ORIGINAL ARTICLE

Blood phenylalanine concentrations in patients with PAH-deficient hyperphenylalaninaemia off diet without and with three different single oral doses of tetrahydrobiopterin: Assessing responsiveness in a model of statistical process control

M. Lindner · G. Gramer · S. F. Garbade · P. Burgard

Summary Tetrahydrobiopterin (BH_4) cofactor loading is a standard procedure to differentiate defects of BH_4 metabolism from phenylalanine hydroxylase (PAH) deficiency. BH_4 responsiveness also exists in PAH-deficient patients with high residual PAH activity. Unexpectedly, single cases with presumed nil residual PAH activity have been reported to be BH_4 responsive, too. BH_4 responsiveness has been defined either by a ≥30% reduction of blood Phe concentration after a single BH_4 dose or by a decline greater than the individual circadian Phe level variation. Since both methods have methodological disadvantages, we present a model of statistical process control (SPC) to assess BH_4 responsiveness. Phe levels in 17 adult PKU patients of three phenotypic groups off diet were compared without and with three different single oral dosages of BH_4 applied in a double-blind randomized cross-over design. Results are compared for ≥30% reduction and SPC. The effect of BH_4 by ≥30% reduction was significant for groups ($p < 0.01$) but not for dose ($p = 0.064$), with no interaction of group with dose ($p = 0.24$). SPC revealed significant effects for group ($p < 0.01$) and the interaction for group with dose ($p < 0.05$) but not for dose alone ($p = 0.87$). After one or more loadings, seven patients would be judged to be BH_4 responsive either by the 30% criterion or by the SPC model, but only three by both. Results for patients with identical PAH genotype were not very consistent within (for different BH_4 doses) and between the two models. We conclude that a comparison of protein loadings without and with BH_4 combined with a standardized procedure for data analysis and decision would increase the reliability of diagnostic results.

Abbreviations

BH_4	tetrahydrobiopterin
bw	bodyweight
HPA	hyperphenylalaninaemia
MHP	mild hyperphenylalaninaemia
PAH	phenylalanine hydroxylase
Phe	phenylalanine
PKU	phenylketonuria
SDS	standard deviation score
SPC	statistical process control
Tyr	tyrosine
U(L)CL	upper (lower) control limit

M. Lindner (✉) · G. Gramer · S. F. Garbade · P. Burgard
Division of Metabolic Disorders, Department of General Paediatrics, University Children's Hospital,
Im Neuenheimer Feld 430,
69120 Heidelberg, Germany
e-mail: Martin.lindner@med.uni-heidelberg.de

Introduction

Phenylketonuria (PKU; OMIM #261600) is an inborn error of metabolism resulting from deficiency of the enzyme phenylalanine hydroxylase (PAH, EC 1.14.16.1), which requires tetrahydrobiopterin (BH_4)

as a cofactor to convert phenylalanine (Phe) into tyrosine. In untreated patients, PAH deficiency leads to accumulation of Phe in blood and other tissues, resulting in mental retardation and severe neurological symptoms. PKU was the first metabolic disorder diagnosed in population-based newborn screening programmes. The therapy consists of a low-protein diet supplemented with a synthetic Phe-free amino acid mixture. Treatment is highly successful (Burgard et al 1999; Scriver and Kaufman 2001) but requires continuous compliance of patients and their families and negatively affects quality of life (MacDonald 2000; MacDonald et al 1997; Weglage et al 1996).

Following the discovery of hyperphenylalaninaemias (HPAs) due to defects in cofactor metabolism, BH$_4$ loading has been performed for differential diagnosis in patients detected by newborn screening (Danks et al 1976; Niederwieser et al 1979). Whereas PAH-deficient HPA does not respond, cofactor deficiencies show a sharp decline in blood phenylalanine concentrations (Niederwieser et al 1985).

In 1999 it was shown for the first time that also patients with PAH-deficient HPA can respond to BH$_4$ with a substantial reduction of blood Phe concentrations (Kure et al 1999). Meanwhile numerous studies and case reports have been published, especially in patients with milder forms of PAH deficiency (Bernegger and Blau 2002; Desivat et al 2004; Fiori et al 2005; Lindner et al 2001; 2003; Matalon et al 2004; Muntau et al 2002; Perez-Duenas et al 2004). BH$_4$ is now available as sapropterin dihydrochloride (KUVAN) in the United States and in Europe.

Two different criteria for assessment of responsiveness to a single BH$_4$ loading dose have been described so far. The most widely accepted criterion was first proposed by an expert group and published by Blau and Erlandsen (2004), defining BH$_4$ responsiveness as a decline in blood Phe concentration of at least 30% from the Phe concentration determined immediately before loading. That definition bears the problems that the cut-off is arbitrary, equal absolute decreases can result in different percentages, and dietary fluctuations during the loading test as well as normal intra-day variation are ignored. Alternatively, Leuzzi and colleagues (2006) calculated for each patient an individual 95% confidence interval for the mean using the standard deviation of six Phe levels (= baseline) measured during the day before BH$_4$ loading. Patients are classified as responders when Phe after BH$_4$ loading falls below the lower limit of the 95% confidence interval. However, transferring the standard deviation of the baseline to determine the confidence interval for the Phe level immediately

before BH$_4$ loading can be misleading in those cases where this level and the mean are different. In order to circumvent this problem, we suggest a model of statistical process control (SPC) to analyse the effects of three different doses of BH$_4$ (10, 20, 30 mg/kg bw) on blood Phe levels. As most previous studies used 20 mg/kg bw, we hypothesized that 10 mg/kg bw could also show an effect in milder forms and possibly 30 mg/kg bw could show an effect in more severe forms. Within subjects, the design would allow to investigate a possible dose-effect relationship.

