Fabry or not Fabry: From genetics to diagnosis
van der Tol, L.

Citation for published version (APA):
van der Tol, L. (2015). Fabry or not Fabry: From genetics to diagnosis

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Download date: 30 Dec 2018
CHAPTER 10

Summary and general discussion
SUMMARY
The aim of this thesis was to improve the diagnosis of Fabry disease (FD), in order to support the diagnostic process in individuals with a genetic variant of unknown significance (GVUS) in the α-Galactosidase A (GLA) gene.

The prevalence of FD has long been reported to be around 1:40,000 males \(^1\). However, it is increasingly suggested that the prevalence of FD may be underestimated, especially as a result of screening studies that assess the prevalence of FD in different subgroups. In chapter 2 we have reviewed the literature on screening for FD. We found that while two screening studies were performed by 2001, following approval of enzyme replacement therapy (ERT) for FD by the European Medicines Agency (Europe) and the Food and Drug Administration (USA), subsequently 49 screening studies were performed in 2002 - 2012. Furthermore, 45% of these studies were funded by at least one of the companies that market ERT, suggesting that besides increased awareness, due to the availability of ERT, the involved pharmaceutical companies may have had an active role in promoting such studies. Screening in newborns suggested an overall prevalence of 1:2500 (0.04%), much higher than based on clinical ascertainment. In high-risk screening cohorts, that is in subgroups with a single non-specific symptom such as left ventricular hypertrophy (LVH), chronic kidney disease (CKD) or stroke, an even higher prevalence of 1:161 (0.62%) was found. However, we noted significant differences between studies.

Subsequently, we applied strict clinical and biochemical criteria to the individuals detected with a GLA gene variant in the high risk screening cohorts. As such, we aimed to identify those with a definite diagnosis and a classical FD phenotype. When applying these criteria, only 0.12% (rather than 0.62%) had a definite diagnosis of FD. It is important to note that in a significant number of individuals, too little information was given for classification. Despite this limitation, it can still be concluded that screening for FD will identify a substantial number of individuals with a GLA variant, who do not demonstrate characteristic features of classical FD. In these individuals, the GLA variant is considered a GVUS. These individuals may either have a non-classical phenotype, or the GLA variant may be neutral. In the latter cases there might be another underlying cause for the symptoms that led to the inclusion of to explain the initial reason to screen that individual in the screening study (e.g. a sarcomeric mutation or severe hypertension that causes LVH).

In order to improve the diagnosis of FD, we first had to establish the criteria that define this disease. We collaborated with seven experts in a Delphi consensus procedure (chapter 3) to develop criteria to select those individuals in whom there is no doubt about the diagnosis. These criteria were made with an emphasis on avoiding false positives. The criteria (chapter 3, table 2) comprised of the presence of a GLA variant and a very low α-Galactosidase A (aGalA) enzyme activity for males; combined with a minimum of one of the characteristic features of classical FD: FD neuropathic pain, angiookeratoma, cornea verticillata or a globotriaosylsphingosine (lysoGb3) or globotriaosylceramide (Gb3) in the range of classical male FD patients; or a family member with a definite diagnosis of FD carrying the same GLA variant. Individuals, who do not fulfil these criteria, have an uncertain diagnosis of FD. Furthermore, the panel endorsed that strict definitions apply to angiokeratoma, FD neuropathic pain and cornea verticillata (CV). Therefore,
the committee emphasized that assessment of angiokeratoma, Fabry neuropathic pain and CV should be performed by a physician with expertise on FD, in order to avoid an incorrect classification (e.g., atypical pain in hand or feet unjustly labeled as FD pain).

The panel furthermore discussed that in case of an uncertain diagnosis, histological evidence for lipid storage in an affected organ should be pursued. The panel agreed that the demonstration of storage characteristic of FD in an affected organ (e.g., heart, kidney, aside from skin) by electron microscopy analysis, serves as the gold standard for FD. Concentric multi lamellated myelin bodies with a zebra like pattern (zebra bodies) and with a periodicity of approximately 5 nm are characteristic of FD, but it remains important that the biopsy is assessed by an experienced pathology team. Lysosomal storage similar to that seen in FD may also be found in other lysosomal storage disorders, however, the storage is considered specific in the context of clinical signs or symptoms compatible with FD. Furthermore, certain medications may cause similar storage, therefore a careful medical history for the use of these medications is indispensable to interpret biopsy results. The criteria for a definite diagnosis and the gold standard were subsequently endorsed and adopted by the experts who participated in the studies presented in chapters 4, 5 and 6.

In chapter 3, 4 and 5, we reviewed the literature to identify clinical and laboratory criteria that may be applied to confirm (entry criterion) or exclude (exit criterion) FD in individuals with LVH, CKD or cerebrovascular disease (transient ischemic attack (TIA), stroke or white matter lesions (WMLs)). The literature search revealed several features that were found to be abnormal in FD patients. However, in order to serve as an entry criterion, a high specificity is warranted. Most often, a relevant control group was not assessed, and specificity remained unclear. This may have influenced the results substantially, and several criteria that were identified may still be valid, despite insufficient data on specificity and sensitivity. In order to overcome this limitation, the experts in the Delphi procedures collaborated on additional analyses in FD cohorts and control groups relevant to assess specificity and to add to the currently available data. Subsequently, criteria were assessed in a modified Delphi consensus procedure and a diagnostic algorithm was constructed for each affected organ system.

