Common Drug Review

Clinical Review Report

September 2017

Drug	Sapropterin dihydrochloride (Kuvan)
Indication	In conjunction with a phenylalanine (Phe)-restricted diet to reduce blood Phe levels in patients with hyperphenylalaninemia (HPA) due to tetrahydrobiopterin-(BH_4)-responsive phenylketonuria (PKU).
Listing request	Ongoing funding of sapropterin (Kuvan) for non-pregnant patients and patients actively planning pregnancy who have a diagnosis of PKU and who have demonstrated a response to the initial 6 month trial of sapropterin and who meet ALL of the following criteria: 1. Compliance with low protein diet, formulas, and treatment with sapropterin; **AND** 2. Has achieved a) normal sustained blood Phe levels [Greater than 120 µmol/L and less than 360 µmol/L] (At least 2 levels measured at least 1 month apart); OR b) sustained blood Phe reduction of at least 30% (At least 2 levels measured at least 1 month apart) compared to baseline if the Phe baseline level is less than 1200 µmol/L; OR c) sustained blood Phe reduction of at least 50% (At least 2 levels measured at least 1 month apart) compared to baseline if the Phe baseline level is greater than 1200 µmol/L; **AND** 3. Demonstrated increase of dietary protein tolerance based on targets set between the clinician and patient; **OR** 4. Clinically meaningful age-appropriate improvement in: a) neurobehavioural or neurocognitive function or impairment for patients with such impairments as determined by peer reviewed clinically validated scales; OR b) demonstrated improvement in Quality of Life using peer reviewed validated scales; AND 5. Managed by a physician specialized in metabolic/biochemical diseases.
Dosage form(s)	100 mg oral tablets
NOC date	April 30, 2010
Manufacturer	BioMarin Pharmaceutical Inc.

This review report was prepared by the Canadian Agency for Drugs and Technologies in Health (CADTH). In addition to CADTH staff, the review team included two clinical experts in metabolic/biochemical diseases who provided input on the conduct of the review and the interpretation of findings.

The information in this document is intended to help Canadian health care decision-makers, health care professionals, health systems leaders, and policy-makers make well-informed decisions and thereby improve the quality of health care services. While patients and others may access this document, the document is made available for informational purposes only and no representations or warranties are made with respect to its fitness for any particular purpose. The information in this document should not be used as a substitute for professional medical advice or as a substitute for the application of clinical judgment in respect of the care of a particular patient or other professional judgment in any decision-making process. The Canadian Agency for Drugs and Technologies in Health (CADTH) does not endorse any information, drugs, therapies, treatments, products, processes, or services.

While care has been taken to ensure that the information prepared by CADTH in this document is accurate, complete, and up-to-date as at the applicable date the material was first published by CADTH, CADTH does not make any guarantees to that effect. CADTH does not guarantee and is not responsible for the quality, currency, propriety, accuracy, or reasonableness of any statements, information, or conclusions contained in any third-party materials used in preparing this document. The views and opinions of third parties published in this document do not necessarily state or reflect those of CADTH.

CADTH is not responsible for any errors, omissions, injury, loss, or damage arising from or relating to the use (or misuse) of any information, statements, or conclusions contained in or implied by the contents of this document or any of the source materials.

This document may contain links to third-party websites. CADTH does not have control over the content of such sites. Use of third-party sites is governed by the third-party website owners' own terms and conditions set out for such sites. CADTH does not make any guarantee with respect to any information contained on such third-party sites and CADTH is not responsible for any injury, loss, or damage suffered as a result of using such third-party sites. CADTH has no responsibility for the collection, use, and disclosure of personal information by third-party sites.

Subject to the aforementioned limitations, the views expressed herein are those of CADTH and do not necessarily represent the views of Canada's federal, provincial, or territorial governments, or any third-party supplier of information.

This document is prepared and intended for use in the context of the Canadian health care system. The use of this document outside of Canada is done so at the user's own risk.

This disclaimer and any questions or matters of any nature arising from or relating to the content or use (or misuse) of this document will be governed by and interpreted in accordance with the laws of the Province of Ontario and the laws of Canada applicable therein, and all proceedings shall be subject to the exclusive jurisdiction of the courts of the Province of Ontario, Canada.

The copyright and other intellectual property rights in this document are owned by CADTH and its licensors. These rights are protected by the Canadian *Copyright Act* and other national and international laws and agreements. Users are permitted to make copies of this document for non-commercial purposes only, provided it is not modified when reproduced and appropriate credit is given to CADTH and its licensors.

TABLE OF CONTENTS

ABBREVIATIONS IV

EXECUTIVE SUMMARY VI

1. INTRODUCTION 1
   - 1.1 Disease Prevalence and Incidence 1
   - 1.2 Standards of Therapy 1
   - 1.3 Drug 2

2. SUBMISSION HISTORY 4
   - 2.1 Basis of Resubmission 5

3. OBJECTIVES AND METHODS 6
   - 3.1 Objectives 6
   - 3.2 Methods 6

4. RESULTS 8
   - 4.1 Findings From the Literature 8
   - 4.2 Included Studies 10
   - 4.3 Patient Disposition 21
   - 4.4 Exposure to Study Treatments 22
   - 4.5 Critical Appraisal 23
   - 4.6 Efficacy 24
   - 4.7 Harms 27

5. DISCUSSION 30
   - 5.1 Summary of Available Evidence 30
   - 5.2 Interpretation of Results 31
   - 5.3 Potential Place in Therapy 35

6. CONCLUSIONS 37

APPENDIX 1: PATIENT INPUT SUMMARY 38
APPENDIX 2: LITERATURE SEARCH STRATEGY 41
APPENDIX 3: EXCLUDED STUDIES 44
APPENDIX 4: DETAILED OUTCOME DATA 45
APPENDIX 5: VALIDITY OF OUTCOMES 61
APPENDIX 6: SUMMARY OF OTHER STUDIES 64
APPENDIX 7: DIETARY TREATMENT FOR PHENYLKETONURIA 80
APPENDIX 8: PROPOSED REIMBURSEMENT CRITERIA FROM ONTARIO SPECIALISTS 83

REFERENCES 90

Tables

Table 1: Summary of Results xii
Table 2: Inclusion Criteria for the Systematic Review 6
Table 3: Details of Included Studies 9
Table 4: Summary of Baseline Demographic and Disease Characteristics (Intention-to-Treat Population) 13
Table 5: Patient Disposition 21
Table 6: Exposure to Study Drug (Intention-to-Treat Population) 22
Table 7: Harms (Safety Population) 28
Table 8: Study PKU-016: ADHD-RS/ASRS Total Score (Phe Responders With ADHD Symptoms) 45
Table 9: Study PKU-016: ADHD-RS/ASRS Inattention Subscale Score (Phe Responders With ADHD Symptoms) 46
Table 10: Study PKU-016: ADHD-RS/ASRS Hyperactivity-Impulsivity Subscale Score (Phe Responders With ADHD Symptoms) 47
Table 11: Study PKU-016: Global Function Evaluation (CGI-I) in Phe Responders 48
Table 12: Study PKU-016: Global Function Evaluation (CGI-S) in Phe Responders 48
Table 13: Study PKU-016: Hamilton Anxiety Rating Scale Total Score in Phe Responders 49
Table 14: Study PKU-016: Hamilton Depression Rating Scale Total Score in Phe Responders 51
Table 15: Study PKU-016: BRIEF-Parent T Scores in Phe Responders < 18 Years of Age 52
Table 16: Study PKU-016: BRIEF-A T Scores in Phe Responders ≥ 18 Years of Age 53
Table 17: Study PKU-016: Mean Phe Blood Levels in Phe Responders 54
Table 18: SPARK Study: Mean Phe Blood Levels (Intention-To-Treat Population) 55
Table 19: SPARK Study: Proportion of Patients Maintaining Phe Blood Levels Within Therapeutic Target (Intention-to-Treat Population) 56
Table 20: SPARK Study: Adjusted Mean Treatment Difference in Dietary Phe Tolerance Based on Prescribed Phe (mg/kg/day) (Intention-To-Treat Population) 57
Table 21: SPARK Study — Mean Change in Dietary Phe Tolerance (mg/kg/day) From Baseline to End of Treatment (Intention-To-Treat Population) 58
Table 22: SPARK Study — Adjusted Means Over Time and Treatment Difference at Week 26 in Change from Baseline in Growth Parameters (Intention-to-Treat Population) 58
Table 23: SPARK Study: Neuromotor Developmental Milestones (Intention-To-Treat Population) 59
Table 24: Instrument Descriptions 63
Table 25: Patient Characteristics and Demographics 64
Table 26: Blood Phe Levels: Pre-Sapropterin to Six Years of Exposure 66
Table 27: Dietary Phe: Prescribed, Actual, and Percent Change in Actual Dietary Phe Intake 67
Table 28: Blood Phe Levels in Patients Aged Zero to Four Years 68
Table 29: Dietary Phe in Patients Aged Zero to Four Years: Prescribed, Actual, and Change in Actual Dietary Phe Intake 69
Table 30: Harms 70
Table 31: Baseline Characteristics and Demographics 72
Table 32: Patient Disposition 73
Table 33: Harms 73
Table 34: Blood Phe Levels, Dietary Phe Intake, and Natural Protein Intake in PAH-Deficient Patients 74
Table 35: Other Secondary Outcomes 75
Table 36: Baseline Characteristics 76
Table 37: Sapropterin Responsiveness 77

Table 38: Blood Phe Concentrations, Neurocognitive Function, and Growth Assessments In Children on Sapropterin Treatment 78
Table 39: Harms 79

Figures

Figure 1: Flow Diagram for Inclusion and Exclusion of Studies 8
Figure 2: Study Design of Study PKU-016 11
Figure 3: Study Design of the SPARK Study 12

ABBREVIATIONS

ACMG	American College of Medical Genetics and Genomics
ADHD	attention-deficit/hyperactivity disorder
ADHD-RS	Attention Deficit Hyperactivity Disorder Rating Scale
ANCOVA	analysis of covariance
ASRS	Adult Attention Deficit Hyperactivity Disorder Self-Report Scale
AE	adverse event
BL	baseline
BMI	body mass index
BRI	Behavioral Regulation Index
BRIEF	Behavior Rating Inventory of Executive Function
CMH	Cochran–Mantel–Haenszel
CGI-I	Clinical Global Impression–Improvement
CGI-S	Clinical Global Impression–Severity
CI	confidence interval
DB	double-blind
GEC	Global Executive Composite
EAP	Exceptional Access Program
HAM-A	Hamilton Anxiety Rating Scale
HAM-D	Hamilton Depression Rating Scale
HPA	hyperphenylalaninemia
IQ	intelligence quotient
ITT	intention-to-treat population
IVRS	interactive voice response system
LOCF	last observation carried forward
MI	Metacognition Index
NA	not applicable
NR	not reported
OL	open-label
PC	placebo-controlled
PG	parallel-group
Phe	phenylalanine
PKU	phenylketonuria
PO	by mouth
PP	per-protocol
RCT	randomized controlled trial

RR	relative risk
SAE	serious adverse event
SAP	sapropterin
SD	standard deviation
SDS	standard deviation score
SE	standard error
WDAE	withdrawal due to adverse event

EXECUTIVE SUMMARY

Introduction

Phenylalanine hydroxylase (PAH) deficiency, traditionally known as phenylketonuria (PKU), is an autosomal recessive inborn error of metabolism caused by mutations in the gene encoding PAH, the enzyme that converts phenylalanine (Phe) to tyrosine.[1,2] If untreated, Phe accumulates to toxic levels in the brain, which can lead to profound neurocognitive, neuropsychiatric, and developmental problems. PKU occurs in approximately one in every 12,000 to 15,000 infants born in North America, accounting for approximately 300 new cases each year.[3] The main treatment for PKU is a Phe-restricted diet; however, the strict diet imposes economic and social hardships and often leads to non-adherence, especially among adolescents and young adults.[4] PKU places a large burden on patients and their families and has a substantial effect on patients' quality of life. Patients' expectations are that sapropterin (SAP) will lower Phe levels and increase Phe tolerance so that patients are able to eat a more varied diet and be less isolated from their peers.

Kuvan (sapropterin dihydrochloride) is a synthetic formulation of tetrahydrobiopterin (BH_4), a necessary co-factor for the PAH enzyme that hydroxylates Phe through an oxidative reaction to form tyrosine.[5] In patients with PKU who are responders, treatment with BH_4 can activate residual PAH enzyme, improve the oxidative metabolism of Phe, and decrease Phe blood levels in some patients.[5] Kuvan is available as 100 mg oral tablets and the recommended starting dose is 10 mg/kg/day daily.[5] Prior to treatment, responders are typically identified by evaluating the reduction in blood Phe following treatment with 20 mg/kg/day of SAP for a period of up to four weeks.[5] A 30% reduction from baseline in mean Phe blood levels is often cited as evidence of effective response.[1] Once responsiveness to SAP has been established, the dosage may be adjusted within the range of 5 to 20 mg/kg/day, according to response.[5]

Submission History

In January 2011, the Canadian Expert Drug Advisory Committee (CEDAC) issued a Final Recommendation that Kuvan not be listed.[6] The key reason for the recommendation was that patient details were insufficient to identify a subpopulation for whom SAP may provide a significant clinical benefit that is cost-effective. The Final CEDAC Recommendation was based on a review of Kuvan in November 2010 and a Request for Reconsideration in January 2011.[6] A Request for Advice was also made by the participating drug plans in October 2011, which did not result in any changes to the recommendation.[7]

The objective of this review is to perform an updated systematic review of the beneficial and harmful effects of SAP (dose: 5 to 20 mg/kg/day), when used in conjunction with a Phe-restrictive diet, to reduce blood Phe levels in patients with HPA due to BH_4-responsive PKU.

