAOD on NBS panels occurred at varying rates [13, 15]. We assumed that almost 100% of patients with LC-FAOD who underwent NBS were diagnosed.

Prior to the adoption of LC-FAOD on NBS panels, the clinical diagnostic rate for a given year was modeled as a composite curve (linear curve, S-curve, or rapid uptake curve) in the basis of clinical expertise, factoring in patient age and symptom intensity. Patients with greater disease intensity were assigned a greater likelihood of diagnosis, and the diagnostic rate was assumed to increase with patient age as the likelihood of recognizing disease characteristics following metabolic decompensation increased. The peak diagnostic rate was defined as the proportion of patients among non-NBS infants that would ever be diagnosed. Here, asymptomatic patients were assumed to achieve a peak diagnostic rate of 20% at age 30 years, whereas symptomatic patients with increasing symptom intensities had corresponding increasing peak diagnostic rates of 40% (peak age 12 years), 80% (peak age 4 years), and 95% (peak age 2 years) for mild, moderate, and intense symptom presentations, respectively. With the gradual standardization of NBS panels, a shift in the diagnostic rate of LC-FAOD was included in the model, nearing a 100% diagnostic rate once fully adopted.

Data on the precise timing of when and where NBS programs were introduced, and the fraction of newborns within the country covered are difficult to come by. Data from the USA demonstrated the adoption of LC-FAOD screening onto NBS programs from 2001 to 2006; however, the USA is a special case as NBS is regulated at the state level. Therefore, uptake in this model was coded as a series of binary events at a point in time in each state. For the rest of the world, NBS adoption rates were based on the review conducted by Therrell et al. [27]. In addition, assumptions of the rate of NBS adoption over time were made on the basis of how well developed individual health care systems are across the world, as the primary technology for LC-FAOD screening, tandem-mass spectrometry, is quite advanced [28]. To be identified by NBS, infants must be born in a state in which their variant of LC-FAOD is part of the NBS panel, and must partake in the screening process. Estimations of these rates are quite high, with 98.3% of infants estimated to have screenable variants, and 99.9% receiving screening at its peak. Together, these estimates provide an overall percentage of patients captured by NBS in the USA of 98.2% at its peak in the year 2000 [15, 18].

Mortality Estimates

Survival is compromised in patients with LC-FAOD; however, the impact on mortality is not equal across all patients. Often, earlier onset of symptoms is associated with higher rates of mortality, whereas later onset, presenting with less intense signs and symptoms, and identification by NBS are linked with decreased mortality [10, 14, 24, 29]. In the current model, the best-case scenario survival curve for patients with LC-FAOD was assumed to be equivalent to the survival rate of the general population (Table 1) [19]. The survival rate of the general population was adjusted with a calculation for LC-FAOD excess mortality, which was the percentage of patients who die each year beyond that expected for their age in the general population (Table 1). These values were based on the survival of infants treated with the current

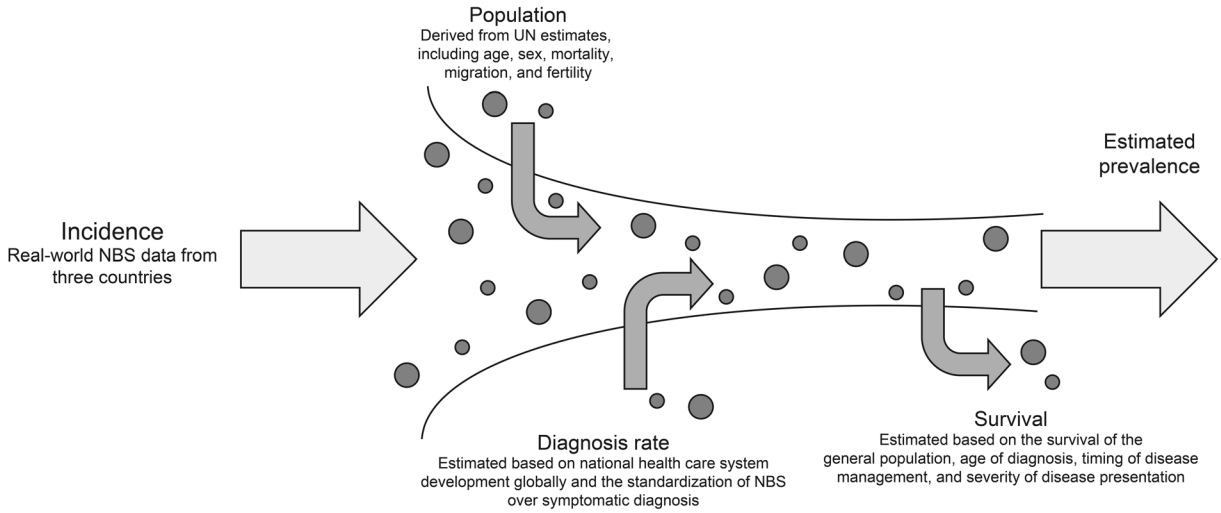

Fig. 1 Model overview. *NBS* newborn screening, *UN* United Nations

standard of care over recent years. Patients with earlier disease onset were assumed to have more intense disease presentation and, consequently, a higher mortality rate [10]. Similarly, those who were diagnosed before the introduction of NBS were also assumed to have a higher rate of mortality, as the initiation of disease management was delayed until a potentially life-threatening presentation.