For individuals with LVH (chapter 3), the panel agreed that, based on the additional analyses, LVH of >15mm in an individual aged <20 years and microvoltages on electrocardiography can both exclude FD. Although data on specificity were limited, the panel considered the presence of a PQ interval < 120 ms, sinus bradycardia, hypertrophied papillary muscle, myocardial late enhancement in the infero-postero-lateral region as ‘red flags’. This indicates that the diagnosis is likely, but further investigations are needed to establish a definite diagnosis. For individuals with CKD (chapter 4), there were few studies in which specificity was assessed and no criteria to confirm or exclude FD nephropathy could be determined. However high urine Gb3 and the maltese cross sign in urine serve as a ‘red flags’. For both individuals with LVH and CKD, and an uncertain diagnosis of FD, the panels fully agreed that a heart or kidney biopsy, respectively, should be performed to confirm or exclude FD. In chapter 5 we described the results for individuals who present with TIA, stroke or WMLs. Additional analyses showed that a pattern of
WMLs, characteristic for cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL or the recessive variant CARASIL) is not seen in FD patients and can thus exclude FD. The pulvinar sign on T1 weighted magnetic resonance imaging (MRI) and increased basilar artery diameter (BAD), assessed with MRI were promising. Additional analyses showed that an increased BAD may possibly identify FD patients, when compared to patients with non-Fabry stroke. The panel agreed that the pulvinar sign and BAD > 4.2 mm can each serve as a ‘red flag’.

In chapter 6 we described the results of a consensus study with the aim to improve the diagnosis of individuals with pain or a small fiber neuropathy (SFN), angiokeratoma or CV, in whom a GLA variant was identified, and in whom there is no heart, brain or kidney involvement. The experts agreed that if the clinical feature meets strict criteria, there is no alternative diagnosis.

Especially in individuals who present with TIA, stroke or WMLs, a definite conclusion cannot always be drawn with the developed algorithm. Also, for individuals with LVH or CKD, a biopsy may not always be feasible. The diagnostic options should be considered on an individual basis and follow-up in a center with expertise on FD is advised for those individuals in whom the diagnosis remains unclear.

As a characteristic feature of classical FD, CV assessment was hypothesized to be an accurate diagnostic tool for FD. In chapter 7, we systematically studied the literature on the prevalence of CV in FD, and found an overall prevalence of 69% (range 36-96), much lower than expected based on clinical experience. The lower prevalence may be accounted for by the inclusion of individuals with a non-classical phenotype or even without FD in the studied cohorts. Only in a subset of studies, data were sufficient to identify patients with a non-classical phenotype or uncertain diagnosis. In order to generate a more accurate estimate of the prevalence, we studied the Dutch FD cohort. Strict criteria (see chapter 3) were applied to identify those with a classical phenotype. The results of organ biopsies were used to identify individuals with a non-classical phenotype and those with a non-pathogenic GLA variant. These data revealed a much higher prevalence in the group of patients with classical FD: 82% (95% confidence interval (95% CI) 71-89) for females, and 94% (95% CI 80-99) for males. In the individuals with a non-classical phenotype, only 16% (95% CI 6-35) had CV, and there were no false positives. The results indicate that the presence of CV could confirm FD. However, with a low prevalence in the non-classical group, being patients who originally presented with an uncertain diagnosis, CV assessment is helpful in only a minority of individuals.

Like CV, FD neuropathic pain is a characteristic feature of FD. The characteristic pain in FD is related to a small fiber neuropathy. In order to assess the diagnostic applicability of quantitative sensory testing (QST) and intraepidermal nerve fiber density (IENFD), we assessed individuals with a non-classical phenotype and individuals with a non-pathogenic variant as described above (chapter 8). In the group of patients with non-classical FD (n=18, 9 males), 29% had one or more abnormal QST modality, while only 2 individuals exhibited an abnormal cold detection threshold (CDT). Most individuals (83%) had an abnormal IENFD. Interestingly, also individuals
with a non-pathogenic variant (n=5, 3 males), showed QST (20%) or IENFD (75%, 1 missing data) abnormalities. For all except 1 individual with a non-classical phenotype, the abnormalities were limited and not sufficient to conclude that a SFN was present. The sensitivity of a combined abnormal QST and IENFD was 28%, with a specificity of 80%. These results indicate that accidental findings of abnormal QST or IENFD may occur in individuals without FD. The cause for the abnormalities in this group remains unclear. We concluded that, in our study, QST and IENFD assessments cannot confirm or exclude FD in individuals with an uncertain diagnosis.

LysoGb3, as a hallmark of FD, may possibly differentiate individuals with a non-classical FD phenotype, from those with a neutral GLA variant. Therefore, we studied lysoGb3 in the Dutch FD cohort, comprising of patients with a classical and non-classical phenotype and individuals with a non-pathogenic GLA variant, again by using the previously described criteria (chapter 9). The data confirmed that males and females with a classical phenotype invariably have elevated plasma lysoGb3 levels (males, n = 38, ≥ 45 nmol/L, females, n = 66, ≥ 1.5 nmol/L, normal reference ≤ 0.6). Also, males with a non-classical phenotype had increased levels, yet considerably lower than patients with classical FD (n = 13, ≥ 1.3 nmol/L). In females with a non-classical phenotype (n = 14), plasma lysoGb3 levels overlapped with healthy controls. In individuals with a non-pathogenic variant (n = 9, 6 males), levels were within the normal range. Furthermore, we assessed lysoGb3 levels in individuals with a persistent uncertain diagnosis, revealing slight elevations up to 1.6 nmol/L. We extensively discussed if a lysoGb3 value ≥ 1.3 nmol/L may be used to confirm FD. This decision was complicated by the small patient numbers and the elevated levels in the “uncertain” group. With these limitations in mind, we concluded that an increased plasma lysoGb3 level supports a diagnosis of FD, but further assessments should be performed to confirm or exclude FD in uncertain cases.

GENERAL DISCUSSION

What have we learnt from screening studies?
Historically, Fabry disease (FD) was reported to have a birth prevalence of approximately 1:40,000 males. Estimations of prevalence, however, differ significantly between studies. For example, calculation of the birth prevalence of FD in the Netherlands was, corrected for the number of births, 1:238,000 males, and 1:476,000 for males and females combined, while Meikle et al reported a birth prevalence of 1:117,000 for Australia in approximately the same study period (males and females combined). These calculations were based on enzymatic diagnoses in, reflecting (mostly) male diagnoses. Before 1999, identification of FD was based on individual case finding. As a consequence, the number of identified cases depended heavily on recognition of the, usually, classical phenotype and family studies. At that time, screening of larger groups had only been performed once by Nakao et al in a cohort of individuals with left ventricular hypertrophy (LVH).