Indication under review
In conjunction with a Phe-restricted diet to reduce blood Phe levels in patients with hyperphenylalaninemia (HPA) due to BH_4-responsive PKU.
Reimbursement criteria requested by sponsor
Ongoing funding of sapropterin (Kuvan) for non-pregnant patients and patients actively planning pregnancy who have a diagnosis of PKU and who have demonstrated a response to the initial 6 month trial of sapropterin and who meet ALL of the following criteria: 1. Compliance with low protein diet, formulas, and treatment with sapropterin; AND 2. Has achieved a) normal sustained blood Phe levels [Greater than 120 µmol/L and less than 360 µmol/L] (At least 2 levels measured at least 1 month apart); OR b) sustained blood Phe reduction of at least 30% (At least 2 levels measured at least 1 month apart) compared to baseline if the Phe baseline level is less than 1200 µmol/L; OR c) sustained blood Phe reduction of at least 50% (At least 2 levels measured at least 1 month apart) compared to baseline if the Phe baseline level is greater than 1200 µmol/L; AND 3. Demonstrated increase of dietary protein tolerance based on targets set between the clinician and patient; OR 4. Clinically meaningful age-appropriate improvement in: a) neurobehavioural or neurocognitive function or impairment for patients with such impairments as determined by peer reviewed clinically validated scales; OR b) demonstrated improvement in Quality of Life using peer reviewed validated scales; AND 5. Managed by a physician specialized in metabolic/biochemical diseases.

Results and Interpretation

Included Studies

Two phase 3b randomized controlled trials (RCTs) met the selection criteria for inclusion in the current systematic review. PKU-016 (N = 206) was a double-blind, placebo-controlled, parallel-arm trial in patients aged eight years and older with PKU who were on a Phe-restricted diet. PKU-016 consisted of two 13-week treatment periods: a double-blind randomized treatment period comparing SAP 20 mg/kg/day and placebo (both in conjunction with a Phe-restricted diet) and an open-label treatment period in which patients randomized to placebo in the first treatment period crossed over to open-label SAP, while maintaining a Phe-restricted diet. The SPARK study (n = 56) was an open-label, parallel-arm trial in patients younger than four years with PKU who were on a Phe-restricted diet and who had Phe blood levels within the target range (120 to 360 µmol/L). The SPARK study compared SAP 10 mg/kg/day plus a Phe-restricted diet with a Phe-restricted diet alone over 26 weeks. There were two co-primary outcomes in PKU-016, which were the change from baseline to week 13 in the Attention Deficit Hyperactivity Disorder Rating Scale (ADHD-RS) or Adult ADHD Self-Report Scale (ASRS) total score in Phe responders with ADHD symptoms and the proportion of patients with a Clinical Global Impression–Improvement (CGI-I) scale rating score of 1 (very much improved) or 2 (much improved) at week 13 in Phe responders with or without attention-deficit/hyperactivity disorder (ADHD) symptoms. The primary outcome in the SPARK trial was the change in dietary Phe tolerance from baseline at week 26. Key limitations of the trials are the lack of a matched placebo in the SPARK study, the small patient numbers and relatively short duration of the trials, lack of validation of measured outcomes or minimal clinically important differences (MCIDs) in patients with PKU, lack of control for multiplicity in the statistical analyses, and the large placebo effect observed in both trials.

Efficacy

Efficacy outcomes identified in the review protocol were quality of life, health care resource utilization, change in Phe blood levels, neurophysiologic and neurocognitive effects, growth parameters, nutritional status, proportion of responders, and change in Phe tolerance. The included trials did not comprise any outcomes pertaining to quality of life, health care resource utilization, or nutritional status.

In Study PKU-016, in the population of Phe responders, the mean (SD) Phe blood levels at baseline were 680.2 (435.44) µmol/L in the SAP + diet arm and 789.5 (464.97) µmol/L in the placebo + diet arm. At week 13, Phe levels decreased by approximately 30% from baseline in the SAP + diet arm and remained largely unchanged in the placebo + diet arm. After week 13, Phe levels remained relatively sTable in the SAP + diet arm, but decreased in the placebo + diet arm as these patients crossed over to receive open-label SAP + diet during the open-label treatment period. At week 26, mean Phe levels were similar between the two arms, although the Phe levels exceeded the upper limit of the Phe target range (360 µmol/L) as recommended by the American College of Medical Genetics and Genomics (ACMG).[1] No statistical comparisons of Phe levels were conducted between treatment arms in PKU-016. In the SPARK study, patients were required to be within the Phe target range (120 to 360 µmol/L) at study entry. Mean (standard deviation [SD]) Phe baseline levels were [redacted] ([redacted]) µmol/L in the SAP + diet arm and [redacted] ([redacted]) µmol/L in the diet alone arm. Phe levels in the SAP + diet and diet alone arms were relatively constant throughout the study. At week 26, the mean (standard error [SE]) change from baseline was –10.1 ([redacted]) µmol/L in the SAP + diet arm and 23.1 ([redacted]) µmol/L in the diet alone arm and the difference was not statistically significantly different. These findings imply that diet alone can maintain Phe blood levels in the target range.

The proportion of Phe responders was not an outcome in either PKU-016 or the SPARK study; however, in PKU-016, those patients who demonstrated a 20% reduction from baseline in Phe blood levels following treatment with SAP 20 mg/kg/day for up to one month were considered to be Phe responders. Of the 206 patients randomized in PKU-016, 118 patients (57.3%) were Phe responders. It is not known how many additional patients may have been considered Phe responders if treated with SAP for longer than one month; however, four weeks is typically considered adequate for determining SAP response.[1] In the SPARK study, all patients entering the trial were required to be BH_4 responders.

Although mean Phe blood levels were reduced from baseline in all treatment arms in both trials, the clinical significance of the magnitude of the reduction is unclear. In PKU-016, the reduction in Phe blood levels, both in patients originally randomized to SAP + diet or who crossed over to open-label SAP + diet, was approximately 30%. The clinical significance of this reduction is difficult to ascertain as the MCID for Phe blood level reduction (i.e., Phe response) is unknown and the per cent reduction in Phe blood levels used to identify Phe responders (i.e., 20% or 30%) is arbitrary, as confirmed by the clinical expert consulted on this review. The ACMG guidelines state that clinical judgment is required to determine what constitutes a significant or beneficial decline in Phe blood levels from baseline in an individual and that 30% is often cited as evidence of effective Phe reduction, although some centres do consider 20% reduction as a measure of response.[1] In both trials, diet remained unchanged; therefore, there is no evidence from the included trials to show that increased diet liberalization was possible due to the addition of SAP to dietary therapy. The target range for mean Phe blood levels of 120 to 360 µmol/L is controversial and it has been suggested that the target range should be less than 360 µmol/L up to age 12 years and then less than 600 µmol/L thereafter.[8] There does, however, appear to be agreement that Phe blood levels of 360 µmol/L and less are not toxic to the brain and the clinical expert advised that in North America, the treatment aim is to maintain levels below 360 µmol/L, although as per the proposed

listing criteria suggested by Ontario PKU specialists (see Appendix 8), target Phe levels could be less than 600 µmol/L in patients aged 16 years and older.

The primary outcome in the SPARK study was dietary Phe tolerance, which was defined as the prescribed amount of dietary Phe (mg/kg/day) tolerated while maintaining mean Phe blood levels within the target range of 120 to 360 µmol/L. At baseline, the mean (SE) Phe tolerance was 35.5 (3.8) mg/kg/day in the SAP + diet arm compared with 42.8 (4.1) mg/kg/day in the diet alone arm. Phe tolerance steadily increased in the SAP + diet arm as opposed to the diet alone arm over the course of the trial. At week 26, the mean (SE) Phe tolerance was 80.6 (4.2) mg/kg/day in the SAP + diet arm and 50.1 (4.3) mg/kg/day in the diet alone arm. The difference between treatment arms at week 26 (30.5 [95% confidence interval (CI), 18.7 to 42.3]) was statistically significant. These findings must be interpreted in the context that the SPARK study included only children aged four years and younger who were already in the target Phe blood level range of 120 to 360 µmol/L with dietary therapy alone. Therefore, adherence to diet would be expected to be very good in this patient population due to vigilant control by parents and caregivers at this age. It would be useful to know if SAP increased Phe tolerance in patients who are not well controlled on diet alone. Nonetheless, although there was a statistically significant increase in dietary Phe tolerance with SAP + diet compared with diet alone in the SPARK study, this was not linked to improved quality of life nor any measure of diet liberalization. In addition, the long-term studies (summarized in Appendix 6) that were provided as new information by the manufacturer also do not provide any evidence that improved Phe tolerance is linked to diet liberalization or improved quality of life.

The effects of SAP treatment on neuropsychiatric and neurocognitive effects using various instruments were investigated in Study PKU-016, whereas in the SPARK study, only effects of treatment on neuromotor developmental milestones were reported. In PKU-016, the primary objective was to evaluate the effect of SAP treatment on ADHD symptoms in patients who were Phe responders with ADHD symptoms at baseline (i.e., n = 19 patients in each arm). This was evaluated by the change from baseline to week 13 in the ADHD-RS/ASRS total score and Inattention and Hyperactivity-Impulsivity subscale scores, where a higher score indicates greater severity of ADHD symptoms. In both treatment arms, the mean (SE) ADHD-RS/ASRS total score decreased from baseline to week 13 (i.e., –9.1 [2.2] in the SAP + diet arm and –4.9 [2.0] in the placebo + diet arm), suggesting improvement, although the MCID is unknown. In each arm, the change from baseline to week 13 was statistically significant, although the difference between arms was not statistically significant at either week 13 or week 26. For the ADHD-RS/ASRS subscale score of Inattention, in both treatment groups, the mean (SE) subscale score decreased from baseline to week 13 (i.e., –5.9 [1.4] in the SAP + diet arm and –2.5 [1.3] in the placebo + diet arm). In each arm, the change from baseline to week 13 was statistically significant and the difference between arms was also statistically significant, although at week 26, the difference between arms was no longer statistically significantly different. For the ADHD-RS/ASRS subscale score of Hyperactivity-Impulsivity, in both treatment arms, the mean (SE) subscale score decreased from baseline to week 13 (i.e., –3.3 [1.1] in the SAP + diet arm and –2.3 [1.0] in the placebo + diet arm). The change from baseline to week 13, however, was statistically significant only in the SAP + diet arm, but not in the placebo + diet arm. The difference between arms was also not statistically significant at either week 13 or week 26. These results are difficult to interpret, as it appears that, in general, ADHD symptoms improved from baseline to week 13 in all treatment groups, but that there were no differences between groups, with the exception of the Inattention subscale score. These results are also complicated by the fact that the MCIDs have not been established for this instrument in patients with PKU and that a correction factor was applied to the ASRS score, which has not been used previously in PKU or other disease areas.[9] In the SPARK study, the only measure of neurodysfunction that was

reported was the proportion of patients who were classified as either normal or abnormal with regard to neuromotor developmental milestones in four areas of assessment: fine motor, gross motor, language, and personal-social. [redacted].

Secondary end points in PKU-016 that also provide information on neuropsychiatric or developmental outcomes are the Hamilton Depression and Anxiety Rating Scales (HAM-D and HAM-A) and the Behavior Rating Inventory of Executive Function (BRIEF), which includes Global Executive Composite (GEC), Metacognition Index (MI), and Behavioral Regulation Index (BRI) scores. The mean change (SE) from baseline to week 13 in the HAM-A in Phe responders was –3.2 ([redacted]) in the SAP + diet arm and –3.6 ([redacted]) in the placebo + diet arm, both of which were statistically significant compared with baseline. A decline in the HAM-A or HAM-D score represents an improvement in symptoms. The difference between treatment arms at either week 13 or week 26, however, was not statistically significant. A similar pattern was observed for the HAM-D results in Phe responders. The mean change (SE) from baseline to week 13 in the HAM-D was –2.1 ([redacted]) in the SAP + diet arm and –2.5 ([redacted]) in the placebo + diet arm, both of which were statistically significant. The difference between treatment arms, however, was not statistically significant at either week 13 or week 26. In Phe responders younger than 18 years of age, the BRIEF-Parent was used (which was completed by parents or caregivers) and in those aged 18 years and older, the BRIEF-A was used (which was self-administered). A higher score on the BRIEF or Index scores indicates greater frequency of the behaviour. For each BRIEF assessment, results are reported separately for the three index scales (GEC, MI, and BRI). For the BRIEF-Parent results, the differences between treatments at week 13 were statistically significantly different for the GEC (i.e., –4.1 [95% CI, –7.9 to –0.3]; $P = 0.034$) and MI (i.e., –4.4 [95% CI, –8.5 to –0.2]; $P = 0.038$, but not for the BRI index scale (i.e., –3.4 [95% CI, –6.8 to 0.0]; $P = 0.053$). For the BRIEF-A results, there were no statistically significant differences between groups for any of the three index scales. One concern expressed by CEDAC in the original review by the CADTH Common Drug Review (CDR) of Kuvan was that no neuropsychological outcomes were measured in the trials included in the review. While the results of the ADHD-RS/ASRS, HAM-D, HAM-A, and BRIEF-Parent GEC and MI index scores appear to support a treatment benefit from baseline with SAP on various neuropsychiatric outcomes, the majority of differences between treatment arms were not statistically significant. It could be inferred from these findings that SAP does not appear to offer a significant additional benefit over and above a well-controlled Phe-restricted diet on neuropsychiatric outcomes, or that the short duration of the trial (26 weeks) precluded observation of any significant changes.

The second co-primary end point in PKU-016 was the proportion of patients with a rating of 1 or 2 in the CGI-I at week 13 in the population of Phe responders. The proportion of patients with this outcome was 26.3% in the placebo + diet group and 21.7% in the SAP + diet group at week 13 and the difference was not statistically significant. At week 26, the proportion of patients with this outcome in the placebo + diet arm (which included patients who crossed over from placebo to open-label SAP) was [redacted]% compared to [redacted]% in the SAP + diet arm, thus favouring the placebo + diet arm. For the secondary outcome of CGI-S (where lower scores indicate improvement), the mean [SE] reduction in scores from baseline to week 13 in both treatment arms was statistically significant (i.e., –0.6 [[redacted]] in the SAP + diet arm and –0.5 [[redacted]] in the placebo + diet arm; $P < 0.001$ for both) but the difference between groups at either week 13 or week 26 was not statistically significantly different.

Another concern expressed by CEDAC in the original CDR review of Kuvan was that no information on growth parameters was provided in the included trials. In the SPARK study, treatment differences in the change from baseline to week 26 were investigated for four different growth parameters: height, weight, BMI, and head circumference, all measured as standard deviation scores (SDS) [redacted]. Thus, although the new clinical evidence does provide some information on growth parameters, the failure to demonstrate a difference could be attributed to the short 26-week duration of the trial, or it may indicate that the addition of SAP does not affect growth in this patient population.