The model input was created to account for excess mortality beyond the general population over time as diagnostic standards transitioned from symptomatic to NBS diagnosis, with the survival of undiagnosed (and therefore untreated) patients being unaffected by changing treatment standards. Excess mortality reflected intensity of symptoms, with undiagnosed patients with the most intense symptoms reflecting the greatest excess mortality (10.0%), and asymptomatic patients who received a diagnosis with the lowest excess mortality (0.0%). To adjust the model over time, other model inputs included year of change, denoting the year at which standards of care started moving from pre-NBS levels to the current level; and T_{max}, the length of time taken for mortality to stabilize to current levels after the year of change. In the USA, the year of change was designated as 1980, with an estimated 30-year period for the diagnostic standardization to peak. Mortality adjustments were also influenced by a composite curve, which could be selected as a linear, S-curve, or rapid uptake curve depending on the country or region. Finally, mortality rates were further adjusted to align with known incidence rates. For example, if a given childhood mortality rate in a certain year resulted in a subsequent prevalence estimate in adult patients that was too low compared with known values, the childhood mortality rate would be lowered, yielding a greater survival rate into adulthood and corresponding increase in prevalence.

Prevalence Calculation

All curves in the model were subject to change over time, where prevalence as a function of age in a given year was a composite effect of all previously mentioned drivers and the estimates of their change over time (past and future). For a given year, estimated disease prevalence was determined by combining the population estimate, the calculated incidence rate, the likelihood of diagnosis in patients based on available technology, the adoption of screening techniques, and the estimated mortality of all previously diagnosed patients at the evaluation year, based on timing of the diagnosis and implementation of disease management. This process is summarized in Fig. 1. Note that the observed prevalence ratio may also be impacted by cure rates (i.e., some people may be cured,

therefore the overall prevalence is reduced). While this is not the case for LC-FAOD currently, this may be relevant for other diseases.

Mortality Scenario Analyses

As a result of the rarity of the disease, there is a dearth of literature on mortality rates of LC-FAOD, particularly for individuals born pre-NBS. The base case mortality estimate was based on published data and clinical expertise, however there was significant uncertainty associated with this estimate. As such, we vary the base case mortality assumption by 50% and 200% to assess the impact of changing mortality assumptions on the number of patients diagnosed over time.

RESULTS

Literature-Reported Incidence

Several sources on LC-FAOD incidence were identified from the literature search. The most robust source was the German National Screening Report, which provides a complete statistical analysis of the near-universal screening results for metabolic and endocrine diseases, including recall rates and confirmed diagnoses (Supplementary Material Table S1a) [23]. Additional incidence data from a Dutch NBS study (October 2007–2010) and US NBS data sources [17, 21, 22, 26] were comparable to those of the German study, despite representing separate countries with independent testing protocols. Results from the Dutch study identified 11 (0.0015%) positive cases of VLCAD, a subtype of LC-FAOD, out of 742,728 screens [17]. The comprehensive German NBS data indicated that approximately 59% of the total LC-FAOD cases were of the VLCAD subtype [23]. Taken together, the 0.0015% VLCAD incidence rate accounts for approximately 59% of the total LC-FAOD incidence rate; from this, the overall LC-FAOD incidence rate can be approximated at 0.0025% (Supplementary Material Table S1b). Although US screening data indicated a higher incidence of LC-FAOD initially (0.0056%), confirmatory diagnostic testing identified the true-positive cases (0.002%), reducing the incidence rate to levels comparable to the German and Dutch findings [17, 20–23, 26]. This agreement from unique sources suggests a global LC-FAOD incidence of approximately 0.002% [17, 21–23, 26]. These data are further supported by the results of independent publications identified in the preliminary literature search, which suggest an incidence between 0.001% and 0.004% [30–32]. The 0.002% incidence rate of LC-FAOD was then applied to the population estimate models created by the UN Population Division for each assessment year to obtain the overall incidence rate [25].