This changed tremendously after enzyme replacement therapy (ERT) was approved by the European Medicines Agency and the Food and Drug Administration in 2001. Nearly 50 studies...
were performed in the 10 years that followed, almost half of these funded or supported by the companies that market ERT. These studies suggested that the prevalence of FD may be much higher than previously thought. Indeed, screening in newborns, by determination of low levels of αGalA enzyme activity, showed an overall prevalence of 1:2500 (0.04%). In adult populations, screening was performed in so-called “high-risk” groups, e.g., individuals with a single non-specific symptom such as LVH, chronic kidney disease (CKD) or stroke. It was suggested that individuals in these subgroups could have unrecognized FD, missing a window of opportunity of adequate treatment. In line with the expectations, the prevalence in these high-risk groups was even higher (combined prevalence of 1:161, 0.62%). There are some major limitations to these studies. Firstly, these studies mainly identified males, as enzymatic analysis was most often used to screen for FD, potentially missing approximately 40% of females \(^\text{11}\). But, secondly, and much more important, it would be incorrect to conclude that this is the true prevalence of FD for each group. For both newborns and the high-risk individuals, these results merely reflect the prevalence of α-Galactosidase A (GLA) gene variants in the different groups. Harboring a GLA variant does not necessarily imply that this individual has or will develop clinical FD. In several studies, individuals with any GLA variant were deemed FD patients and sometimes even treated with ERT. Often, a thorough phenotypic evaluation of these individuals was not performed to confirm the diagnosis with characteristic clinical, biochemical or histological features. Others recognized that these individuals stand out as compared to the clinical pattern of FD seen in classically affected (male) patients. Frequently, these individuals were reported as ‘late onset’, ‘atypical’ patients, or in the presence of solely LVH or CKD as having a cardiac or renal variant of FD. As described in chapter 2, structural categorization of the individuals found by screening, revealed that only 1 out of 6 individuals with a GLA variant had characteristic features of FD and a definite diagnosis according to predefined criteria.

The results of the literature review illustrate that careful considerations should follow before a definite diagnosis can be made when a GLA gene variant is identified. The diagnosis of those who were found to have a GLA variant is often unclear, with a number of individuals not suffering from FD. These individuals have an uncertain diagnosis with a genetic variant of unknown significance (GVUS) in the GLA gene. Over the last decade, several of these patients have been misdiagnosed and have received inappropriate counselling and treatment. This creates an obligation to our patients to critically evaluate all modalities that can aid in making a proper diagnosis. When we address this problem, the first question is how Fabry disease is actually defined.

**What defines Fabry disease?**

Since the first descriptions of FD patients, several clinical features have been associated with what is now referred to as classical FD. The skin lesions, angiokeratoma, were the basis for the initial recognition of FD in the seminal papers of Fabry and Anderson \(^\text{12,13}\). Angiokeratoma in FD are predominantly present in a clustered form, and are located predominantly in the umbilical, perioral and bathing trunk area, but may also cover larger areas such as the trunk and limbs. In addition, patients often report burning pains in the limbs that are exacerbated by heat, fever or exercise. Cornea verticillata (CV) is frequently present and furthermore, patients can have an- or hypohidrosis.
Pathogenic variants in the GLA gene lead to absent or near absent α-Galactosidase A (αGalA) enzyme activity in males with classical FD, while in females, enzyme activity may be normal or slightly reduced \(^{11}\). As a result, globotriaosylceramide (Gb3) accumulates in lysosomes of several tissues and is elevated when measured in plasma or urine in males and some females \(^{14}\). Furthermore, Gb3 accumulation can be demonstrated with histology, and is best assessed with electron microscopy (EM), where concentric multi lamellated myelin bodies with a zebra like pattern (zebra bodies) and with a periodicity of approximately 5 nm are characteristic of FD \(^2\). More recently, globotriaosylsphingosine (lysoGb3) has been identified as an important hallmark of FD \(^{16}\), and can be measured in plasma and urine \(^{16-18}\). The combination of these characteristic clinical and biochemical features reflects patients with an unmistakable classical FD phenotype.

However, problems arise when we are confronted with patients who do not present the clinical and biochemical phenotype that is associated with classical FD. Historically it was already known that there are individuals with GLA variants with a less well defined phenotype. For example, individuals with the IVS4+919G>A, p.M296I, p.N215S or p.R112H variant generally do not exert characteristic features of classical FD \(^{10,19-22}\). However, in those cases, FD was confirmed with histological evaluation of organ biopsies that demonstrated storage characteristic of FD on electron microscopy (EM) \(^{10,19-21,23,24}\). In contrast, in two reports that contain detailed pathological examination by electron microscopy (EM) of some individuals with the p.A143T variant, no characteristic lysosomal inclusions could demonstrated \(^{21,25}\). Hence, individuals with the p.A143T variant are currently thought not to have FD.

Information on the GLA gene provides important information when FD is suspected. Variants leading to absence of the gene product and no residual αGalA enzyme activity in males, generally true for deletions or insertions leading to frame shifts and early truncating (nonsense) variants, are invariably associated with a classical FD phenotype in males. However, the majority are missense variants, some of which are related to classical FD, while others may cause a more attenuated phenotype or no FD at all. Furthermore, many intronic and splice sites variants have been associated with different phenotypes. The prediction of pathogenicity, solely based on the genetic variant can be difficult, and due to the number of different GLA variants and other influences, a clear phenotype-genotype correlation has not been established.

