

University of Groningen

The first European guidelines on phenylketonuria

Evers, Roeland A F; van Wegberg, Annemiek M J; Anjema, Karen; Lubout, Charlotte M A; van Dam, Esther; van Vliet, Danique; Blau, Nenad; van Spronsen, Francjan J

Published in:
Journal of Inherited Metabolic Disease

DOI:
[10.1002/jimd.12173](https://doi.org/10.1002/jimd.12173)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Final author's version (accepted by publisher, after peer review)

Publication date:
2019

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Evers, R. A. F., van Wegberg, A. M. J., Anjema, K., Lubout, C. M. A., van Dam, E., van Vliet, D., Blau, N., & van Spronsen, F. J. (2019). The first European guidelines on phenylketonuria: its usefulness and implications for BH4 responsiveness testing. *Journal of Inherited Metabolic Disease*, 43(2), 244-250. <https://doi.org/10.1002/jimd.12173>

Copyright

Other than for strictly personal use, 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), unless the work is under an open content license (like Creative Commons).

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

The first European guidelines on phenylketonuria: its usefulness and implications for BH₄ responsiveness testing

Roeland A.F. Evers, BSc¹; Annemiek M.J. van Wegberg, MSc¹; Karen Anjema, MD¹; Charlotte M.A. Lubout, MD¹; Esther van Dam, MSc¹; Danique van Vliet, MD¹; Nenad Blau, PhD²; Francjan J. van Spronsen, MD PhD¹.

¹University of Groningen, University Medical Center Groningen, Beatrix Children's Hospital, Division of Metabolic Diseases, The Netherlands.

²University Children's Hospital Zürich, Zürich, Switzerland.

Address correspondence to:

Prof. dr. F.J. van Spronsen
University Medical Center Groningen
Hanzeplein 1
9700 RB Groningen
The Netherlands
Mail: f.j.van.spronsen@umcg.nl
Telephone: +31(0)50-3614147

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/jimd.12173

Abstract

Objective: This study aimed to investigate and improve the usefulness of the 48-hour BH₄ loading test and to assess genotype for BH₄ responsiveness prediction, using the new definition of BH₄ responsiveness from the European guidelines, as well as an amended definition.

Method: Applying the definition of the European guidelines ($\geq 100\%$ increase in natural protein tolerance) and an amended definition ($\geq 100\%$ increase in natural protein tolerance or tolerating a safe natural protein intake) to a previous dataset (Anjema et al 2013), we firstly assessed the positive predictive value (PPV) of the 48-hour BH₄ loading test using a cutoff value of 30%. Then, we tried to improve this PPV by using different cutoff values and separate time points. Lastly, using the BIOPKU database, we compared predicted BH₄ responsiveness (according to genotype) and genotypic phenotypic values (GPVs) in BH₄-responsive and BH₄-unresponsive patients.

Results: The PPV of the 48-hour loading test was 50.0% using the definition of the European guidelines, and 69.4% when applying the amended definition of BH₄ responsiveness. Higher cutoff values led to a higher PPV, but resulted in an increase in false-negative tests. Parameters for genotype overlapped between BH₄-responsive and BH₄-unresponsive patients, although BH₄ responsiveness was not observed in patients with a GPV below 2.4.

Conclusion: The 48-hour BH₄ loading test is not as useful as previously considered and cannot be improved easily, whereas genotype seems mainly helpful in excluding BH₄ responsiveness. Overall, the definition of BH₄ responsiveness and BH₄ responsiveness testing require further attention.

Take-home message: The current 48-hour BH₄ loading test does not effectively predict BH₄ responsiveness as defined by the first European PKU guidelines.

Author contributions

R.A.F. Evers designed the study, analyzed the data, interpreted the results and was the lead writer of the manuscript. A.M.J. van Wegberg designed the study, interpreted the results and was the second lead writer of the manuscript. K. Anjema collected the data, interpreted the results and co-wrote the manuscript. C.M.A. Lubout, E. van Dam, D. van Vliet and N. Blau interpreted the results and co-wrote the manuscript. F.J. van Spronsen interpreted the results, co-wrote the manuscript and supervised the project. All authors read and approved the final version of the manuscript.