Diagnostic Rate

The model output represents the diagnostic rate as the proportion of undiagnosed versus diagnosed patients over time, as diagnostic methodologies improved from clinical diagnosis to NBS. Clinical diagnosis identified a proportion of patients with LC-FAOD on the basis of disease presentation, most commonly at a young age. However, older patients in a given year demonstrated a decreased ratio of undiagnosed to diagnosed patients, as there is a greater likelihood that patients who are initially undiagnosed will show symptoms with age and subsequently receive a diagnosis.

Because the specific global timing of the standardization of LC-FAOD on NBS panels is not readily available, the current model utilizes available literature with regard to its uptake [27], estimating any missing data on the basis of the status of a given region's health care systems. This model demonstrates an increased number of patients diagnosed with LC-FAOD over time, as diagnostic methodology transitions from relying on acute presentation in the clinical setting to near-universal screening at birth, which accurately identifies almost all patients. In turn, this translates to an overall increase in LC-FAOD prevalence with the initiation of NBS and a further increase over time as NBS is standardized [14].

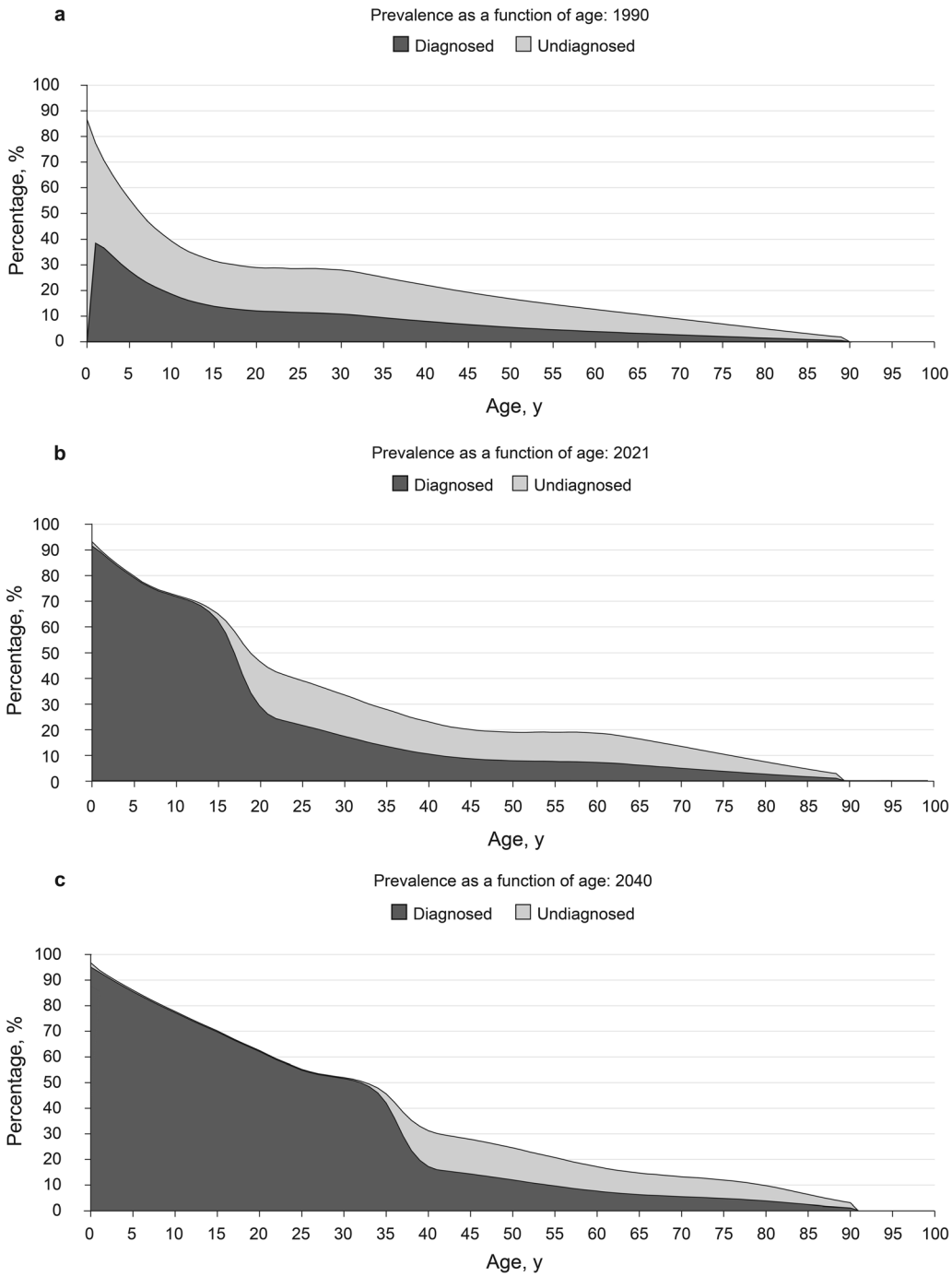

Fig. 2 Estimated percentage of diagnosed versus undiagnosed patients with LC-FAOD by age in **a** 1990, **b** 2021, and **c** 2040 in the USA. *LC-FAOD* long-chain fatty acid oxidation disorders