Several methods are available to further assess the pathogenicity of genetic variants. Prediction models study the structure or function of the gene product and the evolutionary conservation of the affected amino acid. So far, these prediction models only achieve an accuracy of approximately 80%. Recently, Riera et al developed a prediction model specifically for the GLA gene, reaching an accuracy of approximately 90% \(^{26}\). An important limitation is that data from online databased were used to define the pathogenicity of the variants used for validation. These databases do not provide a systematic and robust strategy to designate the pathogenicity of the variants, which is essential for a reliable classification (see section on generalizability of individual findings). In order to define pathogenicity, data from the literature are used, in which a strict definitions to objectify FD, are often lacking. Thus, when strictly defined phenotyping has not been performed, these association studies are not useful. Our current understanding of FD phenotypes should be
implemented into these models to avoid mistakes. An example is the p.D313Y variant: initially, it was believed that this variant was associated with FD, as the variant was identified in a patient with a classical FD. However, additional studies revealed that this variant did not influence the protein structure significantly and that the FD phenotype could be attributed to a second GLA variant. Furthermore, the enzyme activity was related to the pH, resulting in a lower activity at neutral pH as compared to a higher activity when measured at acidic lysosomal pH. The studies on the p.D313Y variant have clearly shown the limitations of enzyme activity measurement: in vitro measurements do not always reflect the in vivo situation, and a deficient enzyme activity in vitro may even be an ‘pseudodeficiency’. It is also important to realize, that a decreased αGalA enzyme activity in vivo, but with a certain amount of residual activity, may be sufficient to prevent clinical disease. While the clinical consequences of residual enzyme activity are often unclear, an absent or nearly absent enzyme activity in males is generally related to a classical FD phenotype.

Further laboratory studies may elucidate the impact of a certain variant. For example, expression studies, often performed in COS cell models, are frequently used to assess the effects of a certain genetic variant on the gene product. Expression of the GLA gene in such a model provides an opportunity to reliably assess residual αGalA enzyme activity, which is not affected by differences between individual patients. However, the benefit for the diagnostic process is limited: while the residual activity is reliably assessed, this does not necessarily predict the clinical outcome. It could be of interest to correlate the clinical phenotype (using predefined criteria to confirm or exclude FD) to the extent of residual enzyme activity in expression studies for different GLA variants. If a correlation is demonstrated, it may be possible to generate a cutoff value for enzyme activity to predict pathogenicity. Besides the enzyme activity, cell models also provide an opportunity to localize the protein, e.g. to study if the protein is correctly localized to the lysosome. However, we should always keep in mind that individual differences occur as a result of genetic, epigenetic and environmental factors. Thus, expression studies may add to the diagnostic process, but cannot be used to predict the phenotype.

Besides information on the GLA gene and αGalA enzyme activity, the in vivo biochemical consequences of a variant are of importance, since storage of Gb3 is the main culprit leading to disease. Substantially increased biomarkers in plasma or urine (e.g. Gb3, lysoGb3) are associated with a classical phenotype as mentioned previously. Because of the importance of biochemical “read-out” of a variant in relation to disease phenotypes, this is further discussed below.

As much relies on thorough phenotyping of individuals, it is of importance to critically evaluate the presence of highly specific imaging, or other characteristics, that can be of help to define FD. Many detailed studies on the affected organs in FD (i.e. heart, kidney, brain) have been performed. For example, a short PQ interval on electrocardiography (EKG) may be seen in FD, possibly caused by involvement of the conductive system of the heart. WMLs may have a distinct pattern, and magnetic resonance imaging (MRI) can reveal increased diameters of the large cerebral arteries. These abnormalities may also be helpful for the differential diagnosis. For a systematic approach to evaluate outcomes of these studies against phenotypes
of FD, the first step that should be taken is to define a definite diagnosis of FD and at the same time delineate criteria for an uncertain diagnosis. For this purpose, a group of international experts discussed the features associated with FD (chapter 3, table 2). The criteria combine the presence of a GLA variant with characteristic clinical and biochemical features to confirm FD. In order to avoid false positives, the definitions were made quite strict, and will most likely identify individuals with a classical FD phenotype only. Individuals who do not fit these criteria, have an uncertain diagnosis of FD. In addition to the establishment of these criteria, the panel agreed that the gold standard for FD is histology, assessed with electron microscopy as describe above. The use of these criteria in the selection of individuals for future research projects will lead to more homogeneous groups and allows for comparison of data between studies.

These criteria have provided much needed clarity and uniformity. It could be discussed, however, if individuals fulfilling these criteria should always be regarded as diseased in the clinical sense. For instance, an individual with a GLA variant, nearly absent aGalA enzyme activity and a very high lysoGb3, will fulfill the criteria for a definite diagnosis, disregarding absence of clinical signs or symptoms such as pain, CKD or LVH. With this profile, it is most likely that a male patient will develop clinical apparent FD in the future. But, similar difficulties also arise for asymptomatic individuals and in particular females, with an affected family member. It is difficult to predict if and when asymptomatic family members will develop signs and symptoms of their own and several factors may be involved. This is an important consideration to take into account when counselling a family.

We recommend to apply these criteria to confirm the diagnosis in individuals with (suspected) FD, while for (yet) asymptomatic family members, caution should be taken before an individual is designated as having FD. In other words, these criteria do not predict the clinical course for each individual patient. In some cases it may not be feasible to confirm or exclude FD in the index patient. Evaluations of (male) family members may be helpful, but it should be discussed that there is no guarantee that this will solve the diagnostic problem, and may add another family member with a GVUS and thus an uncertain diagnosis of FD. Counseling of the family is important to avoid misunderstanding and regrets of decisions made.

Uncertain diagnosis of FD: diagnostic approach

Diagnostic algorithms were developed for individuals who do not have a characteristic phenotype, but in whom a GLA variant is identified (uncertain diagnosis). The algorithms in chapter 3, 4, 5 and 6 were generated using expert panels. Although expert opinion is generally considered low level of evidence, this is not the case for the algorithms developed here. The literature was systematically reviewed in order to obtain a complete overview of the available data. The expert decisions were thus based on the currently best available evidence, and not solely on their clinical experience. Furthermore, before a certain criterion was accepted, a high level of agreement was required (≥75% agreement, no one disagreed). With this approach and with the collaboration of a total of 32 international experts, we believe that these algorithms can be used in clinical practice. However, an open mind to developments that may reveal new insights is an absolute requirement. Along with the discussions, we have identified several caveats that need further
assessment and will help in improving these algorithms further in the future. For example, the literature reviews and diagnostic algorithms clearly demonstrate that the currently available data are not sufficient to confirm or exclude FD in all individuals with an uncertain diagnosis. Several criteria were selected that could exclude FD, i.e. the specific sign or symptom is not found in FD patients. But, the currently best available option to confirm FD is EM assessment of a biopsy specimen from an affected organ. For instance, in individuals who present with transient ischemic attacks (TIA), stroke or white matter lesions (WMLs), a biopsy of the affected organ is not an option, and there were no other criteria to confirm FD. And even for individuals with LVH or CKD, a biopsy may not always be feasible.