Name of the corresponding author

Prof. dr. F.J. van Spronsen

Conflicts of interest statement

R.A.F. Evers has received financial support from Biomarin for attending symposia. A.M.J. van Wegberg has received a research grant from Nutricia, honoraria from Biomarin as speaker, and travel grants from Nutricia and Vitaflo. K. Anjema has received research funding from Merck Serono. C.M.A. Lubout has received a speaker fee from the Recordati Rare Disease Foundation. E. van Dam has received advisory board fees from Merck Serono and Biomarin. D. van Vliet has received speaker's honoraria from Biomarin. N. Blau has no conflicts of interest to declare. F.J. van Spronsen is a member of scientific advisory boards for defects in amino acid metabolism of APR, Arla Food International, BioMarin, Eurocept Int, Lucana, Moderna TX, Nutricia, Rivium and SoBi, has received research grants from Alexion, Biomarin, Codexis, Nutricia, SoBi, and Vitaflo, has received grants from patient organizations ESPKU, Metakids, NPKUA, Stofwisselkracht, Stichting PKU research and Tyrosinemia Foundation, and has received honoraria as consultant and speaker from APR, Biomarin, MendeliKABS and Nutricia.

Funding

The authors received no funding for this research.

Ethical approval

For the original data collection, the BH₄ testing protocol was considered standard patient care by the Medical Ethical Committee of the University Medical Center Groningen. Informed consent was not required for this study.

Availability of data and material statement

Original data will be made available on reasonable request.

Keywords

Phenylketonuria; European guidelines; tetrahydrobiopterin; responsiveness; loading test; definition.

Introduction

The cornerstone of treatment in phenylketonuria (PKU; MIM 261600) is restricting phenylalanine (Phe) intake by a natural protein-restricted diet combined with intake of Phe-free amino acid supplements. Additionally, some patients benefit from pharmacological treatment with tetrahydrobiopterin (BH₄), which can increase residual phenylalanine hydroxylase activity leading to better metabolic control and/or an increase in natural protein tolerance. However, different definitions of BH₄ responsiveness exist (Singh and Quirk 2011, Anjema et al 2013, Vockley et al 2014).

Recently, the first European guidelines on PKU were published (van Spronsen, van Wegberg et al 2017, van Wegberg, A M J et al 2017). In these guidelines, BH₄ responsiveness is defined as “establishing an increase in natural protein tolerance of $\geq 100\%$ with blood Phe concentrations remaining consistently within the target range” or by improved metabolic control, which is defined as “ $>75\%$ of blood Phe levels remaining within target range without any decrease in natural protein intake associated with BH₄ treatment”. Since these criteria are stricter than previously used in the Netherlands (Anjema et al 2013), some patients in our population might no longer be considered BH₄ responsive when applying this definition. We noticed that this would even be the case for some patients who could actually tolerate a safe natural protein intake as a result of BH₄, meaning these patients could meet their protein requirements (according to WHO guidelines) using only natural protein sources, therefore not requiring additional amino acid supplements. Since these patients clearly

benefit from BH₄ treatment, we felt that the definition of BH₄ responsiveness from the European guidelines may need to be amended to include patients who can tolerate a safe natural protein intake due to BH₄.

The European PKU guidelines also give recommendations on the method of BH₄ responsiveness testing. With the exception of patients with a genotype consisting of two null mutations, in whom BH₄ responsiveness does not need to be further considered, it is recommended that BH₄ responsiveness testing is performed by a 48-hour BH₄ loading test. If Phe concentrations decrease with at least 30% during this test, a treatment trial should be performed to evaluate whether the patient is indeed BH₄ responsive. Although the 48-hour BH₄ loading test with a cutoff value of 30% is often cited as a reliable way to predict BH₄ responsiveness (Cerone et al 2013, Anjema et al 2013), its predictive value has not been assessed using the expert-based definition of BH₄ responsiveness that is stated by the European guidelines. Specifically, the study by Anjema et al. is cited as conformation of the utility of the 48-hour BH₄ loading test, but this study defined BH₄ responsiveness as an increase in natural protein intake of $\geq 50\%$ or ≥ 4 g/day, which is a much lower threshold (Anjema et al 2013). Therefore, the present study aimed to investigate and improve the usefulness of the 48-hour BH₄ loading test, and to assess the predictive value of genotype, firstly using the new definition of BH₄ responsiveness from the European guidelines, and secondly using an amended definition that also includes patients who can tolerate a safe natural protein intake due to BH₄.