Harms

Most patients (≥ 75%) in both trials, regardless of treatment arm, experienced treatment-emergent adverse events (AEs); however, the frequencies of AEs were similar between treatment arms and most were mild or moderate in severity. In PKU-016, the most frequent AEs were headache, nasopharyngitis, and vomiting, whereas in the SPARK study they were pyrexia, cough, decreased amino acid level, and vomiting. There were no deaths in either trial, few patients with serious adverse events, and only one patient with a withdrawal due to an adverse event in PKU-016. The safety findings from the new clinical evidence are consistent with those in the original CDR review and support that SAP is generally safe and well tolerated. No new safety issues were identified.

Other Considerations

The current Ontario Drug Program Exceptional Access Program reimbursement criteria for Kuvan and a proposal for alternative listing criteria for Kuvan from Ontario PKU specialists are included in Appendix 8.

Conclusions

Two randomized controlled phase 3b trials met the selection criteria for inclusion in the systematic review: PKU-016 and the SPARK study. In a proportion of patients identified as Phe responders, SAP + a Phe-restricted diet either reduced (PKU-016) or maintained (SPARK) Phe blood levels within the target range. The magnitude of the reduction in Phe levels with SAP + diet was similar to that of placebo + diet, or diet alone, respectively, in the included trials. In the SPARK study, diet alone was able to maintain patients within the target Phe blood level range of 120 to 360 µmol/L, although SAP + diet was associated with a statistically significant increase in dietary Phe tolerance compared with diet alone. In neither trial was a reduction in Phe blood levels or an increase in Phe tolerance associated with diet liberalization or improved quality of life, both of which are outcomes that are important to patients. In PKU-016, improvements from baseline in various neuropsychiatric and/or neurocognitive outcomes were observed with SAP + diet, including ADHD symptoms in a subpopulation of Phe responders, but the differences between treatment arms were either not statistically significant or were inconsistent. Interpretation of the results from the included trials is complicated by a large placebo effect and lack of validation of measured outcomes or identification of MCIDs in patients with PKU. The safety and tolerability profile of SAP was consistent with that reported in the original CDR review, which suggests that SAP is generally safe and well tolerated. Overall, the new clinical evidence does not appear to reduce the uncertainty identified in the original CDR review regarding identification of a subpopulation of patients with PKU who will ultimately benefit from SAP treatment, what is considered to be a clinically meaningful response to SAP, and whether the changes observed in Phe levels are correlated with clinically important outcomes.

TABLE 1: SUMMARY OF RESULTS

	PKU-016		SPARK	
	SAP + Diet N = 98	Placebo + Diet N = 108	SAP + Diet N = 27	Diet Alone N = 29
Phe responders,[a] n (%)	61 (62)	57 (58)	NA	NA
Phe responders + ADHD symptoms at BL,[b] n (%)	19 (19)	19 (18)	NA	NA
ADHD-RS/ASRS Total Score[b]				
BL LSM (95% CI) Week 13 LSM (95% CI) LSM change from BL (95% CI) *P* value	[redacted] [redacted] –9.1 (–13.5 to –4.7) [redacted]	[redacted] [redacted] –4.9 (–8.9 to –0.9) [redacted]	NR NR NR NR	NR NR NR NR
LSM diff from placebo (95% CI); *P* value	–4.2 (–8.9 to 0.6); 0.085		NR	
Global Function Evaluation (CGI-I)[a]				
Patients with rating of 1 or 2,[c] n (%)	13 (21.7)	15 (26.3)	NR	NR
Relative risk (95% CI); *P* value	0.87 (0.46 to 1.64); 0.670			
BRIEF-Parent T Scores in Phe Responders Aged < 18 Years				
LSM diff from BL at week 13, (95% CI); *P* value:				
GEC T score MI T score BRI T score	–4.1 (–7.9 to –0.3); 0.034 –4.4 (–8.5 to –0.2); 0.038 –3.4 (–6.8 to 0.0); 0.053		NR NR NR	NR NR NR
BRIEF-A T Scores in Phe Responders Aged ≥ 18 Years				
LSM diff from BL at week 13, (95% CI); *P* value:				
GEC T score MI T score BRI T score	–1.0 (–5.5 to 3.6); 0.661 –0.5 (–5.3 to 4.2); 0.824 –1.7 (–5.8 to 2.3); 0.396		NR NR NR	NR NR NR
Mean Phe Blood Levels[a] (µmol/L)				
BL, mean (SD) Week 13 or 12,[d] mean (SD) Week 26, mean (SD)	680.2 (435.4) [redacted]	789.5 (464.9) [redacted]	[redacted]	[redacted]
Week 26 adj. txt diff (95% CI) *P* value	NR NR		–33.2 (–94.8 to 28.4) 0.290	
Mean Dietary Phe Tolerance (mg/kg/day)				
Week 2, mean (95% CI) Week 12, mean (95% CI) Week 26, mean (95% CI)	NR NR NR	NR NR NR	[redacted] 80.6 (72.3 to 88.8)	[redacted] 50.1 (41.6 to 58.6)
Week 26 adj. txt diff (95% CI) *P* value	NR NR		30.5 (18.7 to 42.3) < 0.001	

	PKU-016		SPARK	
	SAP + Diet N = 98	Placebo + Diet N = 108	SAP + Diet N = 27	Diet Alone N = 29
Harms				
Deaths, n %)	0	0	0	0
Pts with ≥ 1 AE, n (%)	79 (80.6)	81 (75.0)	27 (100.0)	27 (100.0)
Pts with ≥ 1 SAE, n (%)	0	3 (2.8)	3 (11.1)	1 (3.7)
Pts with ≥ 1 WDAE, n (%)	1 (1.0)	0	0	0

ADHD = attention-deficit/hyperactivity disorder; ADHD-RS = Attention Deficit Hyperactivity Disorder Rating Scale; adj = adjusted; AE = adverse event; ASRS = Adult Attention Deficit Hyperactivity Disorder Self-Report Scale; BL = baseline; BRI = Behavioral Regulation Index; BRIEF = Behavior Rating Inventory of Executive Function; CI = confidence interval; CGI-I = Clinical Global Impression–Improvement; diff = difference; GEC = Global Executive Composite; LSM = least squares mean; MI = Metacognition Index; NA = not applicable; NR = not reported; Phe = phenylalanine; pts = patients; SAE = serious adverse event; SAP = sapropterin; SD = standard deviation; SE = standard error; txt = treatment; WDAE = withdrawal due to adverse event.

[a] Phe responders were defined as 20% reduction in Phe blood levels from BL at 1 month following treatment with SAP 20 mg/kg/day; and were the efficacy population for CGI-I and for mean Phe levels over the duration of the study.

[b] Phe responders with ADHD symptoms at BL were the efficacy population for ADHD-RS/ASRS total score.

[c] CGI-I rating of 1 = very much improved and 2 = much improved at week 13.

[d] Results at week 13 in PKU-016 and week 12 in the SPARK study.

Source: Clinical Study Report (CSR) PKU-016,[10] CSR SPARK study.[11]

1. INTRODUCTION

1.1 Disease Prevalence and Incidence

Phenylalanine hydroxylase (PAH) deficiency, traditionally known as phenylketonuria (PKU), is an autosomal recessive inborn error of metabolism caused by mutations in the gene encoding PAH, the enzyme that converts phenylalanine (Phe) to tyrosine.[1,2] More than 800 mutations that lead to PAH deficiency have been identified, which, in turn, results in a wide range of clinical phenotypes and spectrum of disease severity.[1,2] The most severely affected individuals are those with the phenotype of "classical PKU," whose untreated Phe blood levels are typically greater than 1,200 µmol/L, whereas normal mean Phe levels are 60 µmol/L.[1] Patients with Phe levels of 900 to 1,200 µmol/L are considered to have moderate PKU and those with Phe levels of 600 to 900 µmol/L to have mild or atypical PKU.[1] If untreated, Phe levels accumulate to toxic levels in the brain, which can lead to profound neurocognitive, neuropsychiatric, and developmental problems. Early detection through newborn screening and treatment prevents the most dramatic clinical sequelae, although over time subtle intellectual and neuropsychiatric issues may manifest even with treatment.[1] In addition, patients treated from the early weeks of life with good metabolic control, but who lose control in later childhood or adulthood, may experience both reversible and irreversible neuropsychiatric consequences.[1]

Phenylketonuria occurs in approximately one in every 12,000 to 15,000 infants born in North America, accounting for approximately 300 new cases each year.[3] It is estimated that in Canada, 26 to 32 infants were born with PKU in the past year and approximately 1,250 patients with PKU are currently being actively managed across the provinces.[12] The management of PAH deficiency or PKU places a large burden on patients and their families, which can result in social isolation, reduced quality of life, and financial difficulties. Patient expectations are that sapropterin (SAP), a synthetic form of tetrahydrobiopterin (BH_4), will lower Phe levels and increase Phe tolerance so that patients are able to eat a more varied diet and be less isolated from their peers.

1.2 Standards of Therapy

With the exception of SAP, no pharmacologic therapies are approved in Canada for the treatment of PKU. The mainstay of treatment is dietary manipulation based on a Phe-restricted diet and Phe medical foods, as detailed in Appendix 7. The goals of dietary therapy are to lower Phe levels, while still providing sufficient Phe for protein formation, normal growth, and health maintenance and to prevent catabolism. The Phe-restrictive diet is expensive, unpalatable, and includes three main components: medical food (synthetic Phe-free formula containing other essential amino acids, vitamins, iron, and trace elements), natural foods (strictly "vegan-vegetarian," excluding high-protein foods such as meat, fish, chicken, bread, eggs, cheese, nuts, and certain legumes), and specially designed (and expensive) low-protein foods to give needed variety to the diet. The strict diet imposes an economic and social burden on patients and their families, which often leads to non-adherence, especially among adolescents and young adults.[4] The transition to adulthood is identified as a high-risk period for individuals with PAH deficiency due to the desire for increased independence, peer pressure, rebellious behaviour, loss to follow-up, and the high cost of medical foods.[1]

According to the most recent clinical practice guidelines put forth by the American College of Medical Genetics and Genomics (ACMG), the primary goal of therapy is to lower Phe blood levels and any interventions, including medications or combination of therapies that help to achieve that goal in an individual without other negative consequences, should be considered appropriate therapy.[1] The ACMG guidelines recommend lifelong treatment for patients with untreated Phe levels greater than 360 µmol/L,

with the goal of maintaining Phe blood levels in the range of 120 to 360 µmol/L for patients of all ages.[1] Many treatment centres in North America initiate treatment when the Phe blood level is 360 µmol/L or higher; however, the evidence regarding clinical outcome in untreated patients with Phe levels between 360 and 600 µmol/L is mixed, and published evidence of harm associated with these levels has been inconsistent.[1] Initiation of treatment should be undertaken as early as possible, preferably within the first week of life, with a goal of normalizing Phe blood levels within the first two weeks.[1] The primary treatment is dietary manipulation with frequent modification to respond to growth, life stages, concurrent illness, and comorbidities.[1] The first and only pharmacologic drug currently available for treatment of PAH deficiency is SAP; however, only approximately 25% to 50% of patients are SAP-responsive.[1] The mechanism by which residual PAH activity is enhanced by SAP is unclear, but it is speculated that BH_4 or SAP acts to improve folding and increases stability of the mutant protein; therefore, patients with mild deficiency are most likely to respond because some stable protein is required for SAP to function.[1] Nonetheless, responsive patients are identified even among those with compete PAH deficiency, so the ACMG guidelines recommend that every PAH-deficient patient be offered a trial of SAP.[1] Responsiveness is commonly assessed by starting the patient on 20 mg/kg/day of SAP and obtaining Phe blood levels at regular intervals (e.g., at 24 hours and weekly up to four weeks) with the assumption that the diet remains stable during the testing period.[1] In responders, a decline in Phe blood levels is expected although the ACMG guidelines do not specify an expected range of decline.[1] The ACMG guidelines do state that clinical judgment is required to determine what constitutes a significant or beneficial decline in Phe blood levels from baseline in an individual, but 30% is often cited as evidence of effective Phe reduction.[1] Ongoing treatment in SAP responders is typically 5 to 20 mg/kg/day of SAP in conjunction with a Phe-restricted diet.[1]

1.3 Drug

Sapropterin is a synthetic formulation of BH_4, a necessary co-factor for the PAH enzyme, which hydroxylates Phe through an oxidative reaction to form tyrosine.[5] In patients with PKU, treatment with BH_4 can activate residual PAH enzyme, improve the oxidative metabolism of Phe, and decrease Phe blood levels in some patients.[5] In patients with PKU who are responsive to BH_4, Phe blood levels decrease within 24 hours after a single administration of SAP, although the maximal effect on Phe blood levels may take a month or longer, depending upon the patient.[13]

Kuvan (sapropterin dihydrochloride) is available as 100 mg oral tablets. Kuvan is indicated, in conjunction with a Phe-restricted diet, to reduce Phe blood levels in patients with hyperphenylalaninemia (HPA) due to BH_4-responsive PKU.[5] The recommended starting dose is 10 mg/kg/day and response to therapy is determined by change in blood Phe following treatment with 10 mg/kg/day for a period of up to one month.[5] Blood Phe levels should be checked after one week of initiating treatment and periodically for up to a month. If blood Phe does not decrease from baseline, the dose may be increased weekly to a maximum of 20 mg/kg/day, with frequent monitoring of blood Phe levels over a one-month period. Patients whose blood Phe levels do not decrease after one month of treatment at 20 mg/kg/day are considered non-responders, and treatment should be discontinued in these patients. Once responsiveness to Kuvan has been established, the dosage may be adjusted within the range of 5 to 20 mg/kg/day according to response to therapy. As recommended for clinical management of PKU, blood Phe levels in patients receiving SAP should be tested one or two weeks after each dose adjustment and monitored frequently thereafter. Patients treated with SAP must continue on a Phe-restricted diet.