To improve the diagnostic approach, we have studied several possible diagnostic criteria in detail. Cornea verticillata (CV) has a high prevalence in classically affected patients \(^1\),\(^3\),\(^9\),\(^40\), and importantly, presence of CV is specific for FD (in the absence of the use of medications that may cause CV), i.e. there is no alternative diagnosis \(^6\). With a low prevalence in the non-classical group, being patients who originally presented with an uncertain diagnosis, CV assessment will only be helpful in a minority of individuals with an uncertain diagnosis. Nevertheless, CV assessment is essential in the initial evaluation of each individual with a GLA variant, because if found, and especially in combination with other characteristic features of FD, this information is extremely useful. Another feature that is believed to be characteristic of FD is neuropathic pain. Several studies have shown that there is a relation with affected small nerve fibers \(^41\)-\(^48\). Data from a predominantly classical FD cohort studied by Biegstraaten et al \(^41\),\(^42\) predict that 100% of males and 75% of females can be detected by assessment of temperature perception and IENFD. We found that abnormal QST and IENFD findings were also present in a small cohort of patients with non-classical FD, but the abnormalities were not overwhelming, and false positives were also found (sensitivity 27%, specificity 83%, control group: individuals with a neutral GLA variant). We concluded that small nerve fiber assessment cannot reliably confirm or exclude FD in uncertain cases. However, in the presence of pain with childhood onset, exacerbated with heat of fever and with abnormal QST and IENFD assessments, there is no doubt about the diagnosis of FD.

As mentioned previously, the biochemical characteristics need further discussion as tools to help define FD. The biochemical marker lysoGb3 was identified as a hallmark of FD by Aerts et al in 2008 \(^15\). Subsequent studies have shown that plasma lysoGb3 can reliably detect patients with a classical FD phenotype, in contrast to the primary storage product in FD (Gb3) \(^16\),\(^18\),\(^49\)-\(^51\). Furthermore, clinical studies showed a (limited) correlation of plasma lysoGb3 levels with certain clinical features, disease severity \(^49\) and a response to treatment with ERT was observed \(^52\). These data suggest that plasma lysoGb3 may also be of use for the diagnosis of FD for individuals with an uncertain diagnosis. Our study on plasma lysoGb3 in patients with non-classical FD and in individuals with a non-pathogenic GLA variant, showed that all males with a non-classical phenotype had an elevated lysoGb3 level (≥1.3 nmol/L, normal reference ≤ 0.6), while the values in females overlapped with healthy controls. We extensively discussed if a lysoGb3 value ≥1.3 nmol/L may be used to confirm FD. The small subject numbers, and an elevated level of 1.6 nmol/L in an individual with LVH and with an uncertain diagnosis in whom an organ biopsy was not an option, complicated this decision. A decision to implement lysoGb3 as a confirmative
criterion would also imply that other assessments, such as a biopsy of the affected organ for the purpose of diagnosis, may no longer be indicated in an individual with an increased lysoGb3 in plasma. With these limitations in mind, we concluded that a plasma lysoGb3 ≥1.3 nmol/L is very suggestive of FD, but further assessments should be performed to confirm or exclude FD in uncertain cases. Before we can use lysoGb3 in clinical practice to confirm or exclude FD, additional studies should first confirm our data in larger groups, and should furthermore focus on the specificity of lysoGb3, explicitly regarding the use of medication that may induce a lipidosis. Additionally, the assessment of lysoGb3 is not yet available for all centers and different mass spectrometry assays are in use to quantify plasma lysoGb3, that employ different internal standards. In contrast to the assay used for the studies described in this thesis, some standards show little chemical resemblance to lysoGb3. These differences make it difficult to predict if all assays can reliably quantify slight elevations. Efforts should therefore be made to improve the availability of reliable lysoGb3 measurements before plasma lysoGb3 can replace Gb3 in the diagnostic approach and treatment follow-up.

Besides lysoGb3 itself, several isoforms of lysoGb3 have been identified by the group of C. Auray-Blais. These isoforms have different sphingosine base moieties, and are found in plasma and urine of FD patients. Interestingly, while lysoGb3 is present in urine only in very small amounts, some isoforms are present more abundantly. The profile of lysoGb3 isoforms in individuals with a non-classical FD phenotype is yet unknown. Information on lysoGb3 isoforms may improve our understanding of the different clinical findings in the classical and non-classical patients (e.g. classical versus non-classical or cardiac versus renal involvement). In order to study this, the lysoGb3 isoforms should be measured in treatment naïve patients with a non-classical FD phenotype and compared to a group of age and gender matched patients with a classical FD phenotype, applying the strict definitions to categorize individuals that we applied earlier.

However, measurements in plasma and urine will not demonstrate the origin of lysoGb3 and its isoforms. It would be ideal to assess the different tissues in FD patients with a classical and non-classical phenotype, however, acquiring (sufficient) tissue is not always feasible. Alternatively, tissue could be acquired with obduction. A major limitation is the small number of individuals with a well-defined non-classical phenotype, and the availability of post mortem tissue samples in only a subset of these individuals. To study the lipid profile in different organs, an animal model would be of interest. A knockout mouse model of FD is available and lysoGb3 levels in several tissues of this mouse model have been measured, which gives some indication about possible sources of plasma lysoGb3. However, this model does not demonstrate clear clinical features of FD (i.e. LVH, CKD or stroke), although in a single study, mild LVH was found. Taguchi et al recognized this problem and developed an improved mouse model by combining a GLA knockout with induction of the Gb3 synthase gene. This mouse model showed a level of microalbuminuria, and some systolic and diastolic dysfunctions, but there was no significant increase of cardiac mass. LysoGb3 tissue levels were not reported. In order to study the lipid profiles in affected organs, further efforts should be made to improve an animal model, also considering other animals who may express a classical phenotype with the introduction of a GLA knockout. This would give an opportunity to also study the phenotype of animals with different
GLA variants, to mimic non-classical FD. Additionally, induced pluripotent stem cells, to generate organ specific cell lines such as cardiomyocytes from either human or animal origin could provide interesting possibilities to study lipid accumulation in relation to cell types \(^{61,62}\).