Many studies do not report complete information on genotype, metabolic phenotype, concurrent treatment and absolute phenylalanine concentrations before and after BH$_4$ loading and cover a large age range from birth to adulthood. We therefore investigated a cohort of adult PKU patients off diet, well characterized by genotype and metabolic phenotype according to a standardized protein loading to answer the following questions:

1. How are the results of the 30% criterion related to those of the SPC model?
2. Is BH$_4$ responsiveness dose dependent?
3. Is BH$_4$ responsiveness related to HPA phenotype?
4. Is the effect of BH$_4$ identical in patients with identical genotype?

Patients and methods

Patients

Seventeen patients with PKU (10 male, 7 female; mean age 27.8 years (range 19.3–36.8 years)) were included in the study. All patients had been diagnosed in the neonatal period and were continuously treated during childhood and adolescence. Exclusion criteria were pregnancy or breast-feeding in women, and neuropsychiatric or other chronic disease possibly interfering with study participation (e.g. seizures, asthma, diabetes mellitus). All patients gave written informed consent to participation in the study. The study conformed to the tenets of the Declaration of Helsinki (Version Tokyo 2004) and was approved by the German Federal Institute for Drugs and Medical Devices (Number 4022739) and the institutional review board of the Medical Faculty of the University of Heidelberg (EC I, Number 272/2004).

Classification of metabolic phenotypes

Patients were classified into three phenotypic groups according to their Phe concentration 72 hours (72 h)

after a standardized protein load with 180 mg Phe/kg bw and day performed during a former study (Lutz et al 1990) or during clinical routine as mild PKU, Phe 10–17 mg/dl (600–1020 μmol/L); moderate PKU, Phe 17.1–26 mg/dl (1021–1560 μmol/L); classical PKU, Phe > 26 mg/dl (1560 μmol/L). This classification was verified by comparison with the phenotypic classification of genotypes (Guldberg et al 1998; Güttler et al 1999), where each mutation is assigned a value (AV) of 1 for classic, 2 for moderate, 4 for mild PKU and 8 for mild hyperphenylalaninaemia (MHP). The sum of the AVs of both alleles predicts the metabolic phenotype. Two patients were identical to classical genotype p.R158Q/p.R408W, two to p.R408W/IVS12 + 1G>A and three to mild genotype p.Y414C/IVS12 + 1G>A. Two patients with mutation p.R261Q in combination with two different putative null mutations were evaluated as functionally hemizygous for a moderate genotype.

BH$_4$ loading

Patients on a Phe-restricted diet including synthetic Phe-free amino acid supplements ($n = 10$) stopped the diet 6 days prior to the first day of BH$_4$ loading and kept to a diet with 50 mg Phe/kg bw and day until the end of the study. Patients off diet ($n = 7$) were asked not to change their nutrition and total protein intake. All patients provided nutrition protocols for four days during the study period. Morning and evening Phe concentrations on days −3 to −1 (time −72 h to −12 h, Fig. 1) before BH$_4$ loading were analysed.

BH$_4$ ((6R)-L-*erythro*-5,6,7,8-tetrahydrobiopterin) was obtained in 50 mg tablets from Schircks Laboratories (Jona, Switzerland). On days 0 (time 0 h, symbol ▲ in Fig. 1), 7 (time 168 h, symbol ♦ in Fig. 1) and 14 (time 336 h, symbol ■ in Fig. 1) BH$_4$ was administered after an overnight fast in single oral doses of 10, 20, or 30 mg/kg bw following a randomized double-blind protocol. BH$_4$ was dissolved in water or orange juice immediately before application. Blood samples were taken before (T0) and 4, 8, 12, 16, 24 h (T4, T8, T12, T16, T24) after loading to measure BH$_4$ responsiveness. Blood samples were also taken at times 36, 48, 60, 72, 84 h to monitor the wash-out period after the first and at times 204, 216, 228, 240, 252, and 324 h to monitor the wash-out after the second loading.

Blood sampling and laboratory analysis

For each patient 36 capillary blood samples were spotted on Whatman S&S 903 filter paper for analysis of Phe, Tyr and BH$_4$ metabolites. Samples taken by the patients at home were stored at ambient temperature, collected in the hospital during study visits and then stored at −18°C until analysis. Storage time at home ranged from 0 to 25 days (mean 6.83 days, SD = 4.1); storage time of frozen samples ranged from 36 to 216 days (mean 107.7 days, SD = 45.6). Phe and Tyr levels were analysed using ESI tandem mass spectrometry as previously described (Schulze et al 2003).