Methods

Patients and protocol

We used data that were collected for a previous study on BH₄ responsiveness testing (Anjema et al 2013). Detailed information on the data collection, subjects and protocol were described by Anjema et al. (Anjema et al 2013). Here, the most relevant methodological aspects are summarized.

Data were collected retrospectively from 183 pediatric and adult patients who performed the 48-hour BH₄ loading test between November 2009 and December 2010. None of these tests took place in the neonatal period. For the 48-hour BH₄ loading test, baseline Phe concentrations were required to be over 400 μmol/L. Patients who had Phe concentrations below 400 μmol/L were, therefore, supplemented with Phe. Patients received 20 mg/kg BH₄ for two days (at t = 0 and 24 hours), while blood samples were collected at t = 0, 8, 16, 24 and 48 hours. Patients who showed a reduction in blood Phe concentrations ≥ 30% compared to t = 0 at any moment during the loading test were invited for a BH₄ treatment trial. Three-day dietary records were taken before and after the treatment trial to assess natural protein intake. During this treatment trial, BH₄ was introduced at 20 mg/kg/day (with a maximum of 1400 mg/day), dietary Phe was increased to reach the maximal Phe tolerance, and BH₄ dose was finally decreased if possible. In the original protocol, true BH₄ responsiveness was defined according to previously used guidelines in the Netherlands as “a reduction in blood Phe concentrations of 30% or more compared to mean blood Phe concentrations prior to the 48-hour BH₄ loading test with the same diet, and/or an increase in dietary Phe tolerance of ≥ 50% or ≥ 4 grams of natural protein without increasing the Phe concentrations above the upper target”. Data on genotype was collected if available. In total, 175 PKU patients correctly performed the 48-hour BH₄ loading test, and 65 patients performed the treatment trial (Table 1). Two patients from the original cohort were excluded, since it was found out that these patients had a DNAJC12 deficiency (van Spronsen, Himmelreich et al 2017).

Genotype

Using the BIOPKU database (<http://www.biopku.org>, accessed on January 29, 2019), we assessed two ways for using genotype to predict BH₄ responsiveness. Firstly, we used the BIOPKU database to collect the percentage of BH₄-responsive patients (including “slow” responders) with a corresponding genotype, when information on BH₄ responsiveness was available for ≥ 5 cases. Secondly, we used the BIOPKU database to assign genotypic phenotype values (GPVs) to the genotypes of the patients in this cohort. GPVs are as a numerical representation of predicted PAH activity in PKU patients, ranging from 0 (lowest PAH activity) to 10 (highest PAH activity) (Garbade et al 2018). Since BH₄ responsiveness is associated with higher levels of residual PAH activity, it was hypothesized that GPVs could be helpful in predicting BH₄ responsiveness. Patients who had a 48-hour BH₄ loading test that was considered positive (e.g. $\geq 30\%$ reduction in Phe levels) but did not perform a treatment trial were not included in these analyses.

Analyses

No analyses were copied from the original study. BH₄-responsiveness was assessed based on natural protein intake. The positive predictive value (PPV) of the 48-hour BH₄ loading test was calculated as the number of BH₄-responsive patients (based on the results of the treatment trial, using different definitions) divided by the number of potentially BH₄-responsive patients (based on the results of the 48-hour BH₄ loading test, using different cutoff values). Similarly, the negative predictive value (NPV) was calculated as the number of BH₄-unresponsive patients divided by the number of not potentially BH₄-responsive patients. Descriptive statistics were used to present most of the data. Normality of data was checked visually using histograms and QQ-plots, and tested using the Shapiro-Wilk test. Other statistical tests are mentioned where used. A two-tailed p -value < 0.05 was considered statistically significant. Statistical analyses were performed using IBM SPSS Statistics version 23 and GraphPad Prism version 7 for Windows.

Results

BH₄ responsiveness as defined by the European guidelines: the 48-hour BH₄ loading test and genotype

Of the 65 patients who performed a treatment trial, three had two putative null mutations (here defined as a GPV of 0 in the BIOPKU database) and were therefore excluded from the analyses in this part, as they would not have been involved in any BH₄ responsiveness testing following the recommendations of the European guidelines. All three patients were not considered to be BH₄ responsive in the original protocol. The positive predictive value (PPV) of the 48-hour BH₄ loading test was 50.0% when using the recommended cutoff value of 30% and the definition of BH₄ responsiveness recommended in the European guidelines.