With the currently available data, lysoGb3 isoforms in plasma or urine are subject of important research questions, but have no place (yet) in the clinical setting of diagnosis and follow-up.

**Diagnostic accuracy**

In order to validate the diagnostic accuracy of the algorithms and criteria presented in this thesis, a prospective study would be needed, using uniform in- and exclusion criteria and a standardized approach \(^{63}\). For this purpose, a group of individuals presenting with LVH, CKD, TIA, stroke or WMLs, or CV, SFN or angiokeratoma and a GVUS in the GLA gene should be assessed for the discussed criteria, as well as the gold standard, a biopsy of an affected organ. However, for individuals presenting with TIA, stroke or WMLs, and with CV, SFN/pain or angiokeratoma only, such a study is hampered by impossibility to obtain histological confirmation as a biopsy of the affected organ is not possible.

A biopsy of an affected organ has been designated by the experts as the gold standard for FD. There are, however, several limitations. Firstly, the clinical context is important to consider as similar storage has been described in other lysosomal storage disorders \(^{3-5}\). However, these diseases have no phenotypical overlap, and in the clinical context of signs and symptoms possibly related to FD, there is no discussion about the specificity. Furthermore, it is important to ask patients for the use of lipidosis inducing medication, which can cause storage indistinguishable from FD. Secondly, lamellated structures can be found in various other situations, such as in autophagy/degenerative diseases and rheumatoid arthritis. When there is uncertainty with respect to the clinical phenotype, the characteristic periodicity of the zebra bodies on EM assessment, and in particular the periodicity of the lamellated bodies, can discriminate between FD zebra bodies and other myelin-like bodies. Since these differences can be subtle, this implies that experienced pathology teams need to perform these analyses and that they should be well informed about the clinical context. As an illustration, recently two families (5 patients) have been described by Apelland et al \(^{64}\). Patients presented with early onset severe HCM, including an abnormal fat distribution in 3 patients. Cardiac and kidney biopsies showed lysosomal inclusions on EM with the Fabry related lamellated pattern. However, in the kidney biopsy, podocytes were not affected, which is very unusual for FD. Gb3 was slightly increased in the plasma of 1 patient, while Gb3 in tissue from the affected heart was high in all patients. αGalA leukocyte enzyme activity was slightly reduced in 1 female, but molecular analysis of the GLA gene revealed no variants. It was concluded that the patients have a novel familial Gb3-associated cardiomyopathy, although it is uncertain whether there was accumulation of other lipids. No detailed information about the periodicity of the inclusions was provided. It is quite possible that small elevations of Gb3 are not specific for Fabry disease and can be found in other disorders. This was also illustrated by Schiffmann et al, who demonstrated small increases of urine and tissue Gb3 in patients with non-Fabry heart diseases \(^{65}\). These cases demonstrate that in some cases atypical storage of
Gb3 can occur, which is not necessarily associated with FD. These and upcoming studies also illustrate the need for further refinement of the diagnostic algorithms in the future.

We have now repeatedly argued that a diagnostic accuracy study should abide to a strict protocol to generate homogeneous, comparable, results in all included cases. This protocol should include a structured categorization of individuals, for which the criteria presented in this thesis are very suitable. Because FD is rare, such a study can only succeed with cooperation between (international) expert centers for FD to achieve sufficient power. It would also be an opportunity to study the criteria that were deemed promising, but not yet sufficiently studied (‘red flags’) in patients with a non-classical FD phenotype, and a control group with a non-pathogenic GLA variant. This could subsequently result in important information on specificity. At first, the practical implications of this proposed study seem immense. However, most criteria may be obtained with the standard assessments that are already performed, such as cardiac ultrasound, MRI, and blood and urine laboratory measurements, albeit in a uniform setting for all participating centers.

Equally important, separate studies should focus on specificity of criteria that may confirm FD, by studying these in relevant control groups. LysoGb3 should therefore be assessed in cohorts using certain medications that may induce lipidosis (most importantly amiodarone and chloroquine). Cardiac related criteria on EKG, ultrasound or MRI should be assessed in populations with other causes of LVH/HCM. Correspondingly, criteria related to CKD or TIA, stroke or WMILs should be assessed in groups with the same clinical sign or symptoms. From a methodological point of view, it would be preferred to exclude FD in the patients who are included in the control group, to avoid false positives. However, this is will inevitably lead to unwanted diagnostic dilemmas if GLA GVUS are identified. Therefore, screening for FD should not necessarily be done in all study patients, but may be confined to those in whom the studied criterion is identified.

**Generalizability of individual findings**

Several authors have reported on the characteristics of individuals with a non-classical FD phenotype or with a non-pathogenic variant. To date, reports are consistent in their findings for certain GLA variants with respect to histology. For example, the p.R112H variant is associated with a non-classical phenotype, and biopsy results from different families confirm the presence of characteristic storage on EM, while there are, to our knowledge, no studies that report histology of individuals without Gb3 accumulation. Conversely, for the p.A143T variant, the absence of storage on histology was reported for different families. However, having a certain GLA variant does not always predict disease severity and disease course. Several factors may influence the pathogenicity of a GLA variant. First, cardiovascular risk factors may influence the development of heart, kidney or cerebrovascular disease. And second, differences in genetic background or epigenetics may play a role in the development of a clinical phenotype. To date, the influence of epigenetics and gene interactions on the penetration of monogenetic disorders is largely unknown, hampering an adequate assessment of genetic variants.