Pterin concentrations were measured from the dried blood spots taken at the days of BH$_4$ loadings
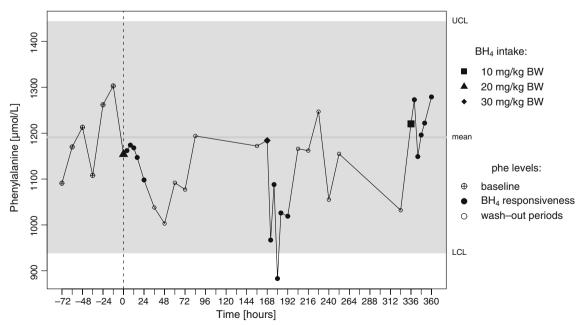

Fig. 1 Data of ID 10, moderate PKU, at baseline, BH$_4$ loadings and wash-outs

(metabolic laboratory, University Children's Hospital Zürich, Switzerland) using HPLC as described previously. As BH_4 is extremely unstable in collected blood and about 30–40% is metabolized to pterin, total biopterin was calculated as the sum of biopterin and pterin (B + P) (Zurflüh et al 2005).

Monitoring of side-effects

Before as well as 30 and 60 minutes after BH_4 loading, blood pressure, heart rate and breathing rate were monitored. Any adverse event occurring during the antecedent study period was recorded at study visits.

Definition of BH_4 responsiveness by the 30% criterion

Patients were defined as BH_4 responsive if blood Phe concentration decreased by ≥30% at any time between 4 and 24 h after BH_4 loading (Blau and Erlandsen 2004).

Definition of BH_4 responsiveness by statistical process control (SPC)

Based on phenylalanine blood levels measured in the morning and in the afternoon of each of three consecutive days off diet and before the first BH_4 application (baseline), we modelled a control chart with the upper (UCL) and lower control limits (LCL) about the mean of all six values. Control limits were defined by the following equations:

$$UCL = mean + 3\,SD \qquad LCL = mean - 3\,SD$$

where SD is the standard deviation of the six values. Under the assumption that the mean is normally distributed, the probability that a value below the LCL occurs is <0.00135. As the same is true for the UCL, the probability for a value outside the control limits is 0.0027. All values within the control limits are interpreted to be part of a normal process of fluctuation, whereas all values outside the control range—at any time—are interpreted as not belonging to the process. This model allows judgement whether the Phe concentration before BH_4 loading lies within the limits of the process and Phe levels after BH_4 application fall below the LCL. To compare the results across individual patients, all Phe levels before and after BH_4 loading were transformed to standard deviation scores (SDSs) by subtracting each Phe level from the mean of the baseline values and dividing the difference by the SD of the baseline values (for a detailed description of the method, see Montgomery (2005; pp. 150)). Figure 1 gives an example of patient 10 (moderate PKU). The patient is diagnosed as responsive after 30 mg BH_4/kg bw but not for 10 and 20 mg.

Results

Pearson correlations between Phe levels 72 h after standardized protein loading and the sum of assigned values (AV1 + AV2) of the genotypes were $r = -0.85$ ($p < 0.0001$), with mean baseline Phe concentrations off-diet $r = 0.76$ ($p < 0.001$), with T0 before 10 mg $BH_4 r = .69$ ($p < 0.01$), with T0 before 20 mg $BH_4 r = 0.81$ ($p < 0.0001$), and with T0 before 30 mg $BH_4 r = 0.81$ ($p < 0.0001$). These correlations indicate the internal validity of the classification systems and the reproducibility of the patients' phenotypes by baseline Phe levels as well as Phe levels immediately before BH_4 loadings.

Paired t-tests of 72 h versus mean baseline Phe levels were not significant (ns) for mild (t (5) = 0.7; ns) and moderate t (1) = 0.8; ns) patients. Mean baseline concentrations were significantly lower than 72 h Phe in the group of classic PKU (t (8) = 14.5; $p < 0.001$). However, it is not surprising that in the present study classic PKU patients did not reach the same Phe levels as in the standardized protein loading with an extremely high amount of 180 mg Phe/kg body weight. Nevertheless, the Phe levels in this group indicate that patients followed study instructions and were indeed off diet. Two-way analysis of variance of Phe levels with phenotype group as one factor and Phe concentrations at baseline, T0 before10 mg, T0 before 20 mg, and T0 before 30 mg BH_4 as a repeated measurement factor revealed a significant main effect for phenotype group ($F(2, 14) = 16.37$; $p < 0.0001$). Main effect for the repeated measurement factor ($F(3, 42) = 2.32$; $p < 0.09$) and its interaction with phenotype ($F(6,$

Table 1 Means (SD) of blood Phe levels (μmol/L) for Phe 72 h, at baseline, and T0 before 10 mg, T0 before 20 mg, and T0 before 30 mg by phenotype