Increasing the cutoff value to 35% led to a PPV of 57.4%, but beyond that, higher cutoff values were associated with a reduction in negative predictive values (NPVs) (Table 2). The PPV of separate time points varied between 51.7% (at t = 8) and 59.1% (at t = 24) using 30% as a cutoff value, with higher cutoff values again resulting in lower NPVs (Supplemental material 1). Furthermore, again looking at separate time points, 51.6%, 29.0%, 16.1% and 6.5% of BH₄-responsive patients showed no Phe decrease $\geq 30\%$ at t = 8, t = 16, t = 24, and t = 48, respectively.

In total, 58 patients had a genotype for which the percentage of BH₄-responsive patients in the BIOPKU database with a corresponding genotype (≥ 5 cases) was available (figure 1A). This BIOPKU responsiveness percentage was compared between patients considered BH₄ responsive and BH₄ unresponsive in this study. Significant differences were found between BH₄-responsive patients versus a total group of BH₄-unresponsive patients (including patients with no potential responsiveness after the loading test), as well as between patients considered BH₄ unresponsive after the loading test versus patients considered BH₄ unresponsive after the treatment trial. However, within the group of patients who performed the treatment trial, there was no difference between BH₄-responsive and BH₄-

unresponsive patients. In 106 patients, it was possible to assign GPVs to their genotype (figure 1B), showing a similar pattern. The lowest GPV in a BH₄-responsive patient was 2.4.

BH₄ responsiveness using an amended definition: the 48-hour BH₄ loading test and genotype

For a second set of analyses, we defined BH₄ responsiveness as “an increase in natural protein tolerance \geq 100% or tolerating a safe natural protein intake”, in which a safe natural protein intake was determined based on gender and age according to recommendations (WHO/FAO/UNU 2007). Using this definition, the number of BH₄-responsive patients in our cohort increased from 31 to 43, which is a significant increase of 38.7% ($p < 0.001$, McNemar test). Equally, again excluding the three patients with two null mutations, the PPV of the 48-hour BH₄ loading test (using the 30% cutoff value) increased from 50.0% to 69.4%.

Higher cutoff values resulted in higher PPVs, but were also associated with a decrease in the NPV (Supplemental material 2). The PPVs of having at Phe decrease \geq 30% at separate time points were between 71.7% (at $t = 48$) and 82.8% (at $t = 8$) (Supplemental material 3), with higher cutoff values again generally resulting in lower NPVs. Of BH₄-responsive patients, 44.2%, 23.8%, 19.0% and 11.6% showed no decrease in Phe \geq 30% at $t = 8$, $t = 16$, $t = 24$, and $t = 48$, respectively.

Parameters of genotype showed similar patterns as with the definition in European guidelines, although significant differences were now found between BH₄-responsive and BH₄-unresponsive patients who performed the treatment trial (Supplemental material 4). With the amended definition, the lowest GPV in a BH₄-responsive patient again was 2.4.

Discussion

Following the publication of the first European guidelines for PKU, this study assessed the usefulness of the 48-hour BH₄ loading test and genotype for BH₄ responsiveness prediction. Our results indicate

that the 48-hour BH₄ loading test is not useful for predicting BH₄ responsiveness as defined by the European PKU guidelines, whereas genotype seems mainly helpful in excluding BH₄ responsiveness. Additionally, we introduced an amended definition of BH₄ responsiveness, which in our opinion better defines which patients benefit from BH₄ treatment. This definition also leads to an increase in BH₄-responsive patients and results in a somewhat more effective 48-hour BH₄ loading test.

Before discussing our findings in more detail, some limitations of this study need to be addressed. Since this study used retrospectively collected patient care data, patients with a negative 48-hour BH₄ loading test did not perform a treatment trial, and the number of false negative tests is therefore unknown. Furthermore, we determined BH₄ responsiveness based on natural protein intake, and therefore did not assess using “>75% of blood Phe levels remaining within target range” as a definition. We anticipate however that this definition will be less regularly used, mostly since increasing the natural protein tolerance is the main goal of BH₄ treatment, as shown by longer-term follow-up studies of BH₄-treated patients (Keil et al 2013, Scala et al 2015, Evers et al 2018). Besides these points, it is important to note that this protocol of the 48-hour BH₄ loading test involved Phe supplementation in case of too low Phe concentrations, and did not take measurements of Phe levels at t = 32 and t = 40.