It would be practical to have a list of GLA variants at hand, in which the pathogenicity of each variant is clearly stated. Several initiatives of databases with FD variants already exist, but
different criteria are used to designate if mutations are pathogenic, a GVUS, or non-pathogenic. With the limitations of generalizability in mind, it would be useful to generate a database in which variants are systematically assessed and scored to add to the diagnostic approach. Several attempts to elucidate the pathogenicity of certain variants have already been made. For example, the p.R118C variant is frequently found in screening studies 70-73, and Ferreira et al recently studied the variant in multiple families in Portugal and concluded that the variant does not segregate with the FD phenotypes in a Mendelian fashion, but that it might be a risk factor of cerebrovascular disease 74. This example also illustrates that a definite conclusion is not always straightforward. However, again, with international collaboration, it ought to be possible to identify variants that will invariably cause clinical apparent FD or no FD at all. Such a project requires active participation of multiple centers, and this database should be updated regularly.

Treatment
The Hamlet study focused on the establishment of a correct diagnosis. It is of great importance to identify those with FD to provide counseling and treatment options, while wrongful labeling of individuals with this chronic, progressive disease should be avoided. A correct diagnosis is also the first, indispensable, step towards the optimal management for a patient. While this seems obvious, this is not always easy in clinical practice. As we described in chapter 2, individuals with a GLA variant are frequently considered to be FD patients without further investigations. As a result, these individuals are sometimes treated with ERT. The p.A143T variant was previously discussed: Smid et al and Terryn et al demonstrated that individuals with this variant do not demonstrate characteristic storage in biopsies of the heart and kidney 21,25. Terryn et al also searched one of the main international FD registries (Fabry Outcome Survey, FOS, Shire HGT), and identified 20 individuals (8 males) with this variant. Interestingly, the majority was treated with ERT (n = 15, 7 males), but data on histology was not presented 25. The current evidence is very convincing that this variant does not lead to clinical apparent FD, and as a result, ERT is not indicated. To put the results of the FOS registry in perspective: 5 individuals have been identified with the p.A143T variant in the Netherlands and treatment with ERT would cost an estimated €200,000 per individual per year, or €1,000,000 per year 75. This is a significant amount for a country with a population of 16 million inhabitants, and only reflects 1 GLA variant, thus underestimating the costs of potential unnecessary treatment significantly. We strongly advocate that treatment with ERT is only considered for those with a definite diagnosis of FD, because initiation of ERT has profound consequences for the individual (invasive treatment, expectations about the future), and society (high costs of ERT) 76-77.

While treatment should obviously be avoided for individuals who do not have FD, there is considerable doubt about the effectiveness of ERT for patients with a non-classical phenotype. These patients, assuming a definite diagnosis of FD is made, do exhibit accumulation of Gb3 in the affected organ 19-21,78. Therefore, it could be argued, that ERT, with the rationale of reduction of Gb3 accumulation, is indicated in these patients. A recent meta-analysis emphasized that ERT may result in (limited) stabilization in subgroups, but complications may still arise despite ERT 79. These studies, like most research projects on FD, were conducted in cohorts of predominantly classically affected patients. It may be that the limited efficacy is partially caused by severe
disease at initiation of ERT. Furthermore, antibody formation against the infused enzyme often occurs in males with classical FD, who have absent or near absent αGalA enzyme activity. Antibodies reduce the effects of ERT on urine and plasma (lyso)Gb3, and are likely to reduce the clinical benefit as well. Following this line of argument, a non-classical patient, with attenuated disease and residual αGalA enzyme activity, may be hypothesized to be the ideal candidate for ERT. Conversely, individuals with a non-classical phenotype will often present later in life. It is questionable if ERT will improve the disease progression substantially in order to benefit from the treatment in terms of life span and quality of life.

Evidence for effectiveness of ERT in non-classical patients is scarce, and the best evidence probably comes from females, who often present with solitary involvement of the heart. In females, the cardiac mass decreased or remained stable with ERT. Besides studies on females, only one study reports on the efficacy of ERT in patients with a non-classical phenotype. This study focusses on one particular GLA variant (IVS4+919G>A / c.936+919G>A) resulting in a cardiac phenotype and reported improvement or stabilization of cardiac status and renal function, improved micro-albuminuria, and a reduction of lysoGb3. However, a later report by the same group reveals an increase of lysoGb3 after the initial decrease. These data suggest a possible beneficial effect for patients with attenuated FD.

The considerations regarding the application of ERT for non-classical FD patients are even more complex for asymptomatic family members with the same GLA variant. There may be a window of opportunity to initiate treatment before extensive organ damage occurs. In the literature, several studies endorse early treatment, with the assumption that a better outcome can be established. A recent study published the results of 6.5 year follow-up of children treated with ERT (agalsidase alfa), reporting a stabilizing effect. Another open label study is currently performed in children with a 5 year follow-up (agalsidase beta). Both studies included patients without severe symptoms. Unfortunately, an untreated control group was not included and moreover, FD has a slow disease progression rate. Interpretation of the results of these studies is therefore challenging. The only way to overcome this issue, is to perform a randomized controlled trial with a long follow-up, for both early treatment of classical FD patients and treatment of patients with non-classical FD patients (symptomatic and asymptomatic).

Although evidence is lacking, there is considerable support for treatment of individuals with non-classical FD. A clinical trial with an untreated arm, may not be supported by all FD experts. Alternatively, criteria may be drafted to allow for a change of study arm when disease progression occurs in the untreated group after a minimal study duration (e.g. 6 months). For example, individual with LVH, only remain untreated with a stable left ventricular mass or septal wall thickness. When progression of, for example, 2 mm occurs, the individual may change to the treated arm.

Besides treatment with ERT, supportive care has become a cornerstone in the management of patients with FD. Treatment with angiotensin converting enzyme inhibitors and Angiotensin II Receptor Blockers of patients with CKD and proteinuria is most likely to reduce progression of CKD. Furthermore, primary and secondary prevention with antiplatelet coagulation therapy
may reduce cardiovascular events. Besides studies on the efficacy of ERT, a third study arm would be needed to assess the effects of maximal supportive care only. It is not inconceivable, that supportive care will achieve similar or even better results when compared to ERT alone. The latter is, however, difficult to study prospectively, as it would be unethical to deny optimal supportive care. But it is feasible to study the efficacy of ERT in a controlled manner: ERT and optimal supportive care versus optimal supportive care only.