Phenotype	Phe 72 h	Phe baseline	Phe T0 before 10 mg	Phe T0 before 20 mg	Phe T0 before 30 mg
Mild $n = 6$	869 (116)	761 (168)	796 (194)	714 (185)	821 (229)
Moderate $n = 2$	1347 (253)	1064 (179)	1170 (70)	1096 (81)	1038 (206)
Classic $n = 9$	2060 (136)	1182 (139)	1192 (190)	1296 (270)	1345 (133)

Table 2 Means (SD) of Phe level change after loading with 10, 20, and 30 mg BH_4/kg bw for the 30% criterion and the SPC model

Phenotype group	Criterion of 30%: mean % ↓Phe (SD)			SPC model: mean SDS ↓Phe (SD)		
	10 mg BH_4	20 mg BH_4	30 mg BH_4	10 mg BH_4	20 mg BH_4	30 mg BH_4
Mild $n = 6$	27.17 (10.26)	27.67 (8.66)	39.17 (13.38)	2.42 (1.60)	2.92 (0.79)	3.16 (1.25)
Moderate $n = 2$	12.50 (9.19)	2.5 (3.54)	20.50 (6.36)	0.37 (0.19)	0.17 (1.33)	2.57 (1.54)
Classic $n = 9$	11.89 (9.92)	16.89 (10.76)	16.11 (12.99)	1.35 (0.95)	1.26 (0.99)	0.22 (1.95)

42) = 1.6; $p < .18$) were not significant. The non-significant repeated measurement factor indicates that all phenotypic groups had similar Phe levels during the baseline period off-diet and immediately before the three BH_4 loadings (Table 1). In four loadings of 3 patients (IDs 14, 4, 18) T0 levels were above the UCL, but leading to a false positive result only in the 20 mg loading of ID 18 (Table 4). Measurement errors as a possible cause for extremely high Phe levels were excluded by a second analysis.

BH₄ responsiveness by dose and PKU phenotype

Means and standard deviations of change scores according to the 30% criterion and the SPC model after loading with 10, 20, and 30 mg BH_4/kg bw are shown in Table 2.

BH₄ responsiveness according to the 30% criterion

Two-way ANOVA of percentages of Phe-level decrease with phenotype as one factor and BH_4 dose as repeated measurement factor revealed a significant main effect for group ($F(2, 14) = 9.03$; $p < 0.01$), and a non-significant dose effect ($F(2,28) = 3.042$; $p = 0.064$). Mean decrease was 17.4% (SD = 12.0) after 10 mg, 19.0% (SD = 12.1) after 20 mg, and 24.8% (SD = 16.3) after 30 mg BH_4/kg bw. The interaction of phenotype group with BH_4 dose was not significant ($F(4, 28) = 1.46$; $p = 0.24$).

BH₄ responsiveness by statistical process control (SPC)

Two-way ANOVA of SPC SDS of Phe-level decrease with phenotype as one factor and BH_4 dose as repeated measurement factor revealed a significant effect for phenotype ($F(2,14) = 6.65$; $p < 0.01$) and a significant interaction of phenotype with BH_4 dose ($F(4,28) = 3.64$; $p < 0.05$). Main effect of BH_4 dose ($F(2, 28) = 0.14$; $p = 0.87$) was not significant. SD scores showed a continuous increase in the mild group, and no effect in the classic group, whereas in the two moderate patients there was a remarkable effect after the 30 mg dose (Table 2).

Similarities and differences between the 30% criterion and the SPC model

Both algorithms identified nearly identical rates of responsive judgements, 13 in the SPC and 12 in the 30% model. Overall correspondence between the two criteria (Table 3) was significant (Fisher' exact test $p < 0.01$). Forty (78.4%) of all 51 BH_4 loadings were scored identical by both models as non-responsive ($n = 33$; 64.7%) or responsive ($n = 7$; 13.7%). However, agreement was better for diagnosing non-responsiveness than for responsiveness. From the 12 loading results judged as responsive by the 30% criterion, only 7 were also scored as responsive by the SPC model; and from the 13 loadings scored as responsive by the SPC model, only 7 were judged as responsive by the 30% criterion. The kappa statistic for agreement of categorical data revealed a $\kappa = 0.42$, i.e. a fair to moderate agreement (Landis and Koch 1977). After one or more loadings, seven patients would be judged to be BH_4 responsive either according to the 30% criterion or to the SPC model, but only three by both (Table 4). The 30% criterion judges 5/6 mild and 0/2 moderate patients to be responsive; the rates for the SPC algorithm are 4/6 for mild and 1/2 for moderate phenotypes. Interestingly both approaches classify two, though different, classic patients to be responsive.

Identical genotypes

Three patients (2, 5, 9) were identical with regard to mild genotype (p.Y414C/IVS12 + 1G>A (AV4/AV1). Two patients with moderate phenotype (7, 10) had functionally identical genotypes p.R261Q/g.822_832del

Table 3 Correspondence of all loading results according to the 30% criterion and the SPC model

	30% responsive	30% non-responsive	Σ
SPC responsive	7	6	13
SPC non-responsive	5	33	38
Σ	12	39	51

Fisher's exact test $p < 0.01$; $\kappa = 0.42$.