Since the introduction of BH₄ as a new treatment option for PKU, different definitions of BH₄ responsiveness have been proposed (Singh and Quirk 2011, Anjema et al 2013, Vockley et al 2014, van Spronsen et al 2017). The definition of BH₄ responsiveness recently proposed in the European guidelines is stricter than the definition previously used in the Netherlands, as shown by a decrease in the number of BH₄-responsive patients in this study compared to the original publication (33 versus 58 BH₄-responsive patients). As mentioned in the introduction, we noticed that some patients who did not meet the criteria of BH₄ responsiveness set by the European guidelines could tolerate a safe natural protein intake due to BH₄. In our cohort, this was the case for 12 patients. These patients could

Accepted Article

increase their tolerance with a mean of 22 grams of natural protein (range: 13 to 35 grams) during the BH₄ treatment trial, enabling them to adopt a more normal diet. Moreover, these patients typically did not need amino acid supplementation anymore. Considering these major advantages of dietary liberalization to this extent, the consensus-based definition in the European guidelines could, in our opinion, be improved by defining BH₄ responsiveness as “an increase in natural protein tolerance \geq 100% *or tolerating a safe natural protein intake*”.

Clearly, this amended definition is arbitrary as well. The criterion of increasing the natural protein tolerance by at least 100% may still be too strict, and seems especially difficult for patients with a relatively high baseline natural protein tolerance. This is also indicated by a significant difference in baseline natural protein intake between BH₄-responsive and BH₄-unresponsive patients (according to the definition of the European PKU guidelines) who performed a treatment trial (Table 1). Ultimately, to be able to give a more definitive definition of BH₄ responsiveness that is evidence-based rather than consensus-based, long-term follow up studies comparing outcomes in BH₄-treated to only dietary treated PKU patients are necessary. Such studies should identify the exact advantages of better metabolic control and/or increased natural protein intake as a result of BH₄ treatment in PKU. Since these data are not yet available, developing a (better) consensus-based definition of BH₄ responsiveness may be the best alternative. However, different definitions of BH₄ responsiveness may require different BH₄ responsiveness testing methods, as shown by the results in this study.

Our results show a low PPV of the 48-hour BH₄ loading test when applying the definition of BH₄ responsiveness from the European guidelines. Using the amended definition resulted in a higher PPV, albeit still much lower than the PPV of 87% that was reported in the original publication. Given the widespread use of the loading test, this is a worrying result, since it impairs the cost-effectiveness of BH₄ testing and may also lead to an increase in disappointed patients with a false-positive BH₄ loading test. Using the amended definition, our data show that a larger reduction in Phe during the

loading test is associated with a higher chance of being BH₄ responsive. However, considering the increase in false negatives with larger cutoff values, using a cutoff value higher than 30% would not be recommended according to the results in this cohort. Equally, the PPV of a decrease in Phe at separate time points may be somewhat higher, but is also associated with an increase in false-negative tests. Overall, the 48-hour BH₄ loading test cannot be improved easily, and its shortcoming should be kept in mind, especially when using the definition for BH₄ responsiveness stated in the European PKU guidelines.

To improve testing for potential BH₄ responsiveness, the simple 48-hour BH₄ loading test may need to be developed further. In this study, we unsuccessfully tried to improve the test by selecting higher cutoff values and by using separate measurements, but we were unable to investigate the predictive value of Phe measurements at t = 32 and t = 40, and Phe-to-tyrosine ratios. These latter approaches may thus deserve further attention. Additionally, extending the BH₄ loading test to 7 days (Nielsen, Nielsen and Güttler 2010) or even longer may be helpful in differentiating between daily Phe variation and a BH₄-induced decrease in Phe. On this topic, a comparison between the outcomes of the BH₄ testing regime in Europe (using a 48-hour BH₄ loading test) and the United States (using a 28-days BH₄ loading test) would be interesting. Recently, a group of experts proposed a BH₄ testing algorithm which combines the 48-hour BH₄ loading tests with testing periods up to 4 weeks (Muntau et al 2019). Apart from the fact that Muntau et al. did not recommend a specific definition of long-term BH₄ responsiveness, the value of the recommended testing protocol remains to be established. Alternatively, it could be argued that completely different testing methods are simply superior. Potential examples of this include comparing simple Phe loading with Phe + BH₄ loading (Porta, Spada and Ponzzone 2017), using a [¹³C]Phe breath test (Muntau et al 2002), or assessing genotype (Garbade et al 2018).