Mortality

As confirmed by the targeted literature search, no true mortality rate has been established for LC-FAOD owing to the high variability among age of onset, symptomology, age at diagnosis, and differences in sampling, e.g., time interval and sample sizes. This problem is evident by the greatly varying rates in reported mortality (disease lethality; summarized in Supplementary Material Table S2). The current model applies a survival curve that accounts for these factors to the known mortality rate of the general population (Table 1) to estimate the survival of patients with LC-FAOD. The resulting estimation illustrates a large increase in LC-FAOD prevalence in younger patients following the introduction of NBS, corresponding with a decreased mortality rate. Over time, this translates into a greater prevalence in older populations as more patients survive into adulthood as a result of earlier disease detection and management.

Estimated Prevalence

The prevalence estimation output of the model was variable by analysis year, as all previously mentioned drivers changed over time. Incidence-derived LC-FAOD prevalence was examined at key points. Figure 2 represents the estimated total prevalence of LC-FAOD as the proportion of diagnosed and undiagnosed patients by patient age. The introduction of LC-FAOD on NBS panels in 1990, prior to broad implementation, shows a sharp increase in diagnosed patients younger than 10 years, with the highest rate in newborns (Fig. 2a). Estimated prevalence of LC-FAOD in the USA in 1990 was 2030 cases, of which 43% (871) were pediatric cases and 57% (1159) were adult cases. The overall diagnosis rate was estimated to be 39%. Estimated prevalence in present-day 2021 is higher, driven by a greater proportion of diagnosed children as types of LC-FAOD have become more common on NBS panels (Fig. 2b); a shift in the proportion of diagnosed versus undiagnosed patients was observed in patients aged 10–20 years, increasing further in younger patients and newborns. We estimated 93 newborns with LC-FAOD in 2021. Further, the prevalence of LC-FAOD was approximately 2871 cases in the USA in 2021, of which 1355 were in children and 1515 in adults; the overall diagnostic rate was estimated at 72%. Finally, we forecast the prevalence of LC-FAOD in 2040, with the assumed near-universal incorporation on NBS panels (Fig. 2c). Over time, the diagnosed versus undiagnosed population continues to increase, nearing a 100% diagnosis rate in newborns. Meanwhile, patients originally identified by NBS at its inception are now aged 35–40, which is represented by the large surge in prevalence in this age group. The model predicted that there would be 3425 cases in the USA in 2040, of which 42% (1452) were pediatric cases and 58% (1973) were adult cases. The overall diagnostic rate was estimated to be 85%. Model outputs showing the estimated number of LC-FAOD cases worldwide over time are shown in Supplementary Material Table S3. Estimated prevalence increased from 56,245 cases with a 39% diagnosis rate in 1990 to 82,333 cases with a 62% diagnosis rate in 2021, and is forecast to reach 95,457 cases with a 71% diagnosis rate in 2040.

Scenario Analyses

In the scenario analysis, a 50% reduction in the base case mortality rates increased the prevalent number of patients to 3645 (127.0% increase from base case population estimate) and 4268 (124.6%) in 2021 and 2040, respectively. Increasing the mortality rate by 200% resulted in an estimated 2164 (75.4%) and 2616 (76.4%) patients with LC-FAOD in 2021 and 2040, respectively. The impact of changing mortality rates was greater in 2021 versus 2040 as there are more pre-NBS individuals, which have a higher mortality rate. Results of the mortality scenario analyses are available in the Supplementary Material.

DISCUSSION

Assumptions, Estimations, and Limitations

This model represents a method of estimating the incidence and prevalence of rare diseases. However, when modeling the epidemiology of rare diseases such as LC-FAOD, one of the greatest challenges is the distinct lack of reliable, broadly applicable data from which to draw accurate estimates. To fill in these missing knowledge gaps, models of rare diseases must make assumptions and estimations on the basis of the limited available data [7]. Each estimate comes with its own unique set of assumptions specific to the rare disease being modeled. The current model is no exception, as it begins with an estimated incidence rate derived from real-world data that is then applied to estimates of population, diagnostic rate, and survival rate over time to generate the best possible estimated prevalence curve on the basis of limited available data.