Table 4 Genotypic and phenotypic description of the sample and results of the 30% and SPC analysis (1 triplet and 3 pairs of identical genotypes in boxes). Negative figures indicate increase of blood Phe levels. Shaded cells indicate BH$_4$ responsiveness

ID	Group	72 h Phe mg/dl	Metabolic phenotype	Genotype	AV 1+2	Classification by Zurflühet al (2008)	%↓Phe 10 mg BH$_4$/kg bw	%↓Phe 20 mg BH$_4$/kg bw	%↓Phe 30 mg BH$_4$/kg bw	↓SD 10 mg	↓SD 20 mg	↓SD 30 mg
17	1	13.8	Mild	p.A395G/IVS10-11G>A	8+1=9	not in OPKUdb/unclear	38	31	44	3.27	3.39	4.02
16	1	11.0	Mild	p.Y414C/p.R261Q	4+2=6	resp/resp	35	43	63	1.79	3.84	3.91
11	1	15.7	Mild	p.R261Q/p.R261Q	2+2=4	resp/resp	26	27	25	4.11	3.28	1.90
9	1	14.1	Mild	p.Y414C/IVS12+1G>A	4+1=5	resp/non-resp	33	23	35	4.01	3.11	4.67
5	1	16.5	Mild	p.Y414C/IVS12+1G>A	4+1=5	resp/non-resp	11	18	38	0.98	1.98	2.91
2	1	15.0	Mild	p.Y414C/IVS12+1G>A	4+1=5	resp/non-resp	20	24	30	0.38	1.93	1.56
10[+]	2	19.3	moderate	p.R261Q/p.K724_Y277>Nfs	2+1=3	resp/non-resp	6	5	25	0.50	1.11	3.66
7[a]	2	25.2	moderate	p.R261Q/p.Y166X	2+1=3	resp/not in BIOPKUdb	19	0	16	0.23	-0.77	1.48
3	3	33.5	Classic	p.R158Q/p.R408W	1+1=2	resp/non-resp	6	26	-6	0.78	1.18	0.28
8	3	33.5	Classic	p.R158Q/p.R408W	1+1=2	resp/non-resp	16	13	29	0.78	0.31	0.29
15	3	33.2	Classic	p.R408W/IVS12+1G>A	1+1=2	non-resp/non-resp	19	15	14	1.41	2.35	0.30
6	3	32.4	Classic	p.R408W/IVS12+1G>A	1+1=2	non-resp/non-resp	-4	18	36	1.13	0.71	0.92
13	3	37.6	Classic	IVS12+1G>A/	1+1=2	non-resp/non-resp	24	7	17	3.23	1.82	1.98
18	3	36.8	Classic	IVS12+1G>A/p.L48S	1+1(4)=2(5)	non-resp/resp	23	39[b]	28[b]	1.60	0.76[b]	1.26[b]
14	3	34.7	Classic	p.P281L/p.R408W	1+1=2	non-resp/non-resp	16	21[b]	8	2.07	0.02[b]	0.36
12	3	30.1	Classic	p.R158Q/p.R158Q	1+1=2	resp/resp	5	8	10	1.25	3.10	1.33
4	3	34.5	Classic	p.P281L/p.P281L	1+1=2	non-resp/non-resp	2	5	9[b]	-0.14	1.10	-4.72*
						Sum responsive	**3**	**3**	**6**	**4**	**5**	**4**

[a] Functionally identical genotypes.

[b] T0 Phe levels above UCL.

Negative figures indicate increase of blood Phe levels. Shaded cells indicate BH$_4$ responsiveness.

and p.R261Q/p.Y166X. Two patients (3, 8) had identical classical genotype p.R158Q/p.R408W, and two patients (15, 6) were identical for classical genotype p.R408W/IVS12 + 1G>A. Results were not very consistent within (for different BH_4 doses) and between the two models. The 30% criterion identified the three mild genotypes p.Y414C/IVS12 + 1G>A as non-responsive after the 20 mg dose but as responsive after the 30 mg dose, whereas the SPC model diagnosed only one of them as responsive, for all three doses. One of the functionally hemizygous moderate genotypes (p.R261Q/p.K724_Y277>Nfs) was classified as responsive by the SPC model, but only for the 30 mg dose. The SPC model rejected responsiveness for both classical pairs, whereas the 30% criterion would define one of the patients with genotype p.R408W/IVS12 + 1G>A as responsive (Table 4).

Pharmacokinetics of BH_4 after 10, 20 or 30 mg BH_4/kg bw

In all patients peaks of B + P levels (measured as nmol/g Hb) were reached 4 h after BH_4 loading showing a dose effect but no phenotype effect. B + P max levels and BH_4 response were not associated (Gramer et al 2009). Tyrosine levels and Phe/Tyr ratios did not show any significant changes following BH_4 application (data not shown).

Side-effects

No side-effects, as reported by Fiege and colleagues (2004), were observed in our patients after BH_4 loading with 10, 20 or 30 mg/kg bw.