With regard to genotype, results of this study showed that all patients with a GPV below 2.4 showed no (potential) BH₄ responsiveness, which is in line with the report of Garbade et al. that mentions a cutoff value of 2 as a minimum for (potential) BH₄ responsiveness (Garbade et al 2018). Nevertheless, in line with that study, we found considerable overlap between BH₄-responsive and BH₄-unresponsive patients, with some unresponsive patients having a GPV as high as 10 (using the definition of the European guidelines) or 8.9 (using the amended definition). It should be noted that many BH₄-unresponsive patients only performed a loading test and no treatment trial, and some of them might therefore be false negatives. Nevertheless, this is an important finding that requires further examination. Overall, GPVs seem to be primarily helpful to exclude BH₄ responsiveness in patients in case of low residual PAH activity. Predicting BH₄ responsiveness using data on BH₄ testing in the BIOPKU database from patients with the same genotype shows similar results. Importantly, the BIOPKU database defines BH₄ responsiveness as a decrease in Phe levels of at least 30% during a short-term BH₄ loading test only, regardless of the specific protocol used. While genotype may be strongly related to the results of 48-hour BH₄ loading test (Garbade et al 2018), our results show that the 48-hour BH₄ loading test is not a good predictor for BH₄ responsiveness. Therefore, although genotype is helpful in excluding BH₄ responsiveness, it is not an overall good predictor of BH₄ responsiveness. These results show the shortcomings of both genotype and the 48-hour BH₄ loading to predict BH₄ responsiveness, thereby underlining the importance of a treatment trial.

To conclude, the 48-hour BH₄ loading test with a cutoff value of 30% does not effectively predict long-term BH₄ responsiveness as defined by the European PKU guidelines, whereas genotype seems mainly helpful to exclude BH₄ responsiveness. We recommend amending the definition of BH₄ responsiveness from the European PKU guidelines to include patients who can tolerate a safe natural protein intake as a result of treatment with BH₄. Methods to predict BH₄ responsiveness require further

attention, although a sound definition of BH₄ responsiveness may be even more important to optimize BH₄ responsiveness testing as well as treatment.

Acknowledgements

We thank everyone who was involved with the original publication, including patients, physicians, dieticians, researchers and laboratory technicians.

Anjema K, van Rijn M, Hofstede FC, et al (2013) Tetrahydrobiopterin responsiveness in phenylketonuria: prediction with the 48-hour loading test and genotype. *Orphanet J Rare Dis* 8: 103.

Cerone R, Andria G, Giovannini M, Leuzzi V, Riva E, Burlina A (2013) Testing for tetrahydrobiopterin responsiveness in patients with hyperphenylalaninemia due to phenylalanine hydroxylase deficiency. *Adv Ther* 30: 212-228.

Evers RAF, van Wegberg, Annemiek M. J., van Dam E, de Vries MC, Janssen MCH, van Spronsen FJ (2018) Anthropomorphic measurements and nutritional biomarkers after 5 years of BH4 treatment in phenylketonuria patients. *Mol Genet Metab* 124: 238-242.

Garbade SF, Shen N, Himmelreich N, et al (2018) Allelic phenotype values: a model for genotype-based phenotype prediction in phenylketonuria. *Genet Med*.

Keil S, Anjema K, van Spronsen FJ, et al (2013) Long-term follow-up and outcome of phenylketonuria patients on sapropterin: a retrospective study. *Pediatrics* 131: 1881.

Muntau AC, Adams DJ, Bélanger-Quintana A, et al (2019) International best practice for the evaluation of responsiveness to sapropterin dihydrochloride in patients with phenylketonuria. *Mol Genet Metab* 127: 1-11.

Muntau AC, Röschinger W, Habich M, Demmelmair H, Hoffmann B, Sommerhoff CP, Roscher AA (2002) Tetrahydrobiopterin as an alternative treatment for mild phenylketonuria. *N Engl J Med* 347: 2122-2132.