One of the most common challenges faced by epidemiological studies is the difficulty in representing all types of a population of interest, which may have inherently different incidence rates or different disease characteristics [6, 7]. The LC-FAOD incidence data used in this model were assumed from three key sources: German birth records with near-universal rates of NBS, US NBS data, and Dutch NBS data, all of which demonstrated consistent disease incidence rates of approximately 0.0020% (200/100,000 births) [17, 21–23, 26]. Therefore, the current model assumes that because these rates were consistent across multiple countries with different populations, diagnostic criteria, and rates of NBS adoption, this incidence rate of approximately 0.0020% can be extrapolated to represent the global disease incidence for LC-FAOD. However, it is important to note that there are certain countries or subpopulations within countries that have shown unique LC-FAOD incidence rates, including those of Chinese, Native Alaskan, Canadian First Nations, Saudi Arabian, and Hutterites descent, and this may be due in part to consanguinity within these populations [11, 33, 34]. The population size, screening rates, and incidence of these subpopulations may have a greater or lesser impact on the disease incidence of a larger assessment population. For example, a smaller subpopulation with a higher incidence rate than the rest of the assessment population due to inherent genetic differences would have an overall smaller impact on the overall incidence rate. However, when a high-incidence subpopulation accounts for a larger proportion of the overall assessment population, the overall incidence estimate will be overinflated. Therefore, to apply the current model globally or to assess the prevalence of other diseases, it must be assumed that incidence rates of distinct subpopulations such as these are comparable to those of the larger, known populations.

The current model applies the estimated diagnostic rate to the incidence population to determine the number of patients with LC-FAOD identified in a given year. Diagnostic rates of LC-FAOD present a particular challenge, as the rate of inclusion of this group of disorders on NBS panels is not uniform across the state or country level. Furthermore, the specific LC-FAOD types included on screening panels may vary, and the diagnostic criteria, rates of misdiagnosis, rates of delayed diagnosis, rates of incomplete diagnosis, and screening sensitivity may differ between testing centers [13]. This variability makes it challenging to estimate specific diagnostic rates both in the present day and in the future because diagnostic techniques evolve. To overcome this, the current model assumes a single diagnostic rate for patients of a given age in a given year, accounting for factors such as patient age, disease intensity, and diagnostic technique, wherever data were not available.

Perhaps the greatest assumption required in estimating the prevalence of LC-FAOD deals with the mortality rate of these patients. Presently, there is no true mortality rate for LC-FAOD available in the literature owing to the rarity of the disease, inconsistencies in diagnostic techniques, heterogeneous disease presentation, and lack of published research. Because the LC-FAOD patient population is so small, mortality rates of individual studies vary

considerably following a single death. Patient populations are further reduced when examining individual disease variants. Without large, targeted studies of mortality, it is difficult to accurately note the true mortality rate of LC-FAOD. Therefore, the mortality rates used in the present model were estimated on the basis of the best available data from a limited pool. Initial mortality rates were generated to align with known available prevalence by patient age. For example, if an estimated mortality rate was too high for a young population, the resulting prevalence in the adult population at a later time point would be lower than known data. To account for this, the earlier mortality rate would be decreased, corresponding with an increased survival into adulthood and corresponding increase in prevalence. Modeling the mortality rate over time required several assumptions to account for changes to diagnostic techniques and subsequent management strategies. It is assumed that, over time, care for patients with LC-FAOD has improved as knowledge of the disease has increased, translating to improved mortality. This model accounts for this by adjusting the curve such that a gradual increase in the survival rate is observed over time. Similarly, mortality rates of patients with LC-FAOD have reportedly improved following the standardization of NBS panels, as timely treatment management following diagnosis can reduce the occurrence of potentially life-threatening metabolic decompensations before symptoms present [10, 12]. However, it is unknown how rapidly the effects of these improvements will manifest as decreases in mortality rate. The current model assumes that in the USA, the effects of the improved diagnostic rate and earlier age of diagnosis, beginning with the implementation of NBS in 1990, will appear approximately 30 years from this date (2021). At this point, NBS for LC-FAOD is mostly standardized and the largest improvements in mortality and diagnostic rates are observed, represented by a large increase in LC-FAOD prevalence.

The lack of reliable data on the incidence, prevalence, and mortality of LC-FAOD is a key factor in one of the major limitations of this model. We have been unable to validate the model or the assumptions made within, as there are insufficient data available on the incidence and prevalence of LC-FAOD, and this problem serves as a common limitation of many models focusing on rare diseases.