Indication under review
In conjunction with a Phe-restricted diet to reduce blood Phe levels in patients with hyperphenylalaninemia (HPA) due to BH_4-responsive PKU.
Reimbursement criteria requested by sponsor
Ongoing funding of sapropterin (Kuvan) for non-pregnant patients and patients actively planning pregnancy who have a diagnosis of PKU and who have demonstrated a response to the initial 6 month trial of sapropterin and who meet ALL of the following criteria: 1. Compliance with low protein diet, formulas, and treatment with sapropterin; AND 2. Has achieved a) normal sustained blood Phe levels [Greater than 120 µmol/L and less than 360 µmol/L] (At least 2 levels measured at least 1 month apart); OR b) sustained blood Phe reduction of at least 30% (At least 2 levels measured at least 1 month apart) compared to baseline if the Phe baseline level is less than 1200 µmol/L; OR c) sustained blood Phe reduction of at least 50% (At least 2 levels measured at least 1 month apart) compared to baseline if the Phe baseline level is greater than 1200 µmol/L; AND 3. Demonstrated increase of dietary protein tolerance based on targets set between the clinician and patient; **OR** 4. Clinically meaningful age-appropriate improvement in: a) neurobehavioural or neurocognitive function or impairment for patients with such impairments as determined by peer reviewed clinically validated scales; OR b) demonstrated improvement in Quality of Life using peer reviewed validated scales; AND 5. Managed by a physician specialized in metabolic/biochemical diseases

2. SUBMISSION HISTORY

In January 2011, the Canadian Expert Drug Advisory Committee (CEDAC) issued a Final Recommendation that Kuvan not be reimbursed.[6] The key reason for the recommendation was that patient details were insufficient to identify a subpopulation for which SAP may provide a significant clinical benefit that is cost-effective. Although there was sufficient evidence that SAP lowers blood Phe levels in certain patients with PKU, the submission did not provide sufficient details of how to identify patients who would benefit in a cost-effective manner. A proposed "Kuvan Starter Program" was suitable only to screen patients to identify "responders," but such response did not differentiate low response from clinically important response. In addition, starting and stopping rules, beyond the screening stage, that are linked to substantive benefit were needed. The Final CEDAC Recommendation was based on a review of Kuvan in November 2010 and a Request for Reconsideration in January 2011.[6] A Request for Advice (RfA) was also made by the participating drug plans in October 2011, which did not result in any changes to the recommendation.[7]

In making its recommendation, CEDAC considered a systematic review of two double-blind, randomized controlled trials (RCTs) in patients with BH_4-responsive PKU (PKU-003 and PKU-006), a critique of the manufacturer's pharmacoeconomic evaluation, and patient input.[6] For both RCTs, eligible patients were required to demonstrate a 30% or greater reduction in Phe blood levels following an eight-day challenge with SAP. PKU-003 enrolled 89 responder patients aged eight years and older who were not adherent to a Phe-restrictive diet who were treated with SAP 10 mg/kg/day or placebo over six weeks. PKU-006 (Part 2) enrolled 46 responder children aged four to 12 years who were adherent to a Phe-restricted diet who were treated with SAP 20 mg/kg/day or placebo over 10 weeks. The primary end point in PKU-003 was change in Phe blood level from baseline to week 6 and in PKU-006 was the Phe supplement tolerated at week 10 while maintaining Phe blood levels at less than 360 µmol/L. Results demonstrated that in PKU-003, SAP-treated patients experienced statistically significantly greater mean reductions in Phe blood levels compared with placebo at six weeks: –235.9 µmol/L versus +2.9 µmol/L, respectively. In addition, a statistically significantly greater proportion of SAP-treated patients, compared with placebo-treated patients, achieved blood Phe levels of 600 µmol/L or less (54% versus 23%) and blood Phe levels of 360 µmol/L or less (32% versus 2%). In PKU-006 (Part 2), the mean Phe supplement tolerated was 21 mg/kg per day for SAP compared with 2.9 mg/kg per day for placebo. In general, adverse events (AEs) were mild, with the most frequent AEs observed in SAP-treated patients including headache, upper respiratory tract infection, and cough.

The RfA that was made regarding the aforementioned recommendation sought to clarify the "Of Note" section of the CEDAC Final Recommendation for SAP, as the jurisdictions participating in the CADTH Common Drug Review (CDR) indicated that the statements "Starting and stopping rules…are needed. It would be advisable for the manufacturer to work with provinces to establish these requirements with respect to age, PKU classification, and specific benefits to be achieved that support the price of this product" conflicted with the "Do Not List" Recommendation and the usual role of CEDAC in providing coverage criteria recommendations. A reassessment of the clinical evidence (including new evidence available since the original CDR review), however, did not lead to a change in the recommendation. Specifically, the CDR RfA report concluded, "In the absence of evidence demonstrating benefit of sapropterin therapy on neurocognitive outcomes or diet liberalization, there is insufficient evidence to support specific starting and stopping rules for sapropterin treatment (beyond percentage reductions in blood Phe levels). There is no direct evidence to support the rules proposed by the manufacturer in response to the RfA, which are based on arbitrary percentage changes in blood Phe levels and dietary

Phe tolerance, and unspecified improvements in neurocognitive or behavioural outcomes." Despite this, several jurisdictions did subsequently proceed to develop criteria to allow access to SAP (e.g., see Appendix 8A for the criteria in Ontario), and these criteria are largely the subject of the manufacturer's current resubmission.

2.1 Basis of Resubmission

According to the manufacturer, the rationale for filing a resubmission for Kuvan is the availability of new clinical evidence. Following the CEDAC recommendation, provincial reimbursement of Kuvan has occurred in Ontario (as of February 2013) and Saskatchewan (as of September 2013).[12] The manufacturer states that listing criteria for reimbursement in these provinces were established with the aim of providing access and with the understanding that new data would be forthcoming that may affect knowledge about the effectiveness and appropriate use of SAP to treat patients with PKU.

The new clinical evidence submitted by the manufacturer includes the following studies:

- PKU-016, which is a phase 3b, double-blind RCT comparing SAP and placebo on symptoms of attention-deficit/hyperactivity disorder (ADHD) and executive and global functioning in PKU patients who are responders to SAP (i.e., defined in this study as a ≥ 20% reduction in Phe blood levels after one month of SAP therapy).[9,10]
- PKU-015, which is an ongoing, phase 3b, two-part, multi-centre, open-label, single-arm study to evaluate the safety of SAP and its effect on maintaining neurocognitive function, blood Phe levels, and growth in children aged zero to six years with PKU.[14]
- The SPARK study, which is a phase 3b, multi-centre, open-label, RCT of the efficacy, safety, and population pharmacokinetics of SAP plus a Phe-restricted diet compared with a Phe-restricted diet alone in children with PKU who are younger than four years of age.[11]
- The PKUDOS registry study is a phase 4 voluntary observational study designed to provide longitudinal safety and efficacy data for up to 15 years in patients with PKU who are (or have been) treated with SAP.[15]

Both Study PKU-016 and the SPARK study met the selection criteria for inclusion in the systematic review of the resubmission. Study PKU-015 and the PKUDOS registry study are summarized in Appendix 6.

The price of Kuvan has not changed from the price considered in the original review (i.e., $33.00 per 100 mg tablet). A new pharmacoeconomic analysis was submitted utilizing new clinical data available from the PKU-016 study to determine the cost-effectiveness of SAP in conjunction with a Phe-restricted diet compared with diet alone for patients with PKU over a lifetime horizon.

It does not appear that reimbursement criteria for Kuvan were put forth in the previous submission. In the resubmission, the manufacturer has suggested reimbursement criteria for Kuvan (as detailed in section 1.3) that support providing an initial trial of Kuvan therapy to determine a patient's response.[12] The manufacturer has also stated that the existing provincial reimbursement criteria are not supported by the metabolic geneticist community in Ontario and have been a deterrent to physicians to trial patients on SAP, as no patients have met the criteria since their inception.[12] The current provincial reimbursement criteria and a submission made directly to CADTH with proposed reimbursement criteria for SAP from a group of metabolic physician specialists from Ontario are included in Appendix 8.

3. OBJECTIVES AND METHODS

3.1 Objectives

To perform an updated systematic review of the beneficial and harmful effects of sapropterin dihydrochloride (dose: 5 to 20 mg/kg/day), when used in conjunction with a Phe-restrictive diet, to reduce blood Phe levels in patients with HPA due to BH_4-responsive phenylketonuria.

3.2 Methods

All manufacturer-provided trials considered pivotal by Health Canada were included in the systematic review. Phase 3 studies were selected for inclusion based on the selection criteria presented in Table 2.

The review protocol for the resubmission is largely unchanged from that of the original Kuvan submission. The only differences are the addition of the outcomes of health care resource utilization and proportion of responders to better inform the review. As well, examples of neurophysiologic and neurocognitive effects have been added to clarify this outcome. Any studies included in the previous CDR review were excluded from the current review.

TABLE 2: INCLUSION CRITERIA FOR THE SYSTEMATIC REVIEW

Patient Population	Patients with a diagnosis of HPA due to BH_4-responsive PKU on a Phe-restrictive diet. Subpopulations: age, Phe level at BL
Intervention	Sapropetrin dihydrochloride 5 to 20 mg/kg/day
Comparators	Phe-restrictive diet
Outcomes	**Efficacy Outcomes** – Quality of life[a] – Health care resource utilization (e.g., physician visits, hospitalization) – Change in Phe blood levels – Neurophysiologic and neurocognitive effects[a] (i.e., IQ; speech; executive and global functioning; seizures; tremors; psychological, behavioural, and social abnormalities) – Growth parameters (i.e., weight, height, z score) – Nutritional status (i.e., BMI, energy, nutrient intake) – Proportion of responders (e.g., 30% reduction in Phe level from baseline) – Change in Phe tolerance[a] **Harms Outcomes** – Mortality – AEs – SAEs – WDAEs
Study Design	Published and unpublished phase 3 RCTs

AE = adverse events; BL = baseline; BMI = body mass index; DB = double-blind; HPA = hyperphenylalaninemia; IQ = intelligence quotient; Phe = phenylalanine; PKU = phenylketonuria; RCT = randomized controlled trial; SAE = serious adverse event; WDAE = withdrawal due to adverse events.
[a] Identified as important outcomes in patient input received by CADTH for this review.

The literature search was performed by an information specialist using a peer-reviewed search strategy.

Published literature was identified by searching the following bibliographic databases: MEDLINE (1946–) with in-process records and daily updates via Ovid; Embase (1974–) via Ovid; and PubMed. The search strategy consisted of both controlled vocabulary, such as the National Library of Medicine's MeSH (Medical Subject Headings), and keywords. The main search concepts were Kuvan (sapropterin dihydrochloride) and phenylketonuria.

No methodological filters were applied to limit retrieval. Where possible, retrieval was limited to the human population. Retrieval was limited by publication year (2011 to 2016 to update a previous Kuvan submission from 2011), but not by language. Conference abstracts were excluded from the search results. See Appendix 2 for the detailed search strategies.

The updated search was completed on March 1, 2016. Regular alerts were established to update the search until the meeting of the CADTH Canadian Drug Expert Committee (CDEC) on July 20, 2016. Regular search updates were performed on databases that do not provide alert services.

Grey literature (literature that is not commercially published) was identified by searching relevant websites from the following sections of the *Grey Matters* checklist (www.cadth.ca/grey-matters):

- Health Technology Assessment Agencies
- Health Economics
- Clinical Practice Guidelines
- Drug and Device Regulatory Approvals
- Advisories and Warnings
- Drug Class Reviews
- Databases (free)
- Internet Search.

Google and other Internet search engines were used to search for additional Web-based materials. These searches were supplemented by reviewing the bibliographies of key papers and through contacts with appropriate experts. In addition, the manufacturer of the drug was contacted for information regarding unpublished studies.

Two CDR clinical reviewers independently selected studies for inclusion in the review based on titles and abstracts, according to the predetermined protocol. Full-text articles of all citations considered potentially relevant by at least one reviewer were acquired. Reviewers independently made the final selection of studies to be included in the review, and differences were resolved through discussion. Included studies are presented in Table 3; excluded studies (with reasons) are presented in Appendix 3.

4. RESULTS

4.1 Findings From the Literature

A total of two studies were identified from the literature for inclusion in the systematic review (Figure 1). The included studies are summarized in Table 3 and are described in section 4.2. There were no excluded studies, as all potentially relevant studies were included in the systematic review.