Nielsen JB, Nielsen KE, Güttler F (2010) Tetrahydrobiopterin responsiveness after extended loading test of 12 Danish PKU patients with the Y414C mutation. *J Inherit Metab Dis* 33: 9-16.

Porta F, Spada M, Ponzzone A (2017) Early Screening for Tetrahydrobiopterin Responsiveness in Phenylketonuria. *Pediatrics* 140.

Scala I, Concolino D, Della Casa R, et al (2015) Long-term follow-up of patients with phenylketonuria treated with tetrahydrobiopterin: a seven years experience. *Orphanet journal of rare diseases* 10: 14.

Singh RH, Quirk ME (2011) Using change in plasma phenylalanine concentrations and ability to liberalize diet to classify responsiveness to tetrahydrobiopterin therapy in patients with phenylketonuria. *Mol Genet Metab* 104: 485-491.

van Spronsen FJ, Himmelreich N, Rüfenacht V, et al (2017) Heterogeneous clinical spectrum of DNAJC12-deficient hyperphenylalaninemia: from attention deficit to severe dystonia and intellectual disability. *J Med Genet*.

van Spronsen FJ, van Wegberg AM, Ahring K, et al (2017) Key European guidelines for the diagnosis and management of patients with phenylketonuria. *Lancet Diabetes Endocrinol*.

van Wegberg, A M J, MacDonald A, Ahring K, et al (2017) The complete European guidelines on phenylketonuria: diagnosis and treatment. *Orphanet J Rare Dis* 12.

Vockley J, Andersson HC, Antshel KM, et al (2014) Phenylalanine hydroxylase deficiency: Diagnosis and management guideline. *Genetics in Medicine* 16: 188-200.

WHO/FAO/UNU (2007) Protein Amino Acid Requirements in Human Nutrition.

Table 1 Demographic and clinical details of the study cohort.

48-hour BH ₄ loading test result	No potential BH ₄ responsiveness ¹ (n = 97)	Potential BH ₄ responsiveness ¹ (n = 78)		
		BH ₄ responsive ² (n = 31)	BH ₄ unresponsive ² (n = 34)	Other ³ (n = 13)
Treatment trial result				
Gender (% female)	51.5	64.5	47.1	46.2
Age (yr)	14.7 (10.3 – 23.0)	12.0 (6.6 – 17.1)	12.5 (7.9 – 22.2)	14.1 (10.4 – 19.4)
Baseline Phe (μmol/L) ⁴	663 (530 – 915)	458 (362 – 549)***	509 (408 – 623)**	485 (386 – 590)*
Phe supplementation (%)	29.9	71.0***	58.8*	92.3***
Phe supplementation (mg/day) ⁵	200 (125 – 300)	335 (202 – 500)	350 (150 – 1000)	400 (210 – 1000)*
Natural protein intake prior to treatment trial (g/kg bw)		0.34 (0.23 – 0.43) [§]	0.52 (0.29 – 0.70)	

Data are presented as median (IQR) or as percentage of patients. ¹Based on a decrease in Phe levels of 30% as cutoff value. ²Based on an increase in natural protein tolerance of 100% as cutoff value. ³No treatment trial despite a positive 48-hour BH₄ loading test. ⁴At t = 0 during the 48-hour BH₄ loading test. ⁵Amount of Phe supplementation in patients who were supplemented with Phe. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ compared to patients with no potential BH₄ responsiveness (Mann-Whitney U test, corrected for multiple comparisons according to Bonferroni). [§] $p < 0.05$ compared to BH₄-unresponsive patients (Mann-Whitney U test).

Table 2 Predictive values using different cutoff values for a decrease in Phe levels during the 48-hour BH₄ loading test.

Phe decrease^a	PPV^b	Phe decrease^a	NPV^c
≥ 30%	50.0%		
≥ 35%	57.4%	30-35%	100.0%
≥ 40%	58.7%	30-40%	75.0%
≥ 45%	57.5%	30-45%	63.6%
≥ 50%	56.8%	30-50%	60.0%

^aMaximum Phe decrease compared to the baseline value. ^bPositive predictive value (PPV) using the corresponding cutoff value. ^cNegative predictive value (NPV) for a decrease in Phe in the corresponding range.