Model Interpretation

Currently, some of the best sources of epidemiological data for LC-FAOD are national databases, including NBS records. However, these sources are restricted to specific regions and subpopulations and do not fully represent the global state of LC-FAOD [17, 21–23, 26]. This highlights the unmet need for an accurate, reliable method of estimating the prevalence of LC-FAOD that is applicable to the global population to perform risk assessment analyses; compare disease rates between different subgroups of people; support diagnostic decisions; plan health care needs; compare disease burden, including cost; and determine the economic benefit of treatment [1–3].

By applying the methodology described in this article to known LC-FAOD epidemiological data, we are able to develop a best estimate of the prevalence of this group of rare diseases in the past, and forecast it in the future. The estimated prevalence in 1990 represents a time prior to the standardization of LC-FAOD on NBS panels. The greater proportion of diagnosed versus undiagnosed patients younger than 10 years coincides with the historical method of diagnosing patients with LC-FAOD only after symptoms first present. Older patients depicted in the 1990 prevalence estimation demonstrate a greater proportion of undiagnosed versus diagnosed patients compared with those younger than 10 years. This represents the idea that patients who presented later in life or who remained asymptomatic often went undiagnosed prior to NBS standardization, as LC-FAOD often present with the most serious and obvious disease early in life [11]. In addition, relying on initial clinical presentation for diagnosis often meant that serious, life-threatening disease presented before treatment management could begin, increasing early mortality and decreasing the number of patients with LC-FAOD reaching

older age, as evidenced by the overall smaller prevalence estimate in patients older than 35 years.

The estimation of LC-FAOD prevalence in 2021 represents a time point approximately 20 years after the broad availability of LC-FAOD on NBS panels in the USA, but before its universal adoption. By 2005, 36 US states had incorporated LC-FAOD in their NBS panels, but even these only included the more common LC-FAOD types, VLCAD and LCHAD [13]. Patients who would once have been diagnosed in the clinical setting at an older age can now be diagnosed sooner owing to NBS, as illustrated by an even greater proportion of diagnosed versus undiagnosed patients between the ages of 10 and 20 in the 2021 prevalence estimation. Those patients who never developed clinical disease or who presented later in life would now be diagnosed at birth and are included in the prevalence estimation. In addition, patients diagnosed earlier in life by NBS can begin early disease management, preventing life-threatening disease manifestations and increasing survivability, as shown by a greater LC-FAOD prevalence in older patients in the 2021 prevalence estimate compared with the 1990 time point [10, 11, 35].

Finally, we forecast the future LC-FAOD prevalence in 2040. With near-universal adoption of LC-FAOD on NBS panels comes an increased diagnostic rate, as evidenced by the proportion of diagnosed versus undiagnosed patients nearing 100% in newborns. The continuation of prompt disease management after diagnosis minimizes life-threatening metabolic decompensations, leading to increased survivability into adulthood. The first patients to receive a diagnosis as a result of NBS are now reaching 35 years of age, represented by a far greater prevalence compared with patients of that age in the 1990 estimate. It is expected that with continued improvements in diagnostic techniques, rates, and disease management, disease prevalence will plateau, as nearly all patients with LC-FAOD receive prompt diagnosis at birth and proper disease management through life.

CONCLUSIONS

Although the current model still provides only an estimated prevalence for LC-FAOD, we recommend that the underlying methodology is applied to other rare diseases in future studies, whether or not NBS is available. In the case of LC-FAOD, the availability of NBS data has increased and with it the confidence with which incidence and prevalence can be estimated. We therefore recommend that NBS programs are extended to all rare diseases that can be detected this way. It is important that similar studies of rare diseases begin with quality, reproducible data as a foundation. From there, we recommend that drivers, including diagnostic rate, survivability, and population growth, are estimated on the basis of real-world events specific to the rare disease in question to derive an accurate prevalence estimate. It is important to note that these drivers are variable over time and should be calculated as such. Finally, we recommend that key assumptions and estimations must be included transparently for the disease being assessed. Any estimation of prevalence is only as accurate as the available source data, and, in the case of rare diseases, these data are quite limited owing to small sample sizes, incomplete data sets, and a lack of published research.

In the case of LC-FAOD, early diagnosis and prompt disease management have greatly improved survivability in patients; however, there still exists premature mortality beyond that of the general population, and these patients experience an overall decreased quality of life. Moreover, there exists a population of undiagnosed patients who live with impaired lives. Early diagnosis by NBS gives all patients with LC-FAOD, both symptomatic and asymptomatic, the opportunity to initiate early disease management and potentially increase the quality of their lives. Together, these observations emphasize the need for further research on LC-FAOD, particularly surrounding mortality and the impact of treatment on outcomes in these patients.

ACKNOWLEDGEMENTS

Funding. Sponsorship for this work and Rapid Service and Open Access Fees were funded by Ultragenyx Pharmaceutical Inc.