FIGURE 1: FLOW DIAGRAM FOR INCLUSION AND EXCLUSION OF STUDIES

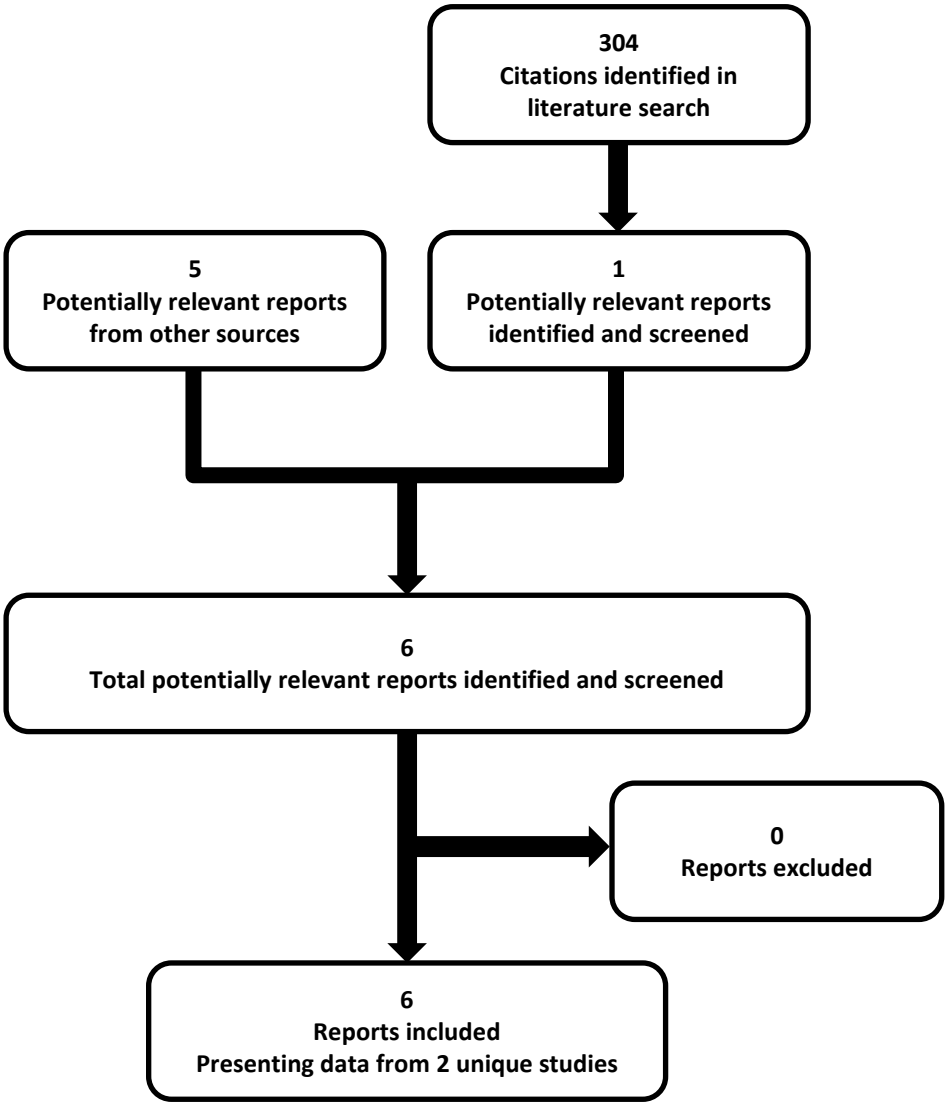


TABLE 3: DETAILS OF INCLUDED STUDIES

		PKU-016	SPARK
DESIGNS & POPULATIONS	**Study Design**	DB, PC, PG, phase 3b RCT	OL, PG, phase 3b RCT
	Locations	Canada and US	Europe, Slovakia, Turkey, United Kingdom
	Randomized (N)	206[a]	56[d]
	Inclusion Criteria	Patients aged ≥ 8 years with confirmed PKU and willingness to continue their Phe-restrictive diet unchanged for the duration of the study	Patients aged < 4 years with confirmed clinical and biochemical PKU, previous response to a BH_4 test, a defined level of dietary Phe tolerance, good dietary adherence, Phe levels between 120 and 360 µmol/L over the 4 months prior to screening
	Exclusion Criteria	SAP within past 16 weeks, initiated or adjusted ADHD, depression or anxiety medication within ≤ 8 weeks, interfering medication or concurrent disease or condition	Prior use or exposure to SAP or any registered or unregistered preparation of BH_4, previous diagnosis of BH_4 deficiency, or interfering medication or concurrent disease or condition
DRUGS	**Intervention**	SAP 20 mg/kg/day PO + Phe-restrictive diet	SAP 10 mg/kg/day PO + Phe-restrictive diet
	Comparator(s)	Placebo + Phe-restrictive diet	Phe-restrictive diet alone
DURATION	Run-in	2-week screening	42 days screening
	DB	13-week treatment	NA
	OL	13-week treatment	26 weeks
	Follow-up	30 ± 7 days	4 weeks
OUTCOMES	**Primary End Points**	– Change in ADHD-RS[b] or ASRS[c] total score (and separate Inattention and Hyperactivity-Impulsivity subscale scores) from BL to week 13 in Phe responders with ADHD symptoms – Proportion of patients with CGI-I scale rating of 1 (very much improved) or 2 (much improved) at week 13 in Phe responders with or without ADHD symptoms	Dietary Phe tolerance (i.e., defined as the daily amount of Phe in mg/kg/day ingested while maintaining average Phe levels within ≥ 120 to < 360 µmol/L)
	Other End Points	Change in HAM-D or HAM-A total score, BRIEF-GEC T score, BRIEF-BRI T score, BRIEF-MI T score, CGI-S scores from BL to week 13, CGI-I score of 1 or 2 at week 26, change in ADHD-RS, ASRS and above instrument scores from week 13 to week 26 and from BL to week 26	Change in Phe blood levels and dietary Phe tolerance from baseline, physical growth parameters (e.g., height, weight, head circumference), age-related neuromotor developmental milestones and standardized neurodevelopment test results

		PKU-016	SPARK
NOTES	**Publications**	Burton et al., 2015[9]	Unpublished

ADHD = attention-deficit/hyperactivity disorder; ADHD-RS = Attention Deficit Hyperactivity Disorder Rating Scale; ASRS = Adult Attention Deficit Hyperactivity Disorder Self-Report Scale; BH_4 = tetrahydrobiopterin; BL = baseline; BRI = Behavioral Regulation Index; BRIEF = Behavior Rating Inventory of Executive Function; CGI-I = Clinical Global Impression–Improvement; CGI-S = Clinical Global Impression–Severity of Illness; DB = double-blind; GEC = Global Executive Composite; HAM-A = Hamilton Anxiety Rating Scale; HAM-D = Hamilton Depression Rating Scale; NA = not applicable; OL = open-label; PC = placebo-controlled; PG = parallel-group; Phe = phenylalanine; PKU = phenylketonuria; PO = oral; RCT = randomized controlled trial; SAP = sapropterin.
[a] Randomization was stratified by presence of ADHD symptoms (yes or no), age at entry (< 18 and ≥ 18 years), and use of ADHD medications (yes or no).
[b] Patients aged 8 to 17 years.
[c] Patients aged 18 years and older.
[d] Randomization was stratified by age (< 12 months, 12 to < 24 months, and 24 to < 48 months).
Note: 6 additional reports were included (manufacturer's submission,[12] Clinical Study Report [CSR] PKU-016,[10] CSR SPARK study,[11] FDA Medical and Statistical reports).[16,17]
Source: Clinical Study Report (CSR) PKU-016,[10] CSR SPARK study.[11]

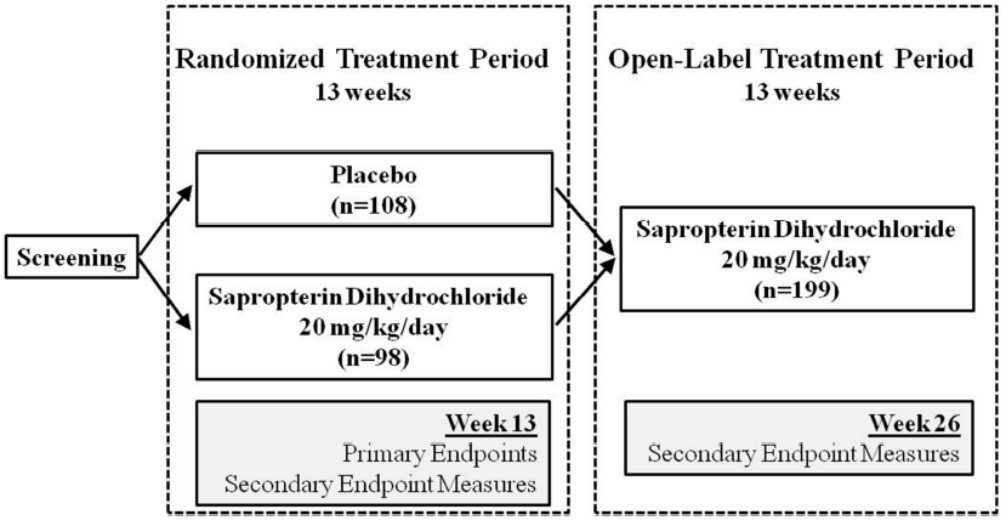
4.2 Included Studies

4.2.1 Description of Studies

The primary objective of PKU-016 was to evaluate the therapeutic effects of SAP compared with placebo on the symptoms of ADHD and global and executive function of patients with a confirmed diagnosis of PKU and symptoms of ADHD. Patients were required to have a mean decrease in Phe blood level of 20% or more from baseline to participate in the treatment periods of the trial. This was calculated as the difference between the mean of the baseline and screening values and the mean of the three lowest Phe blood levels during the first four weeks after initiating SAP.

PKU-016 (N = 206) was a double-blind, placebo-controlled, parallel-arm, phase 3b study consisting of two treatment periods of 13 weeks each: a double-blind randomized treatment period and an open-label treatment period, as illustrated in Figure 2. Eligible patients were stratified on the basis of presence of ADHD symptoms (yes or no), age (< 18 or ≥ 18 years), and use of ADHD medication (yes or no) and then randomized 1:1 by an interactive voice response system (IVRS) to receive SAP 20 mg/kg/day or placebo.

FIGURE 2: STUDY DESIGN OF STUDY PKU-016



Source: Clinical Study Report PKU-016.[10]

Patients who completed the 13-week randomized treatment period crossed over to the open-label treatment period where all patients were treated with SAP 20 mg/kg/day for an additional 13 weeks. A safety follow-up assessment was to be performed at 30 ± 7 days after completion of the open-label treatment period or after the early withdrawal visit.

For patients randomized to SAP in the randomized treatment period, assessments of Phe levels during the first four weeks of treatment (weeks 1 to 4) were used to determine the efficacy analysis population (Phe responders) based on the level of Phe reduction from baseline. For patients randomized to receive placebo in the randomized treatment period, assessments of Phe levels during the first four weeks of SAP treatment in the open-label treatment period (weeks 14 to 17) were used to determine the efficacy analysis population of Phe responders based on the level of Phe reduction from baseline.

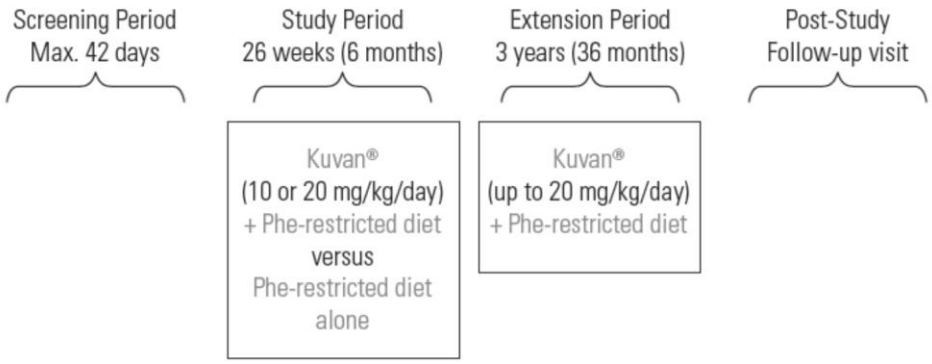
[redacted]. Blood Phe levels were monitored throughout the study for safety and efficacy.

The primary objective of the SPARK trial was to evaluate the efficacy of SAP in increasing dietary Phe tolerance, as compared with dietary therapy alone, in infants and children with PKU who were aged four years or younger. In contrast to Study PKU-016, patients were required to be Phe responders (i.e., defined as a previous response to a BH_4 test of 30% reduction from baseline in Phe levels) and to be in good control (i.e., Phe blood levels between 120 and 360 µmol/L) at study entry.

SPARK (n = 56) was an open-label, parallel-arm, phase 3b study with a duration of 26 weeks, as shown in Figure 3. Following completion of the 26-week treatment period, patients were able to enrol in an extension phase in which all patients were eligible to receive SAP plus a Phe-restricted diet. Patients were randomized 1:1 to either SAP plus a Phe-restrictive diet or a Phe-restricted diet alone. Randomization was stratified by age (i.e., < 12 months, 12 months to < 24 months, and 24 to < 48 months). According to the Clinical Study Report (CSR) for the SPARK study,[11] the trial was open-label due to the inability to conceal Phe blood tests and results from the investigators, or parents or guardians of all

enrolled patients, as the blood Phe levels and Phe response are essential in the care of infants and children with PKU.

FIGURE 3: STUDY DESIGN OF THE SPARK STUDY



Source: Clinical Study Report, SPARK study.[11]

4.2.2 Populations

a) Inclusion and Exclusion Criteria

For PKU-016, the inclusion criteria stipulated that patients with a confirmed diagnosis of PKU were eligible to participate if they were at least eight years of age and willing to continue their current diet (i.e., typical diet for the three months prior to study entry) unchanged while participating in the trial. Key exclusion criteria were prior use of SAP within 16 weeks of randomization, initiation or adjustment of medication for the treatment of ADHD, depression, or anxiety within eight weeks of randomization, or medication that could interact with SAP (e.g., methotrexate, levodopa, or any phosphodiesterase type 5 [PDE-5] inhibitor) or a concurrent disease or condition that could interfere with study participation.

For the SPARK study, the inclusion criteria stated that infants and children younger than four years of age with confirmed clinical and biochemical PKU (including at least two previous Phe blood levels ≥ 400 µmol/L obtained on two separate occasions) and previous response to a BH_4 test could be enrolled in the trial. Previous response to a BH_4 test required that the BH_4 dose was 20 mg/kg/day, the duration of the test was at least 24 hours, and that a 30% decrease in Phe blood levels occurred. Additional inclusion criteria were that the child have a defined level of dietary Phe tolerance, good adherence to dietary treatment, and maintenance of Phe blood levels within the therapeutic target range (120 to 360 µmol/L) over a four-month period prior to screening. Key exclusion criteria were previous use or exposure to SAP or any registered or unregistered preparation of BH_4, previous diagnosis of BH_4 deficiency, or medication or concurrent disease or condition that could affect the trial outcomes.

b) Baseline Characteristics

TABLE 4: SUMMARY OF BASELINE DEMOGRAPHIC AND DISEASE CHARACTERISTICS (INTENTION-TO-TREAT POPULATION)

	PKU-016		SPARK	
	SAP + Diet N = 98	Placebo + Diet N = 108	SAP + Diet N = 27	Diet Alone N = 29
Gender, n (%)				
Female Male	41 (41.8) 57 (58.2)	54 (50.0) 54 (50.0)	11 (40.7) 16 (59.3)	15 (51.7) 14 (48.3)
Race, n (%)				
White Black Asian	96 (98.0) 0 1 (1.0)	102 (94.4) 2 (1.9) 1 (0.9)	[redacted]	[redacted]
Age (Years or Months)[a]				
Mean (SD) Median (min, max) < 18 years, n (%) ≥ 18 years, n (%) < 12 months, n (%) 12 to < 24 months, n (%) 24 to < 48 months, n (%)	23.6 years (12.7) [redacted]	22.5 years (10.4) [redacted]	21.1 months (12.3) [redacted]	21.2 months (12.0) [redacted]
Standing Height (cm)				
Mean (SD) Median (min, max)	[redacted]	[redacted]	82.0 (11.3) [redacted]	82.3 (11.6) [redacted]
Weight (kg)				
Mean (SD) Median (min, max)	[redacted]	[redacted]	11.3 (3.1) [redacted]	11.3 (2.8) [redacted]
ADHD Symptoms, n (%)				
Yes No	[redacted]	[redacted]	NR NR	NR NR
ADHD Medication, n (%)[b]				
Yes No	[redacted]	[redacted]	NR NR	NR NR
Phe Blood Level (µmol/L)[c]				
Mean (SD) Median (min, max)	[redacted]	[redacted]	780.3 (480.7) [redacted])	879.9 (596.5) [redacted]
PKU Severity, n (%)				
Classic Mild Moderate	NR NR NR	NR NR NR	[redacted]	[redacted]

ADHD = attention-deficit/hyperactivity disorder; ITT = intention-to-treat; NR = not reported; Phe = phenylalanine; PKU = phenylketonuria; SAP = sapropterin; SD = standard deviation.