Discussion

The present study is the first conducted in a double-blind randomized controlled repeated design to measure the effects of single oral dosages of 10, 20 or 30 mg BH_4/kg body weight on blood Phe concentrations in three well-defined phenotypic groups of patients with hyperphenylalaninaemia off diet in a metabolic state corresponding to their metabolic phenotype and confirmed by genotype information. The study introduced for the first time the concept of statistical process control to the analysis of BH_4 responsiveness in hyperphenylalaninaemias in comparison with the criterion of 30% blood Phe level reduction.

In a substantial number of observations both criteria came to identical results in diagnosing non-responsiveness. However, although the absolute numbers of

patients as well as of Phe loading results judged as responsive were nearly identical, agreement with respect to individual patients or loadings was roughly only 50%.

With one exception (ID 12), loadings judged as responsive by the SPC algorithm but not by the 30% criterion showed substantial decreases of Phe concentrations between 20% and 30% (IDs 16, 11, 9, 10, 13). In contrast, with one exception (ID 5), loadings judged as responsive by the 30% criterion but not by the SPC algorithm showed relatively small decreases in SD values (0.38–1.93; IDs 2, 6, 18). This was particularly true for ID 18 with classical PKU, in whom the 30% criterion would have predicted a false-positive response because the Phe concentration immediately before the application of BH_4 was above the patient's SPC domain. The patient's nutrition protocol for this day reports adherence to the recommended daily Phe intake. Thus, a low intake during the study day cannot explain the extremely rapid decline after BH_4 loading. Use of the evening level before BH_4 loading instead of the high morning Phe level before BH_4 loading resulted in non-responsiveness. Although classified as classical PKU phenotype, one of this patient's mutations (p.L48S) has been described as variable in severity (Guldberg et al 1998) with AVs of 1, 2 and 4. Several patients homozygous for p.L48S have been found to be BH_4 responsive with a maximum Phe decrease of 60% 24 h after BH_4 loading (Fiege et al 2005; Leuzzi et al 2006) and recently the mutation has been classified as BH_4-responsive (Zurflüh et al 2008). Although it was not possible to definitively resolve this issue, it is evident that careful interpretation of single loading tests is critical in patients with classical PKU.

ANOVA showed different patterns of significance for the two approaches with non-significant effects for dose and interaction of dose with phenotypes of mild, moderate and classic PKU according to the 30% criterion but a clearly significant interaction of dose by phenotype in the SPC model. As the significant interaction corresponds with the postulated mechanisms to explain the effect of BH_4 (Blau and Erlandsen 2004), this result can be regarded as an argument for the superior validity of the SPC model. A dose effect of BH_4 irrespective of the genotype and the predicted residual PAH activity and of BH_4 dose in particular would not be a rational biochemical expectation.

ANOVA main effects for phenotype groups were significant in both approaches, which is corroborated well by the results of larger studies and can be seen as an argument for the internal validity of our design. However, this result should be interpreted in relation to BH_4 dose.

Most BH$_4$-responsive patients defined by the 30% criterion reported so far have had mild phenylketon-uria or even mild hyperphenylalaninaemia with no need for treatment. In our mild phenotype group only one out of three patients with genotype p.Y414C/ IVS12 + 1G>A (ID 9) was consistently responsive in the SPC model, but only in two of three loadings according to the 30% model. The other two responded only to the highest BH$_4$ dose in the 30% model but not at all in SPC. There is only one other report on this particular genotype describing a patient as BH$_4$-responsive with a Phe reduction of 82% 15 h after BH$_4$ loading (Muntau et al 2002). The mutation p.Y414C with a high residual activity in vitro is common in BH$_4$-responsive PKU patients when present in homozygous or in compound heterozygous state (Bernegger and Blau 2002; Fiori et al 2005; Lambruschini et al 2005; Lindner et al 2001, 2003; Matalon et al 2005).

The response to BH$_4$ is smaller in patients with moderate PKU and in most studies negligible with classical PKU (Bernegger and Blau 2002; Desviat et al 2004; Fiori et al 2005; Kure et al 1999; Lindner et al 2003; Perez-Duenas et al 2004; Weglage et al 2002). However, some authors have reported BH$_4$ respon-siveness in patients with 'classical PKU phenotype' carrying the mutation p.R158Q (Perez-Duenas et al 2004; Spaapen and Rubio-Gozalbo 2003), but for this mutation in vitro expression revealed a residual enzyme activity up to 10% and responsive patients always showed PAH mutations on the second allele, e. g. p.R261Q and p.A300S, that have been reported with variable BH$_4$ responsiveness. ID 15 with genotype p. R158Q/IVS12 + 1G>A was not responsive in either algorithm and the same was found for the two patients with genotype p.R158Q/p.R408W (IDs 3, 8). Given the observed variable responses within phenotypic groups and even within individuals in many cases, a single trial approach to diagnose BH$_4$ responsiveness should be regarded as unreliable.

The model of statistical process control shows many similarities to the kinetic analysis of BH$_4$ response described as a single exponential decay (SED) by Langenbeck (2008). That author criticizes the 30% criterion for the same reasons as we do. Percentages are difficult to compare, use only limited information of a BH$_4$ loading, and lack the contrast of blood Phe decreases with and without BH$_4$. However, regardless of the elegance of this method, the statistical evalua-tion of the contrast of data with and without BH$_4$ is not solved by the SED approach.