Medical Writing, Editorial, and Other Assistance. The authors would like to acknowledge Kate Morse and Jennifer Black-Shinn for their contributions to the present analysis. Third-party writing assistance was provided by Jack Pike, PhD, CMPP, of Envision Pharma Group, and funded by Ultragenyx Pharmaceutical Inc.

Authorship. All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this article, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

Author Contributions. All authors contributed to the study conception and design, model creation, interpretation of the results. All authors contributed to the drafting of the manuscript. All authors reviewed the manuscript and approved the final version of the manuscript.

Disclosures. Deborah Marsden and Eliza Kruger are Employees of Ultragenyx Pharmaceutical Inc. Paul McNiven is a consultant for Ultragenyx and was paid as a vendor to assist with the production of the LC-FAOD Disease Prevalence Model.

Compliance with Ethics Guidelines. This article is based on previously conducted studies and does not contain any new studies with human participants or animals performed by any of the authors.

Data Availability. The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

REFERENCES

1. Ward MM. Estimating disease prevalence and incidence using administrative data: some assembly required. J Rheumatol. 2013;40(8):1241–3. https://doi.org/10.3899/jrheum.130675.

2. Burns CJ, Wright JM, Pierson JB, et al. Evaluating uncertainty to strengthen epidemiologic data for use in human health risk assessments. Environ Health Perspect. 2014;122(11):1160–5. https://doi.org/10.1289/ehp.1308062.

3. Pearson I, Rothwell B, Olaye A, Knight C. Economic modeling considerations for rare diseases. Value Health. 2018;21(5):515–24. https://doi.org/10.1016/j.jval.2018.02.008.

4. European Parliament and of the Council Regulation (EC) 141/2000 of 16 December 1999 on Orphan Medicinal Products. OJL 018/1, 22 Jan 2000.

5. U.S. Food and Drug Administration. Office of Orphan Products Development. https://www.fda.gov/about-fda/office-clinical-policy-and-programs/office-orphan-products-development.

6. Whicher D, Philbin S, Aronson N. An overview of the impact of rare disease characteristics on research methodology. Orphanet J Rare Dis. 2018;13:14. https://doi.org/10.1186/s13023-017-0755-5.

7. Auvin S, Irwin J, Abi-Aad P, Battersby A. The problem of rarity: estimation of prevalence in rare disease. Value Health. 2018;21(5):501–7. https://doi.org/10.1016/j.jval.2018.03.002.

8. Rowe AK. Analysis of deaths with an unknown cause in epidemiologic analyses of mortality burden. Trop Med Int Health. 2006;11(4):540–50. https://doi.org/10.1111/j.1365-3156.2006.01581.x.

9. Bonnett D, Martin D, de Lonlay P, et al. Arrythmias and conduction defects as presenting symptoms of fatty acid oxidation in children. Circulation. 1999;100(22):2248–53. https://doi.org/10.1161/01.cir.100.22.2248.

10. Wajner M, Amaral AU. Mitochondrial dysfunction in fatty acid oxidation disorders: insights from human and animal studies. Biosci Rep. 2015;36(1):e00281. https://doi.org/10.1042/BSR20150240.

11. Merritt JL, 2nd, Norris M, Kanungo S. Fatty acid oxidation disorders. Ann Transl Med. 2018;6(24):473. https://doi.org/10.21037/atm.2018.10.57.

12. Spiekerkoetter U, Bastin J, Gillingham M, Morris A, Wijburg F, Wilcken B. Current issues regarding treatment of mitochondrial fatty acid oxidation disorders. J Inherit Metab Dis. 2010;33(5):555–61. https://doi.org/10.1007/s10545-010-9188-1.

13. Shekhawat PS, Matern D, Strauss AW. Fetal fatty acid oxidation disorders, their effect on maternal health and neonatal outcome: impact of expanded newborn screening on their diagnosis and management. Pediatr Res. 2005;57(5 Pt 2):78R-86R. https://doi.org/10.1203/01.PDR.0000159631.63843.3E.

14. Spiekerkoetter U. Mitochondrial fatty acid oxidation disorders: clinical presentation of long-chain fatty acid oxidation defects before and after newborn screening. J Inherit Metab Dis. 2010;33(5):527–32. https://doi.org/10.1007/s10545-010-9090-x.

15. NewSTEPs. Newborn screening status for all disorders. 2021. https://www.newsteps.org/resources/data-visualizations/newborn-screening-status-all-disorders?q=resources/newborn-screening-status-all-disorders. Accessed 11 April 2022.

16. Bhattacharya K, Carpenter K, Ho G, Devanapalli B, Wilcken B, Wiley V. SSIEM 2016 annual symposium - abstracts:Rome, Italy, September 2016. J Inherit Metab Dis. 2016;39(Suppl 1):35–284. https://doi.org/10.1007/s10545-016-9969-2.