[a] In PKU-016, age is reported in years and in the SPARK study, age is reported in months.

[b] In PKU-016, Phe blood level is at baseline and in the SPARK study it is at diagnosis (i.e., because patients had to have Phe levels within the target range of 120 to 360 µmol/L to enter the SPARK study).

[c] In PKU-016, 5 patients did not have a baseline Phe blood level and were included as non-Phe responders for the primary efficacy population.

Source: Clinical Study Report (CSR) PKU-016,[10] CSR SPARK study.[11]

In PKU-016, the mean age of included patients was 23.1 ± 11.5 years, whereas in the SPARK study, the mean age of enrolled patients was 21.2 ± 12.1 months, as shown in Table 4. In both trials, general baseline demographic and disease characteristics were similar between treatment groups and the majority of patients (~94% to 98%) were white, reflective of the geographic location of the trials.

In PKU-016, more males (53.9%) enrolled in the study than females (46.1%), with the difference most apparent in the SAP + diet group (i.e., 58.2% males and 41.8% females), as shown in Table 4. There was also a higher proportion of patients in the age range of ≥ 18 years (n = 120) compared with < 18 years (n = 86); however, the relative proportions within the treatment groups were similar. Although the mean (SD) Phe blood levels at baseline were higher in the placebo + diet group compared with the SAP + diet group, this difference was not statistically significant or clinically relevant. Approximately one-third of patients in both treatment arms had previous ADHD symptoms and [redacted]% [redacted]% of patients were on ADHD medications at the time of study enrolment.

In PKU-016, the primary efficacy population consisted of 118 patients (49 females and 69 males) who had a Phe blood level reduction ≥ 20% from baseline within the first four weeks of SAP treatment (i.e., Phe responders). Of these, 57 (52.8%) patients were in the placebo + diet arm and 61 (62.2%) patients were in the SAP + diet arm. A total of 38 (32.2%) Phe responders (n = 19 in each treatment arm) had ADHD symptoms present at baseline. In these patients, the mean ADHD-RS/ASRS total score at baseline was 31.2 (placebo + diet) and 28.9 (SAP + diet). The contribution of the Inattention subscale score was greater than the Hyperactivity-Impulsivity subscale score at baseline. The mean Inattention subscale scores were 19.2 (placebo + diet) and 18.0 (SAP + diet), whereas the mean Hyperactivity-Impulsivity subscale scores were 12.0 (placebo + diet) and 10.9 (SAP + diet). Seven (12.3%) of the 57 Phe responders in the placebo arm and five (8.2%) of the 61 Phe responders in the SAP arm were on ADHD medications at baseline. The mean ± SD Phe blood level at baseline in the Phe responders was 789.5 ± 464.97 µmol/L (placebo + diet) and 680.2 ± 435.44 µmol/L (SAP + diet).

In the SPARK study, there were also more males than females enrolled, with the difference most apparent in the SAP + diet arm (i.e., 59.3% males and 40.7% females). Other demographic and disease characteristics were balanced and the majority of patients had either mild ([redacted]%) or moderate ([redacted]%) PKU disease severity.

4.2.3 Interventions

In PKU-016, patients received either SAP 100 mg tablets ([redacted] to the commercially available Kuvan 100 mg tablets except for the addition of a blue coating to ensure blinding) or matched placebo. All doses (20 mg/kg) were administered orally once daily after a meal. The IVRS calculated the daily dose of SAP or placebo by multiplying each patient's total body weight (in kg) at screening by 20 mg/kg and then rounding up to the nearest 100 mg to accommodate a 100 mg unit dose. Patients were required to swallow the tablets intact.

Patients who were stable on concomitant medications for the treatment of ADHD, anxiety, depression, or other neuropsychiatric illness were to continue the medications unchanged unless medically warranted. The Investigator was able to prescribe additional medications during the trial, as long as the prescribed medication was not prohibited by the protocol. Patients with mental health and/or behavioural issues were allowed to continue psychotherapy or psychosocial counselling as part of their management plan. Patients were to remain stable on all such therapies unless medically warranted.

Patients were not to deviate from their typical dietary Phe intake consumed prior to entering the trial. The only exception was for those patients who demonstrated a significant decrease in blood Phe concentration of < 120 µmol/L. If the blood Phe level decreased to < 120 µmol/L, the Investigator was notified and the patient managed according to the medical judgment of the Investigator.

[redacted].

[redacted]).

In the SPARK study, patients initially received SAP 10 mg/kg/day plus a Phe-restricted diet or a Phe-restricted diet alone. If after approximately four weeks, a patient's Phe tolerance had not increased by more than 20% compared with baseline, the SAP dose could be increased in a single step to 20 mg/kg/day. To administer the appropriate dose of SAP, patients received the corresponding number of SAP tablets (i.e., based on the dose calculated according to the patient's weight and rounded to the closest number of tablets), which were dissolved in the protocol-defined volume of water (20 to 120 mL depending upon body weight) and given to the child during breakfast. The solution was to be ingested within 15 to 20 minutes after dissolution.

For those patients randomized to a Phe-restricted diet alone during the 26-week study period who continued in the open-label extension, their starting SAP dose was 10 mg/kg/day with a maximum dose increase to 20 mg/kg/day permitted during the extension period. Patients who were randomized to the Phe-restricted diet arm received only the Phe-restricted diet and no matched placebo to SAP.

[redacted]

4.2.4 Outcomes

In PKU-016, there were two primary end points:

- Symptoms of ADHD as measured by the change in the ADHD Rating Scale (i.e., ADHD-RS for patients eight to 17 years of age) or the ASRS (i.e., ASRS for patients ≥ 18 years of age) total score and the two separate components of the total score: the Inattention and Hyperactivity-Impulsivity subscale scores from baseline to week 13 in Phe responders with symptoms of ADHD.
- Global function as measured by the proportion of patients with CGI-I scale rating of 1 (very much improved) or 2 (much improved) at week 13 in Phe responders, with or without ADHD symptoms.

a) ADHD Rating Scale and ADHD Self-Report Scale

The ADHD-RS is an 18-item scale based on Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) criteria that was developed and standardized for caregivers and educators of children to rate ADHD symptoms. The adult ASRS also includes 18 questions about the frequency of symptoms of adult ADHD. Details of both instruments are provided in Appendix 5. In PKU-016, the ADHD-RS was completed by the same parent or legal guardian for patients aged eight to 17 years, at every visit throughout the study to maintain consistency, whereas the ASRS was self-administered by patients aged 18 years and older. For both instruments, higher scores indicate worse severity of ADHD.

In PKU-016, the ADHD-RS and ASRS were combined to create a total score (ADHD-RS/ASRS) to reflect PKU-associated ADHD symptomatology in both children and adults. As the total possible scores for the ASRS and ADHD-RS are 72 and 54, respectively, due to the different rating scales, a correction factor or multiplier of 0.75 was applied to the ASRS score. Combining the scales was based on a clinical rationale that both scales include 18 questions, and each question assesses a specific DSM-IV–defined clinical symptom or behaviour. According to the CSR, the combination of both parent and self-rated scores allows analysis of all patients together, regardless of age, and increased the size of the analysis population.[10]

b) Clinical Global Impression

The Clinical Global Impression–Improvement (CGI-I) is a seven-point scale that requires the clinician to assess how much the patient's illness has improved or worsened relative to a baseline state at the beginning of the intervention. It is rated as follows: 1, very much improved; 2, much improved; 3, minimally improved; 4, no change; 5, minimally worse; 6, much worse; or 7, very much worse. In PKU-016, the CGI rater was a qualified clinician who had access to all rating scales administered at the same visit (e.g., ADHD-RS or ASRS, HAM-A, HAM-D, BRIEF, etc.). Following a review of these rating instruments, patient interviews, and discussions with other health care practitioners who had seen the patient, the CGI rater assessed the CGI score. Raters were required to be trained and certified prior to assessing patients in the study and the same rater was to conduct the CGI at all visits for a subject.

The CGI-S and CGI-I are two separate, but related instruments that assess improvement over time following the intervention (CGI-I) and the severity of disease (CGI-S). The CGI-S is also a seven-point scale that requires the clinician to rate the severity of the patient's mental illness at the time of assessment, relative to the clinician's past experience with patients who have the same diagnosis. Considering total clinical experience, a patient is assessed on severity of mental illness at the time of rating as follows: 1, normal, not at all ill; 2, borderline ill; 3, mildly ill; 4, moderately ill; 5, markedly ill; 6, severely ill; or 7, among the most extremely ill.

The secondary end points in PKU-016 were:

- Change in HAM-D total score, HAM-A total score, BRIEF Global Executive Composite (GEC) T score, BRIEF Behavior Regulation Index (BRI) T score, BRIEF Metacognition Index (MI) T score, and CGI-S scores from baseline to week 13
- Proportion of patients with CGI-I scale of 1 or 2 at week 26
- Change in ADHD-RS/ASRS total score, ADHD-RS/ASRS Inattention subscale score, ADHD-RS/ASRS Hyperactivity-Impulsivity subscale score, HAM-D total score, HAM-A total score, BRIEF-GEC T score, BRIEF-BRI T score, BRIEF-MI T score, and CGI-S scores from week 13 to week 26
- Change in ADHD-RS/ASRS total score, ADHD-RS/ASRS Inattention subscale score, ADHD-RS/ASRS Hyperactivity-Impulsivity subscale score, HAM-D total score, HAM-A total score, BRIEF-GEC T score, BRIEF-BRI T score, BRIEF-MI T score, and CGI-S score from baseline to week 26.

The HAM-A and HAM-D scores in PKU-016 were primarily analyzed for the Phe responder study population, regardless of age. The rationale was that by combining the age populations, this allowed analysis in a meaningful population size as PKU is a rare disease.

c) Hamilton Anxiety and Depression Rating Scales

The HAM-A is a clinician-administered scale designed to assess anxiety symptoms not specific to any disorder. It has 14 items, each measuring specific anxiety symptom clusters (e.g., tension, insomnia, respiratory). Each item is given a five-point score, as follows: 0, absent; 1, mild; 2, moderate; 3, severe; or 4, incapacitating. The HAM-D is a 17-item depression rating scale and nine of the items are scored on a five-point scale, as follows: 0, absence of the depressive symptom being measured; 1, doubt concerning the presence of the symptom; 2, mild symptoms; 3, moderate symptoms; or 4, severe symptoms. The remaining eight items are scored on a three-point scale as follows: 0, absence; 1, doubt on the presence of the symptom; or 2, clear presence of symptoms. The clinicians administering the HAM-A and HAM-D were required to be trained and certified prior to assessing patients and the same clinician was to conduct the HAM-A and HAM-D at all visits. Detailed interview guides were provided to the qualified clinicians to assist with administering the HAM-D. Question 3 of the HAM-D is an assessment of suicidal impulses. Every site was required to have a defined process by which a study patient at risk for suicide would be referred for expert consultation in the event it was deemed appropriate.

d) Behaviour Rating Inventory of Executive Function Scale

The BRIEF–Adult Version (BRIEF-A) is a self-reported questionnaire for patients aged 18 years and older and the Parent Form of the BRIEF (BRIEF-Parent) is a parent- or legal guardian–reported questionnaire for patients younger than 18 years. If the patient was 17 at the start of the study, but turned 18 during the course of the study, the Parent Form of the BRIEF continued to be parent- or guardian-reported for the remainder of the study. A higher score on the BRIEF or index subscales indicates greater frequency of the behaviour.

Study site personnel were to remain available for questions during administration of the BRIEF. A quiet, private, and non-distracting environment was to be provided during the tests.

e) Phe Blood Levels

In PKU-016, Phe blood levels were measured throughout the study for safety monitoring and to determine the efficacy analysis population based on the level of Phe reduction after SAP treatment. Laboratory assessments of Phe levels by blood spot testing were to be performed at screening, at baseline (unless the screening and baseline visits were combined), and at weeks 4, 8, and 13 of the randomized treatment period and at weeks 17, 21, and 26 of the open-label treatment period. Weekly home Phe level testing by blood spot test was to be performed at weeks 1, 2, and 3 during the randomized treatment period and at weeks 14, 15, and 16 during the open-label period for the purpose of determining the primary efficacy population (Phe responders).

[redacted]

f) Phe Tolerance

In the SPARK study, the primary outcome was dietary Phe tolerance after 26 weeks (six months), which was defined as the daily amount of Phe (mg/kg/day) that could be ingested in the diet while maintaining average Phe blood levels within the specified therapeutic target range (defined as ≥ 120 to < 360 µmol/L).

[redacted]. The assessment of Phe tolerance at the final visit (week 26) was used for the primary efficacy analysis, unless the patient was assessed to be not maintaining control at that visit. If this occurred, the last visit at which the patient was assessed to be maintaining overall control was used for the analysis.

Safety parameters included AEs, SAEs and deaths, vital signs, findings on physical examination and changes in clinical laboratory tests (e.g., chemistry, hematology, and urinalysis) in both included trials.

4.2.5 Statistical Analysis

In PKU-016, the planned enrolment was approximately 200 total patients with 100 patients in each treatment group in order to enrol approximately 50 patients who had a blood Phe level reduction of ≥ 20% after SAP treatment in each treatment group. At least 20 of these 50 patients in each treatment group were to have symptoms of ADHD with enrolment extended to include a minimum of 20 patients (out of the expected 50 patients with a ≥ 20% drop in Phe) in each treatment group with symptoms of ADHD. Based on a type I error rate of 0.05 (two-sided) and assuming a mean improvement in ADHD-RS/ASRS score of 13 in the SAP-treated patients, a mean improvement of 5 in the placebo-treated patients, and with a common SD of 9, a total of 20 patients in each treatment group who had a blood Phe level reduction after SAP treatment and also had symptoms of ADHD provided approximately 80% power to detect the estimated difference between SAP and placebo-treated patients.

This sample size calculation assumed that approximately 40% of the patients had a blood Phe level reduction after SAP treatment and have symptoms of ADHD, and there would be a minimum of 50 patients who had a blood Phe level reduction after SAP treatment enrolled in each treatment group. A sample size of 50 patients who had a blood Phe level reduction after SAP treatment in each group yielded approximately 80% power to detect a 30% difference in the proportion of patients with a CGI-I scale rating of 1 or 2 between the two treatment groups. The proportion of patients with a CGI-I scale rating of 1 or 2 was assumed to be 60% in the SAP-treated patients and 30% in the placebo-treated patients in this power calculation.