Recently Levy and colleagues (2007) have proposed an algorithm to evaluate BH$_4$ responsiveness in PKU and its use in treatment. The logic of the suggested diagnostic algorithm is very similar to our approach but could be quantitated by statistical process control, taking into account changes in the central tendency and variability of blood Phe concentrations before and after supplementation of BH$_4$. In particular the significance of effects after changes of BH$_4$ dosage could be evaluated more precisely. In addition, the algorithm allows for the evaluation of the clinical significance of an observed BH$_4$ responsiveness. A transition of a patient's clinical phenotype—e.g. from moderate to mild PKU after BH$_4$ supplementation resulting in a meaningful change of the dietary regimen or even a change from mild PKU to MHP making further dietary treatment dispensable—would be a powerful argument for treatment with BH$_4$.

References

Bernegger C, Blau N (2002) High frequency of tetrahydrobiop-terin-responsiveness among hyperphenylalaninemias: a study of 1919 patients observed from 1988 to 2002. *Mol Genet Metab* **77**: 304–313. doi:10.1016/S1096-7192(02)00171-3

Blau N, Erlandsen H (2004) The metabolic and molecular bases of tetrahydrobiopterin-responsive phenylalanine hydroxy-lase deficiency. *Mol Genet Metab* **82**: 101–111. doi:10.1016/j. ymgme.2004.03.006

Burgard P, Bremer HJ, Bührdel P, et al (1999) Rationale for the German recommendations for phenylalanine level control in phenylketonuria 1997. *Eur J Pediatr* **158**: 46–54. doi:10.1007/s004310051008

Danks DM, Cotton RGH, Schlesinger P (1976) Variant forms of phenylketonuria. *The Lancet* **307**(7971): 1236–1237. doi:10.1016/S0140-6736(76)92179-6

Desviat LR, Perez B, Belanger-Quintana A, et al (2004) Tetrahydrobiopterin responsiveness: results of the BH$_4$ loading in 31 Spanish PKU patients and correlation with their genotype. *Mol Genet Metab* **83**: 157–162. doi:10.1016/j. ymgme.2004.06.007

Fiege B, Ballhausen D, Kierat L, et al (2004) Plasma tetrahy-drobiopterin and its pharmacokinetic following oral admin-istration. *Mol Genet Metab* **81**: 45–51. doi:10.1016/j. ymgme.2003.09.014

Fiege B, Bonafé L, Ballhausen D, Baumgartner M, Thöny B, Meili D (2005) Extended tetrahydrobiopterin loading test in the diagnosis of cofactor-responsive phenylketonuria: a pilot study. *Mol Genet Metab* **86**(Supplement 1): S91–95. doi:10.1016/j.ymgme.2005.09.014

Fiori L, Fiege B, Riva E, Giovannini M (2005) Incidence of BH4-responsiveness in phenalalanine-hydroxylase deficient Italian patients. *Mol Genet Metab* **86**(Supplement 1): S67–74. doi:10.1016/j.ymgme.2005.06.017

Gramer G, Garbade SF, Blau N, Lindner M (2009) Pharmacokinetics of tetrahydrobiopterin following oral loadings with three single dosages in patients with phenylketonuria. *J Inherit Metab Dis* **32**(1): 52–57. Epub 2008 Nov 21(in press).

Guldberg P, Rey F, Zschocke J, et al (1998) European multicenter study of phenylalanine hydroxylase deficiency: classification of 105 mutations and a general system for genotype-based prediction of metabolic phenotype. *Am J Hum Genet* **63**: 71–79. doi:10.1086/301920

Güttler F, Azen C, Guldberg P, et al (1999) Relationship among genotype, biochemical phenotype, and cognitive performance in females with phenylalanine hydroxylase deficiency: report from the maternal phenylketonuria collaborative study. *Pediatrics* **104**(2 Pt 1): 258–262. doi:10.1542/peds.104.2.258

Kure S, Hou DC, Ohura T, et al (1999) Tetrahydrobiopterin-responsive phenylalanine hydroxylase deficiency. *J Pediatr* **135**: 375–378. doi:10.1016/S0022-3476(99)70138-1

Lambruschini N, Perez-Duenaz B, Vilaseca MA, et al (2005) Clinical and nutritional evaluation of phenylketonuric patients on tetrahydrobiopterin monotherapy. *Mol Genet Metab* **86**(Supplement 1): S54–S60. doi:10.1016/j.ymgme.2005.05.014

Langenbeck U (2008) Classifying tetrahydrobiopterin responsiveness in the hyperphenylalaninaemias. *J Inherit Metab Dis* **31**: 67–72. doi:10.1007/s10545-007-0572-4

Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* **33**: 159–174. doi:10.2307/2529310

Leuzzi V, Carducci C, Carducci C, et al (2006) The spectrum of phenylalanine variations under tetrahydrobiopterin load in subjects affected by phenylalanine hydroxylase deficiency. *J Inherit Metab Dis* **29**: 38–46. doi:10.1007/s10545-006-0096-3