17. Diekman E, de Sain-van der Velden M, Waterham H, et al. The newborn screening paradox: sensitivity vs. overdiagnosis in VLCAD deficiency. JIMD Rep. 2016;27:101–6. https://doi.org/10.1007/8904_2015_476.

18. National Institutes of Health. How many newborns are screened in the United States? 2017. https://www.nichd.nih.gov/health/topics/newborn/conditioninfo/infants-screened. Accessed 11 April 2022.

19. Heron M, Hoyert DL, Murphy SL, Xu J, Kochanek KD, Tejada-Vera B. Deaths: final data for 2006. Natl Vital Stat Rep. 2009;57(14):1–134.

20. Martin JA, Hamilton BE, Sutton PD, et al. Births: final data for 2006. Natl Vital Stat Rep. 2011;57(7):1–104.

21. Merritt JL, 2nd, Vedal S, Abdenur JE, et al. Infants suspected to have very-long chain acyl-CoA dehydrogenase deficiency from newborn screening. Mol Genet Metab. 2014;111(4):484–92. https://doi.org/10.1016/j.ymgme.2014.01.009.

22. National Newborn Screening & Genetics Resource Center (NNSGRC). National newborn screening 2006 incidence report. 2006. https://genes-r-us.uthscsa.edu/sites/genes-r-us/files/resources/genetics/2006datareport.pdf. Accessed 11 April 2022.

23. German Society for Newborn Screening (DGNS). National screening reports. https://www.screening-dgns.de/reports.php. Accessed 11 April 2022.

24. Sander J, Sander S, Steuerwald U, et al. Neonatal screening for defects of the mitochondrial trifunctional protein. Mol Genet Metab. 2005;85(2):108–14. https://doi.org/10.1016/j.ymgme.2005.02.002.

25. United Nations Population Division. World population prospects 2019. 2019. https://population.un.org/wpp/. Accessed 11 April 2022.

26. Lindner M, Hoffmann GF, Matern D. Newborn screening for disorders of fatty-acid oxidation: experience and recommendations from an expert meeting. J Inherit Metab Dis. 2010;33(5):521–6. https://doi.org/10.1007/s10545-010-9076-8.

27. Therrell BL, Padilla CD, Loeber JG, et al. Current status of newborn screening worldwide: 2015. Semin Perinatol. 2015;39(3):171–87. https://doi.org/10.1053/j.semperi.2015.03.002.

28. Mittal RD. Tandem mass spectroscopy in diagnosis and clinical research. Indian J Clin Biochem. 2015;30(2):121–3. https://doi.org/10.1007/s12291-015-0498-9.

29. Spiekerkoetter U, Lindner M, Santer R, et al. Management and outcome in 75 individuals with long-

chain fatty acid oxidation defects: results from a workshop. J Inherit Metab Dis. 2009;32(4):488–97. https://doi.org/10.1007/s10545-009-1125-9.

30. Boneh A, Andresen BS, Gregersen N, et al. VLCAD deficiency: pitfalls in newborn screening and confirmation of diagnosis by mutation analysis. Mol Genet Metab. 2006;88(2):166–70. https://doi.org/10.1016/j.ymgme.2005.12.012.

31. Hencher Lee H-C, Mak CM, Lam C-W, et al. Analysis of inborn errors of metabolism: disease spectrum for expanded newborn screening in Hong Kong. Chin Med J (Engl). 2011;124(7):983–9. https://doi.org/10.3760/cma.j.issn.0366-6999.2011.07.006.

32. Joost K, Ounap K, Zordania R, et al. Prevalence of long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency in Estonia. JIMD Rep. 2012;2:79–85. https://doi.org/10.1007/8904_2011_51.

33. Chien Y-H, Lee N-C, Chao M-C, et al. Fatty acid oxidation disorders in a Chinese population in Taiwan. JIMD Rep. 2013;11:165–72. https://doi.org/10.1007/8904_2013_236.

34. Obaid A, Nashabat M, Alfadhel M, et al. Clinical, biochemical, and molecular features in 37 Saudi patients with very long chain acyl CoA dehydrogenase deficiency. JIMD Rep. 2018;40:47–53. https://doi.org/10.1007/8904_2017_58.

35. Bleeker JC, Kok IL, Ferdinandusse S, et al. Impact of newborn screening for very-long-chain acyl-CoA dehydrogenase deficiency on genetic, enzymatic, and clinical outcomes. J Inherit Metab Dis. 2019;42(3):414–23. https://doi.org/10.1002/jimd.12075.