[redacted]. For the primary efficacy outcome (i.e., ADHD-RS or ASRS and separate Inattention and Hyperactivity-Impulsivity subscale scores), the changes from baseline to week 13 were analyzed by the analysis of covariance (ANCOVA) with treatment group, ADHD medication, age group at study entry (< 18 or ≥ 18 years), and ADHD symptoms as factors and baseline total score as a covariate. For the analysis of change in ADHD-RS/ASRS total score, the total score was rescaled by multiplying 0.75 to reduce the heterogeneity between the variances of ADHD-RS total score and ASRS total score in order to make them comparable. The stratified Wilcoxon rank-sum test (van Elteren test) was performed as supportive analyses for the change from baseline in ADHD-RS/ASRS total scores across strata (age, ADHD symptom, and ADHD medication). The proportion of patients with a CGI-I scale rating of 1 or 2 in the two treatment groups was compared by the Cochran–Mantel–Haenszel (CMH) test adjusted for age group, ADHD symptom, and ADHD medication, as appropriate. The 95% CIs were provided for the proportion by treatment group.

Statistical analysis of the secondary end points was similar to that used for the primary efficacy end point. For the within-treatment comparison, a pair-wise two-sided t test was used. There was no type I error rate adjustment for multiple tests for the secondary end points. For each treatment arm, additional exploratory analyses were performed on ADHD-RS/ASRS total score by subgroup to determine the possible interaction of subgroups with treatment using the ANCOVA model of the primary analysis. These subgroups were age (< 18 years versus ≥ 18 years), gender, whether on ADHD medication at baseline, and presence of ADHD symptoms at baseline.

In the SPARK study, the calculation of sample size for the study period was based on the assumption that at the week 26 (month 6) visit, the dietary Phe tolerance for the SAP-treated group would be 75% greater than the dietary Phe tolerance for the group treated with dietary therapy alone. a sample size of 23 patients per group would yield 80% power for testing the null hypothesis of no treatment difference between the treatment groups.

To compensate for possible dropouts, a total of 50 patients (25 per treatment group) were to be randomized to treatment. The rationale for the above estimates was based on the results of trial PKU-006 Part 2, which evaluated the effects of SAP therapy on Phe tolerance in children aged four to 12 years who were under adequate control of Phe blood levels.

The dietary Phe tolerance during the study period was to be analyzed using a repeated-measures ANCOVA on the observed records applying direct likelihood method.

The adjusted means and their 95% CIs were to be derived for each time point within each treatment group. Additionally, the adjusted treatment difference (SAP plus Phe-restricted diet and Phe-restricted diet alone) at week 26, its two-sided 95% CI, and the associated *P* value were derived.

Changes from baseline in Phe blood levels, blood pressure, physical growth parameters, and age-related neuromotor developmental milestones and standardized neurodevelopmental results were analyzed using repeated-measures ANCOVA.

a) Analysis Populations

In PKU-016, the enrolled population or intention-to treat (ITT) population included all patients randomized to treatment. The safety population included all patients who were randomized to treatment, received any study drug (either SAP or placebo), and had any safety data collected after first dose.

The primary efficacy population comprised patients who had a blood Phe level reduction of ≥ 20% from baseline within four weeks of SAP treatment (Phe responders) and had symptoms of ADHD at screening. For patients randomized to SAP treatment during the first 13 weeks, Phe levels obtained weekly during the first four weeks of treatment were used to determine the primary efficacy population based on the level of Phe level reduction after SAP treatment. For patients randomized to placebo who completed the randomized treatment period, Phe levels obtained weekly during the first four weeks of SAP treatment in the open-label treatment period (weeks 14 to 17) were used to determine the primary efficacy population based on the level of Phe level reduction from baseline after SAP treatment.

A 20% blood Phe level decrease was defined as the difference between 1) the mean of the baseline and screening values, and 2) the mean of three out of the four lowest values taken during the first four weeks after starting SAP. The lowest three values were chosen to best represent clinically stable conditions. For example, a markedly elevated Phe level observed during an acute febrile illness would thus be disregarded. Similarly, if the patient's Phe level decreased abruptly to below the locally acceptable safety threshold and supplemental dietary Phe was administered, a value obtained immediately following the administration of supplemental Phe would similarly be discarded.

In the SPARK study, the analysis populations were defined as follows:

ITT population: consisted of all patients who were randomized at the start of the study period and analyzed according to the group allocated.

Per-protocol (PP) population: consisted of those patients from the ITT population without a major protocol deviation.

Safety population: consisted of all patients who had some safety assessment data available (at least one visit in vital signs, AE, or laboratory results) in the study period and:

- who received at least one dose of SAP in the study period, or
- who were randomized to the Phe-restricted diet alone.

[redacted]

4.3 Patient Disposition

Patient disposition data are provided in Table 5 in both the PKU-016 and SPARK trials; the proportion of patients who discontinued in either treatment arm did not exceed approximately 10%. The majority of patients withdrew in PKU-016 due to loss to follow-up during the open-label treatment phase or withdrawal by the patient in the SPARK study. In the PKU-specific analysis populations for Study PKU-016, more patients in the SAP + diet arm (62%) compared with the placebo + diet arm (53%) were considered to be Phe responders (i.e., based on a 20% reduction in Phe levels from baseline following a SAP challenge). It should be noted that the Phe-responder population for the placebo + diet arm includes patients initially randomized to placebo during the randomized treatment phase who were then crossed over to SAP during the open-label treatment phase. Of the patients who were Phe responders, only about 18% to 19% also had ADHD symptoms at baseline.

TABLE 5: PATIENT DISPOSITION

	PKU-016		SPARK	
	SAP + Diet	Placebo + Diet	SAP + Diet	Diet Alone
Screened, N	NR		109	
Randomized, N	98	108	27	29
ITT, N	98	108	27 (100.0)	29 (100.0)
PP, N	[redacted]	[redacted]	23 (85.2)	19 (65.5)
Safety, N	98	108	27 (100.0)	27 (93.1)[a]
Discontinued, N (%)	3 (3.1)	8 (7.4)	2 (7.4)	3 (10.3)
Randomized treatment period	1 (1.0)	3 (2.8)	2 (7.4)	3(10.3)
Open-label treatment period	2 (2.0)	5 (4.6)		
Reason for discontinuation, N (%)				
Lost to follow-up	2 (2.0)	4 (3.7)	0	0
Withdrawal by patient	1 (1.0)	2 (1.9)	1 (3.7)	2 (6.9)
Pregnancy	0	1 (0.9)	0	0
Protocol violation	0	0	1 (3.7)	1 (3.4)
Other	0	1 (0.9)	0	0

	PKU-016		SPARK	
	SAP + Diet	Placebo + Diet	SAP + Diet	Diet Alone
PKU-Specific Analysis Populations				
Primary efficacy population (Phe responders), N (%)	61 (62.2)	57 (52.8)	NA	NA
Primary efficacy population with ADHD symptoms (Phe responders with ADHD), N (%)	19 (19.4)	19 (17.6)	NA	NA
Primary efficacy population without ADHD symptoms (Phe responders without ADHD), N (%)	42 (42.9)	38 (35.2)	NA	NA

ADHD = attention-deficit/hyperactivity disorder; ITT = intention-to-treat; NR = not reported; Phe = phenylalanine; PKU = phenylketonuria; PP = per-protocol; SAP = sapropterin.
[a] Two patients withdrew consent (family did not want to continue) after day 1 and are not included in the safety population.
Source: Clinical Study Report (CSR) PKU-0016,[10] Burton et al. 2015,[9] CSR SPARK study.[11]

4.4 Exposure to Study Treatments

In PKU-016, exposure to study drug in the ITT population (safety population) was similar between patients in each treatment group and each period, as detailed in Table 6. In addition, the mean daily dose was also similar between treatment arms and study periods. The mean ± SD daily dose on a mg/kg basis adjusted by planned exposure days was 18.8 ± 3.7 mg/kg/day for patients in the placebo group and 19.2 ± 2.5 mg/kg/day for patients in the SAP treatment group.

TABLE 6: EXPOSURE TO STUDY DRUG (INTENTION-TO-TREAT POPULATION)

	PKU-016			
	Randomized Treatment (Week 0 to Week 13)		Open-Label Treatment (Week 13 to Week 26)	
	SAP + Diet [redacted]	Placebo + Diet [redacted]	SAP + Diet [redacted]	Placebo to SAP + Diet [redacted]
Total duration of exposure (days)				
Mean (SD) Median (min, max)	[redacted]	[redacted]	[redacted]	[redacted]
Total duration of exposure, n (%)				
< 46 days 46 to 91 days > 91 days	[redacted]	[redacted]	[redacted]	[redacted]
Amount of drug taken (mg)				
Mean (SD) Median (min, max)	[redacted]	[redacted]	[redacted]	[redacted]
Average daily dose (mg/day)				
Mean (SD) Median (min, max)	[redacted]	[redacted]	[redacted]	[redacted]

SAP = sapropterin; SD = standard deviation.
Note: Treatment exposure was not analyzed in the Spark study, but at the end of the study period, 25 (92.6%) of patients were on SAP 10 mg/kg/day + diet and 2 (7.4%) of patients were on SAP 20 mg/kg/day + diet.
Source: Clinical Study Report PKU-0016.[10]

4.5 Critical Appraisal

4.5.1 Internal Validity

The methods used for randomization and allocation concealment in PKU-016 were appropriate (i.e., IVRS and used of matched placebo tablets); however, in the SPARK study, the lack of a matched placebo to SAP may have introduced performance bias resulting from potential confounding factors associated with the administration of SAP.

A large placebo effect was observed in both the PKU-016 and the SPARK study for various outcomes. It is possible that more frequent clinic visits may have increased patient and provider attention to dietary treatment, which, in turn, contributed to improved Phe blood levels in both treatment groups. It is not uncommon for a large placebo effect to be observed in the evaluation of mental health outcomes (e.g., ADHD symptoms and executive functioning were assessed using parent-reported and self-reported measures, which could differ from assessments made by a health care professional). In addition, self-reporting of ADHD symptoms over time could have been affected by the patient's improved ability to focus and monitor themselves over the course of treatment.[9] It may have been advisable to also have an independent rater provide collateral information in this regard for self-reported outcomes.[9] In addition, patients with mental health and/or behavioural issues were allowed to continue psychotherapy or psychosocial counselling as part of their management plan; therefore, such interventions could have contributed to the observed results.

Baseline demographic and disease characteristics were generally balanced between treatment arms in each trial, with the possible exception of more males in the SAP treatment arms of both trials and a greater proportion of older patients (aged ≥ 18 years) compared with < 18 years of age in PKU-016. It was not possible to compare patient populations between the two trials due to study design differences such as the age of enrolled patients (i.e., mean age of enrolled patients in PKU-016 was approximately 23 years compared with 21 months in the SPARK study), which precluded making comparisons.

The use of a population after the conclusion of the actual primary end point (i.e., week 13) to select the control groups may have introduced bias, because the responses of the treatment and control group are being determined at two different time points. This is based on the possibility that defining one group of responders during the double-blind phase and the other group during the open-label phase could potentially result in there being differences between the two subsets of patients, such as differences related to patient behaviour (e.g., adherence to diet).

In general, the statistical methods and analyses appeared to be appropriate; however, there was no control for multiplicity in the analyses of the two primary end points in PKU-016 or for the analyses of the secondary end points in both trials. Therefore, the risk of type 1 error is very high.

The small patient numbers in the primary efficacy populations (e.g., 19 patients in each treatment arm for the primary efficacy analysis population of Phe responders with ADHD symptoms at baseline in PKU-016) are expected to have had only limited statistical power, although based on the manufacturer's sample size calculations detailed in section 4.2.5, PKU-016 should have had approximately 80% power to detect treatment differences.

In PKU-016, following the randomized treatment period, patients who were initially randomized to placebo + diet crossed over to open-label SAP + diet after week 13 in the open-label treatment period. Thus, comparisons of results at week 26 are between patients who were originally randomized to SAP + diet in the randomized treatment period and continued on this regimen in the open-label

treatment period with patients who crossed over to open-label SAP + diet, thus complicating the interpretation of the results after week 13. Furthermore, because the analysis population is not the population that was randomized, the benefits of randomization are lost, thereby subjecting the results to the limitations commonly observed with non-randomized studies. For instance, any differences between the treatment and control groups may be driven by potential differences in patient characteristics as the benefit of randomization has been removed. Furthermore, interpreting patient-reported outcomes for the open-label phase is particularly problematic, as these outcomes are more susceptible to bias in open-label studies.

It does not appear that any of the outcomes used to assess ADHD symptoms or neurocognitive or psychological effects of treatment in the trials (with the exception of the ADHD-RS/ASRS Inattention subscale score) have been validated in patients with PKU or the minimal clinically important differences (MCIDs) for these outcomes established, which complicates interpretation of the results. Furthermore, for the primary end point analysis in the PKU-016 trial, the ADHD-RS (parent-reported in patients younger than 18 years) and ASRS (self-reported in patients aged 18 years and older) were combined to derive an overall score and the ASRS was corrected by multiplying by 0.75. The appropriateness of doing so is questionable, especially because it is a new approach that has not been used previously in PKU or in other disease areas.[9]

4.5.2 External Validity

The PKU-016 trial included clinical sites in Canada, whereas the SPARK study was conducted primarily in Europe. Nonetheless, based on the inclusion criteria and given the rare nature of PAH deficiency PKU, it is expected that the patient populations in the included trials would be similar to the target treatment population in Canada.

The small number of patients in the primary efficacy populations in the trials affects the generalizability of results to a broader patient population and precludes identification of a subpopulation of patients who may benefit from SAP treatment. In PKU-016, the objective of the trial was to assess the impact of SAP + diet on ADHD symptoms in patients with PKU who are Phe responders; however, only approximately 30% of all enrolled patients had ADHD symptoms at baseline and only 8% to 9% were on ADHD medication at study entry.

Phe responders were defined differently in the two trials (i.e., 20% versus 30% reduction from baseline in Phe levels following SAP challenge, in PKU-016 and the SPARK study, respectively). In addition, patients in the SPARK study were required to be within the target Phe range of 120 to 360 µmol/L at study entry as opposed to the patients entering PKU-016 who were not well controlled by diet alone. According to the clinical expert consulted on this review, the per cent reduction thresholds for Phe blood levels are arbitrary, so the clinical significance of the threshold difference is uncertain.