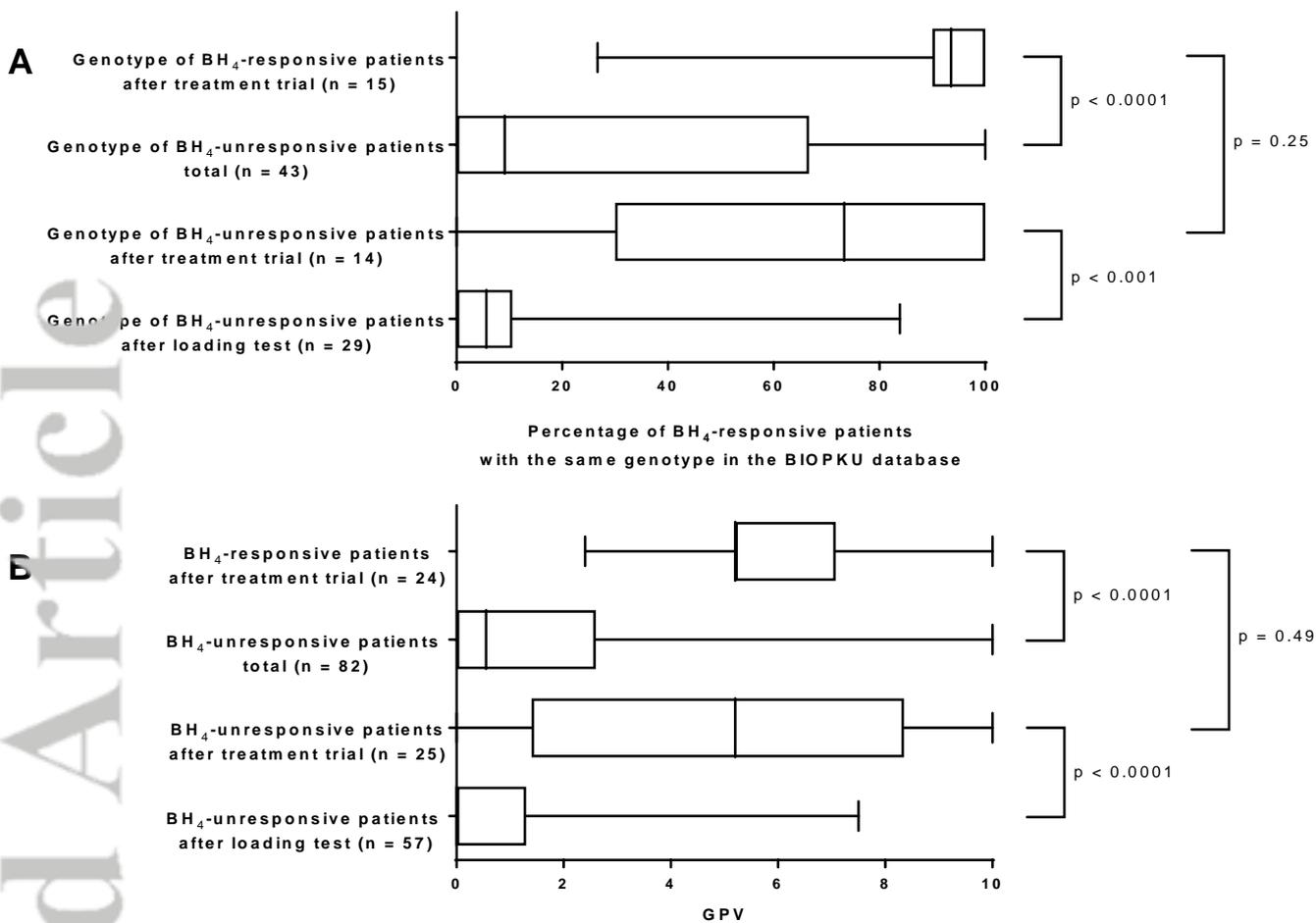


Figure 1 Association between genotype and BH₄ responsiveness.

Legend to figure 1:

A. Boxplots (median, 25th – 75th percentile, min, max) of the percentage of BH₄-responsive patients with a similar genotype in the BIOPKU database in BH₄-unresponsive and BH₄-responsive patients according to the definitions of the European PKU guidelines. Differences are tested with the Mann-Whitney U test. B. Boxplots (median, 25th – 75th percentile, min, max) of the genotypic phenotype value (GPV) in BH₄-unresponsive and BH₄-responsive patients according to the definitions of the European PKU guidelines. Differences are tested with the Mann-Whitney U test.