Levy H, Burton B, Cederbaum S, Scriver C (2007) Recommendations for evaluation of responsiveness to tetrahydrobiopterin (BH4) in phenylketonuria and its use in treatment. *Mol Genet Metab* **92**: 287–291. doi:10.1016/j.ymgme. 2007.09.017

Lindner M, Haas D, Zschocke J, Burgard P (2001) Tetrahydrobiopterin responsiveness in phenylketonuria differs between patients with the same genotype. *Mol Genet Metab* **73**: 104–106. doi:10.1006/mgme.2001.3168

Lindner M, Steinfeld R, Burgard P, Schulze A, Mayatepek E, Zschocke J (2003) Tetrahydrobiopterin sensitivity in German patients with mild phenylalanine hydroxylase deficiency. *Hum Mutat* **21**: 400. doi:10.1002/humu.9117

Lutz P, Schmidt H, Batzler U (1990) Study design and description of patients. *Eur J Pediatr* **149**(Supplement 1): S5–12. doi:10.1007/BF02126292

MacDonald A (2000) Diet and compliance in phenylketonuria. *Eur J Pediatr* **159**: S136–141. doi:10.1007/PL00014375

MacDonald A, Harris G, Rylance G, Asplin DA, Booth IW (1997) Abnormal feeding behaviours in phenylketonuria. *J Hum Nutr Diet* **10**(3): 163–170. doi:10.1046/j.1365-277X.1997.00050.x

Matalon R, Koch R, Michals-Matalon K, et al (2004) Biopterin responsive phenylalanine hydroxylase deficiency. *Genet Med* **6**: 27–32. doi:10.1097/01.GIM.0000108840.17922.A7

Matalon R, Michals-Matalon K, Koch R, Grady J, Tyring S, Stevens RC (2005) Response of patients with phenylketonuria in the US to tetrahydrobiopterin. *Mol Genet Metab* **86**(Supplement 1): S17–21. doi:10.1016/j.ymgme.2005.06.024

Montgomery DG (2005) Introductions to statistical Statistical quality Quality controlControl, 5th edition. New York:John Wiley & Sons, New York.

Muntau AAC, Röschinger W, Habich M, et al (2002) Tetrahydrobiopterin as an alternative treatment for mild phenylkentonuria. *N Engl J Med* **347**: 2122–2132. doi:10.1056/NEJMoa021654

Niederwieser A, Curtius HC, Viscontini M, Schaub J, Schmidt H (1979) Phenylketonuria variants. *The Lancet* **313**(8115): 550. doi:10.1016/S0140-6736(79)90966-8

Niederwieser A, Ponzone A, Curtius HC (1985) Differential diagnosis of tetrahydrobiopterin deficiency. *J Inherit Metab Dis* **8**(Supplement 1): 34–38. doi:10.1007/BF01800657

Perez-Duenas B, Vilaseca MA, Mas A, et al (2004) Tetrahydrobiopterin responsiveness in patients with phenylketonuria. *Clin Biochem* **37**: 1083–1090. doi:10.1016/j.clinbiochem.2004.09.005

Schulze A, Lindner M, Kohlmüller D, Olgemöller K, Mayatepek E, Hoffmann GF (2003) Expanded newborn screening for inborn errors of metabolism by electrospray ionization-tandem mass spectrometry: results, outcome, and implications. *Pediatr Pediatrics* **111**(6 Pt 1): 1399–406. doi:10.1542/peds.111.6.1399

Scriver CR, Kaufman S (2001) Hyperphenylalaninemias: phenylalanine hydroxylase deficiency. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds; Childs B, Kinzler KW, Vogelstein B, assoc. eds. *The Metabolic and Molecular Bases of Inherited Disease*, 8th edn. New York: McGraw-Hill, 1667–1724.

Spaapen LJ, Rubio-Gozalbo ME (2003) Tetrahydrobiopterin-responsive phenylalanine hydroxylase deficiency, state of the art. *Mol Genet Metab* **78**: 93–99. doi:10.1016/S1096-7192(02)00229-9

Weglage J, Funders B, Ullrich K, Rupp A, Schmidt E (1996) Psychosocial aspects in phenylketonuria. *Eur J Pediatr* **155**: S101–104. doi:10.1007/PL00014225

Weglage J, Grenzebach M, von Teeffelen-Heithoff A, et al (2002) Tetrahydrobiopterin responsiveness in a large series of phenylketonuria patients. *J Inherit Metab Dis* **25**: 321–322. doi:10.1023/A:1016514727870

Zurflüh MR, Giovannini M, Fiori L, et al (2005) Screening for tetrahydrobiopterin deficiencies using dried blood spots on filter paper. *Mol Genet Metab* **86**(Supplement 1): S96–103. doi:10.1016/j.ymgme.2005.09.011

Zurflüh MR, Zschocke J, Lindner M, Feillet F, Chery C, Burlina A (2008) Molecular genetics of tetrahydrobiopterin-responsive phenylalanine hydroxylase deficiency. *Hum Mutat* **29**: 167–75. doi:10.1002/humu.20637