For the long-term follow-up, studies are performed with data from the international registries (FOS, Shire HGT and Fabry registry, Genzyme, a Sanofi company). However, those studies are limited by the observational nature and missing data, as well as the different protocols for follow-up at different centers 95. These registries cannot serve as a substitute of prospective clinical trials. Clinical trials for FD are complicated by the low patient numbers and need for long-term follow-up. This is true for ERT, but also for supportive care and new therapeutic options. Such studies are costly, and with the limited benefit for pharmaceutical companies, it will be challenging to find financial support. However, despite the limitations, in the current climate of critical appraisal of budgets for healthcare, proper controlled studies in large groups should be performed to assess efficacy of orphan drugs. It is essential that expert centers collaborate (internationally) to achieve this common goal.

To improve the evidence of efficacy of ERT, supportive care and future treatment options, we also need to invest in the understanding of the natural history of non-classical FD. To study the natural history of symptomatic individuals in a prospective manner is probably not feasible, as a long-term follow-up is needed, and this would imply that treatment (with ERT) is denied. However, asymptomatic family members are an interesting group to study prospectively, allowing for identification of the onset of LVH, CKD or WMLs in the regular assessments. A retrospective study is hampered by limited available data. However, the first identification and progression of cardiac, kidney or cerebrovascular disease can probably be retrieved. A combined retrospective approach for symptomatic, and a prospective approach for asymptomatic individuals will likely retrieve valuable data. The use of cross-sectional data is less attractive, because of the phenotypical heterogeneity. Importantly, the use of strict criteria to identify individuals with a non-classical phenotype essential, and as mentioned before, a sufficient dataset can only be achieved with international collaboration.

Who to test for FD?
Population screening for FD is extensively performed in research settings, and recently, newborn screening for FD was implemented in the state of Missouri, USA 96. Furthermore, technical developments have made it possible to apply screening of several genes at once to individual patients. And with the emerging possibilities to perform whole exome and whole genome sequencing (WES and WGS, respectively) for individual patients, it is expected that GVUS in the GLA gene will increasingly be identified 97.
Symptomatic individuals
It is interesting, that as a result of these developments, the diagnostic approach of an individual with a certain sign or symptom has changed. While earlier individuals were extensively ‘phenotyped’, it has become possible to test for several genetic causes at once, without first fully elucidating the signs and symptoms present in an individual.

A ‘conservative’ approach may be advocated in order to avoid the identification of GVUS, and the difficulties that are subsequently faced. An individual would be fully assessed, and if FD was suspected, an expert on FD is consulted. If characteristic features of classical FD are present, or if the family history is suggestive of FD, diagnostic tests for FD can subsequently be performed. However, with this approach, patients with a non-classical FD phenotype may easily be missed.

It is more realistic to assume that screening for FD in high-risk groups will continue and even increase. We should therefore strive to achieve protocols for genetic testing, to identify those who may benefit from genetic testing. In the Netherlands, screening of individuals with cardiomyopathy is performed using a cardio chip, testing over 40 genes at once. It was decided on a national level, that only individuals with HCM, in contrast to those with dilated cardiomyopathy (DCM), are tested for the GLA gene. The rationale behind this decision was that primary DCM is not seen in FD patients, unless with the coexistence of a second genetic cause for cardiomyopathy. In fact, one individual with DCM was identified to have the p.A143T GLA variant, prior to the implementation of this guideline. This has caused tremendous distress for this individual, and a cardiac biopsy was not feasible in this case due to his anxiety for the procedure. The distress could have been easily avoided, with the initial identification of individuals who are actually ‘at risk’ for FD, i.e. those with HCM only, and not those with DCM. As genetic testing for FD of individuals with HCM has also revealed patients with true FD (non-classical and classical phenotypes) in the Netherlands, it is acceptable that this is continued in clinical practice, but it is important to keep in mind that the yield is low, and we should accept that a cardiac biopsy is warranted for several individuals to confirm/exclude a diagnosis of FD. For individuals with CKD, it is yet unclear if screening will identify patients with true FD. In a research setting, a large screening cohort should first assess this, and such a study should include kidney biopsies for all individuals with a GLA variant. For individuals with TIA, stroke, WMLs or small fiber neuropathy, screening for FD is not indicated, as there are no tools to reliably confirm FD if a GLA variant is found. The best approach in these individuals is to consider the likelihood of FD on an individual basis. Of course, in individuals with CV, in the absence of medication that may cause CV, FD should always be considered. Angiokeratoma are more difficult to interpret, and testing for FD should be considered by an expert. Again, advantages and disadvantages apply and should always be considered carefully when testing for FD is considered. At least, it is important to inform all individuals that accidental findings may occur with genetic testing, and that further assessment in an expert center may be needed, before a definite diagnosis can be established. Importantly, they should also be informed that the significance of a genetic alteration may remain unclear. This statement relates to FD, but also applies, of course, to genetic testing in general.
Newborns

While a sign or symptom is the reason to test for FD in high-risk groups, this is not the case in newborn screening. Clinical features of FD are not yet present in the first years of life, complicating the assessment of pathogenicity of a GLA variant even more. Sometimes, family screening may be helpful, but this is often complicated by the coexistence of comorbidities (e.g. hypertension, diabetes mellitus). Furthermore, the disease has a heterogeneous phenotype, and even in the presence of a pathogenic variant, the disease course cannot be predicted, especially in females: individuals may remain asymptomatic for decades. Moreover, if the pathogenicity of a variant is established, there is insufficient evidence that early initiation of treatment is beneficial (see the section on treatment). Altogether, newborn screening will often result in significant distress for parents and their families and ultimately for the child himself, while a substantial number of GLA variants will not (or only very late) cause clinical manifestations of FD. We strongly discourage to implement FD in newborn screening programs, as the disease does not fit the criteria for population screening as defined by Wilson and Jungner ⁹⁸ and the benefits are, at best, very limited.
REFERENCES


84. Liao HC, Huang YH, Chen YJ, et al. Plasma globotriaosylphosphoglycerol (lysoGb3) could be a biomarker for Fabry disease with a Chinese