The duration of the trials (26 weeks) is insufficient to appropriately assess the efficacy and safety of a medication intended for chronic, potentially lifelong administration.

4.6 Efficacy

Only those efficacy outcomes identified in the review protocol are reported below (section 3.2, Table 2). See Appendix 4 for detailed efficacy data.

4.6.1 Change in Phe Levels

In Study PKU-016, in the population of Phe responders, the mean (SD) Phe blood levels at baseline were 680.2 (435.44) µmol/L in the SAP + diet arm and 789.5 (464.97) µmol/L in the placebo + diet arm (Table 17). At week 13, levels in the SAP + diet arm decreased by approximately 30% from baseline to [redacted] ([redacted]) µmol/L and remained largely unchanged in the placebo + diet arm (i.e., [redacted] [[redacted]] µmol/L). After week 13, levels remained relatively stable in the SAP + diet arm, but decreased in the placebo + diet arm as these patients crossed over to receive open-label SAP + diet during the open-label treatment period. At week 26, mean (SD) Phe levels were similar between the two groups: [redacted] ([redacted]) µmol/L in the SAP + diet arm and [redacted] ([redacted]) µmol/L in the placebo + diet arm. No statistical comparisons were conducted between the treatment groups.

In contrast to Study PKU-016, patients enrolled in the SPARK study were required to be within the Phe blood level target range of 120 to 360 µmol/L at study entry. The mean (SD) baseline values were [redacted] ([redacted]) µmol/L in the SAP + diet arm and [redacted] ([redacted]) µmol/L in the diet alone arm (Table 18). Phe blood levels remained relatively constant in both treatment arms throughout the duration of the study. At week 26, the mean (SE) change from baseline was –10.1 ([redacted]) µmol/L in the SAP + diet arm and 23.1 ([redacted]) µmol/L in the diet alone arm. The LS or adjusted mean difference between the treatment arms was not statistically significant (i.e., –33.2 [95% CI, –94.8 to 28.4]; *P* = 0.290). In the SPARK study, nine patients (33.3%) in the SAP + diet arm maintained Phe blood levels in the target range (120 to 360 µmol/L) throughout the study compared with three patients (10.3%) in the diet alone arm (Table 19).

4.6.2 Neuropsychiatric and Neurocognitive Effects

The effects of SAP treatment on neuropsychiatric and neurocognitive effects using various different instruments were investigated in Study PKU-016, whereas in the SPARK study, only effects of treatment on neuromotor developmental milestones were reported. The primary objective of PKU-016 was to evaluate the effect of SAP treatment on ADHD symptoms in patients who were Phe responders with ADHD symptoms at baseline (i.e., n = 19 patients in each treatment arm). The proportion of patients in the Phe responder population with CGI-I ratings of 1 (very much improved) or 2 (much improved) was the second primary end point in PKU-016. Other instruments included as secondary end points were the CGI-S, HAM-A, HAM-D, and the GEC, MI, and BRI index T scores of the BRIEF rating scale. In the SPARK study, neuromotor status assessment was performed using the standardized Bayley-III Scales of Infant and Toddler Development for patients younger than 3.5 years of age and the Wechsler Preschool and Primary Scale of Intelligence for patients between 3.5 and 4 years of age.

The first primary end point in PKU-016 was the change from baseline to week 13 in the ADHD-RS/ASRS total score and a higher score on either rating scale indicates greater severity of ADHD symptoms (Table 8). In both treatment groups, the mean (SE) ADHD-RS/ASRS total score decreased from baseline to week 13 (i.e., –9.1 [2.2] in the SAP + diet arm and –4.9 [2.0] in the placebo + diet arm), suggesting improvement, although the MCID is unknown. In each arm, the change from baseline to week 13 was statistically significant ([redacted]), although the difference between arms was not (i.e., –4.2 [95% CI, –8.9 to 0.6]; *P* = 0.085). At week 26, the difference between arms was also not statistically significantly different ([redacted]).

For the ADHD-RS/ASRS subscale score of Inattention, in both treatment arms, the mean (SE) subscale score decreased from baseline to week 13 (i.e., –5.9 [1.4] in the SAP + diet arm and –2.5 [1.3] in the placebo + diet arm) (Table 9). In each arm, the change from baseline to week 13 was statistically significant ([redacted]) and the difference between arms was also statistically significant (i.e., –3.4 [95% CI, –6.6 to –0.2]; *P* = 0.036), although the MCID is unknown. At week 26, the

difference between arms was no longer statistically significantly different ([redacted]).

For the ADHD-RS/ASRS subscale score of Hyperactivity-Impulsivity, in both treatment arms, the mean (SE) subscale score decreased from baseline to week 13 (i.e., –3.3 [1.1] in the SAP + diet arm and –2.3 [1.0] in the placebo + diet arm) (Table 10). The change from baseline to week 13, however, was statistically significant only in the SAP + diet arm ([redacted]), but not in the placebo + diet arm ([redacted]). The difference between arms was also not statistically significant at either week 13 (i.e., –1.0 [95% CI, –3.4 to 1.4]; *P* = 0.396) or week 26 ([redacted]).

The second primary end point in PKU-016 was the proportion of patients with a rating of 1 or 2 in the CGI-I at week 13 in the population of Phe responders (Table 11). The proportion of patients with this outcome was 26.3% in the placebo + diet group and 21.7% in the SAP + diet group at week 13 and the difference was not statistically significantly different (i.e., 0.87 [95% CI, 0.46 to 1.64]; *P* = 0.670). At week 26, however, the proportion of patients with this outcome in the placebo + diet arm (which included patients who crossed over from placebo to open-label SAP) was [redacted]% compared to [redacted]% in the SAP + diet arm. The treatment difference was statistically significant in favour of the placebo + diet arm ([redacted]). For the secondary outcome of CGI-S (where lower scores indicate improvement), the mean [SE] reduction in scores from baseline to week 13 in both treatment arms was statistically significant (i.e., –0.6 [0.2] in the SAP + diet arm and –0.5 [[redacted]] in the placebo + diet arm; [redacted]); however, the difference between arms was not statistically significantly different at week 13 or week 26 (Table 12).

In PKU-016, the mean change (SE) from baseline to week 13 in the HAM-A in Phe responders was –3.2 ([redacted]) in the SAP + diet arm and –3.6 ([redacted]) in the placebo + diet arm, both of which were statistically significant ([redacted]) (Table 13). A decline in the HAM-A or HAM-D score represents an improvement in symptoms. The difference between treatment arms, however, was not statistically significant (i.e., 0.4 [95% CI, –1.5 to 2.3]; *P* = 0.669). Similarly, the treatment difference at week 26 was also not statistically significant (i.e., –0.5 [95% CI, –2.4 to 1.4); *P* = 0.590). A similar pattern was observed for the HAM-D results in Phe responders. The mean change (SE) from baseline to week 13 in the HAM-D was –2.1 ([redacted]) in the SAP + diet arm and –2.5 ([redacted]) in the placebo + diet arm, both of which were statistically significant ([redacted]) (Table 14). The difference between treatment arms, however, was not statistically significant (i.e., 0.4 [95% CI, –1.1 to 1.9]; *P* = 0.588). Similarly, the treatment difference at week 26 was also not statistically significant ([redacted]).

In Study PKU-016, in Phe responders younger than 18 years of age, the BRIEF-Parent was used, which was completed by parents (Table 15). In those aged 18 years and older, the BRIEF-A was used, which was self-administered (Table 16). For each BRIEF assessment, results are reported separately for the three index scales (GEC, MI, and BRI). For the BRIEF-Parent results (i.e., patients younger than 18 years), the differences between treatments at week 13 were statistically significantly different for the GEC (i.e., –4.1 [95% CI, –7.9 to –0.3]; *P* = 0.034) and MI (i.e., –4.4 [95% CI, –8.5 to –0.2]; *P* = 0.038, but not for the BRI index scale (i.e., –3.4 [95% CI, –6.8 to 0.0]; *P* = 0.053), although the MCID is unknown. For the BRIEF-A results (i.e., patients ≥ 18 years), there were no statistically significant differences between groups for any of the three index scales, GEC, MI, or BRI.

In the SPARK study, the only measure of neurodysfunction that was reported was the proportion of patients who were classified as either normal or abnormal with regard to neuromotor developmental milestones in four areas of assessment: fine motor, gross motor, language, and personal-social (Table 23).

In all four areas, the majority of children were classified as normal and there were no statistically significant differences found between the SAP + diet and diet alone arm for any of the areas of assessment.

4.6.3 Growth Parameters

In the SPARK study, treatment differences in the change from baseline to week 26 were investigated for four different growth parameters: height, weight, BMI, and head circumference, all measured as standard deviation scores (SDS), as detailed in Table 22. For all four growth parameters, there were no statistically significant differences observed between the SAP + diet and diet alone arm.

4.6.4 Proportion of Responders

The proportion of Phe responders was not an outcome in either of the included trials; however, in PKU-016, those patients who demonstrated a 20% reduction from baseline in Phe blood levels following treatment with SAP 20 mg/kg/day for up to one month comprised the Phe responder population. This was the primary efficacy population for the co-primary end point of CGI-I, whereas those patients in the Phe responder population who also had ADHD symptoms at baseline comprised the primary efficacy population for the other co-primary end point of ADHD-RS/ASRS. Of the total patients randomized in PKU-016 (n = 206), 118 patients (57.3%) were Phe responders. Of these, 61 patients (62.2%) in the SAP + diet arm and 57 patients (52.8%) in the placebo + diet arm were Phe responders. In the SPARK study, this outcome was not applicable as patients entering the trial were required to be BH_4 responders.

4.6.5 Change in Phe Tolerance

The primary outcome in the SPARK study was dietary Phe tolerance, which was defined as the prescribed amount of dietary Phe (mg/kg/day) tolerated while maintaining mean Phe blood levels within the target range of 120 to 360 µmol/L. At baseline, the mean (SE) Phe tolerance was 35.5 (3.8) mg/kg/day in the SAP + diet arm compared with 42.8 (4.1) mg/kg/day in the diet alone arm (Table 20). Phe tolerance steadily increased in the SAP + diet arm as opposed to the diet alone arm over the course of the trial. At week 26, the mean (SE) Phe tolerance was 80.6 (4.2) mg/kg/day in the SAP + diet arm and 50.1 (4.3) mg/kg/day in the diet alone arm. The difference between treatment arms at week 26 was statistically significant (i.e., 30.5 [95% CI, 18.7 to 42.3]; $P < 0.001$). Mean change in dietary Phe tolerance from baseline to week 26 was also statistically significant within each treatment arm (i.e., 36.9 [95% CI, 26.1 to 47.7]; $P < 0.001$ in the SAP + diet arm and 13.1 [95% CI, 5.4 to 20.9]; $P = 0.002$ in the diet alone arm) (Table 21).

There were no data reported in the included trials for the following outcomes identified in the review protocol: quality of life, health care resource utilization, and nutritional status.

4.7 Harms

Only those harms identified in the review protocol are reported below (see section 2.2.1, Protocol). See Appendix 4 for detailed harms data.

4.7.1 Adverse Events

The majority of patients in both trials (≥ 75%), regardless of treatment arm, experienced treatment-emergent AEs, as detailed in Table 7. In PKU-016, the frequency of AEs was similar in the randomized treatment arm and the open-label treatment arm. Overall, the most frequent AEs were headache, nasopharyngitis, and vomiting in Study PKU-016 and pyrexia, cough, decreased amino acid level, and vomiting in the SPARK study.

TABLE 7: HARMS (SAFETY POPULATION)

	PKU-016				SPARK	
	Randomized Treatment		Open-Label Treatment			
	SAP + Diet N = 98	Placebo + Diet N = 108	SAP + Diet N = 95	Placebo to SAP + Diet N = 104	SAP + Diet N = 27	Diet Alone N = 27
Deaths						
Deaths, n (%)	0	0	0	0	0	0
AEs						
Patients with ≥ 1 AE, n (%)	79 (80.6)	81 (75.0)	71 (74.7)	79 (76.0)	27 (100.0)	27 (100.0)
Most frequent AEs (≥ 10% in any group), n (%)						
Headache	25 (25.5)	28 (25.9)	17 (17.9)	16 (15.4)	[redacted]	[redacted]
Nasopharyngitis	11 (11.2)	9 (8.3)	11 (11.6)	12 (11.5)	13 (48.1)	11 (40.7)
Vomiting	4 (4.1)	14 (13.0)	3 (3.2)	12 (11.5)	10 (37.0)	9 (33.3)
Nasal congestion	7 (7.1)	11 (10.2)	12 (12.6)	4 (3.8)	0	2 (7.4)
Diarrhea	10 (10.2)	4 (3.7)	4 (4.2)	8 (7.7)	9 (33.3)	6 (22.2)
Oropharyngeal pain	6 (6.1)	10 (9.3)	11 (11.6)	11 (10.6)	[redacted]	[redacted]
Rhinitis	[redacted]	[redacted]	[redacted]	[redacted]	8 (29.6)	6 (22.2)
Pharyngitis	[redacted]	[redacted]	[redacted]	[redacted]	7 (25.9)	3 (11.1)
Upper respiratory tract infection	4 (4.1)	7 (6.5)	3 (3.2)	10 (9.6)	1 (3.7)	6 (22.2)
Gastroenteritis	[redacted]	[redacted]	[redacted]	[redacted]	4 (14.8)	3 (11.1)
Otitis media	[redacted]	[redacted]	[redacted]	[redacted]	2 (7.4)	3 (11.1)
Pyrexia	1 (1.0)	5 (4.6)	7 (7.4)	5 (4.8)	17 (63.0)	18 (66.7)
Abdominal pain	3 (3.1)	2 (1.9)	0	4 (3.8)	3 (11.1)	2 (7.4)
Stomatitis	[redacted]	[redacted]	[redacted]	[redacted]	3 (11.1)	1 (3.7)
Cough	7 (7.1)	8 (7.4)	8 (8.4)	8 (7.7)	13 (48.1)	13 (48.1)
Rhinorrhea	[redacted]	[redacted]	[redacted]	[redacted]	2 (7.4)	8 (29.6)
Amino acid level, decreased	[redacted]	[redacted]	[redacted]	[redacted]	10 (37.0)	9 (33.3)
Amino acid level, increased	0	1 (0.9)	0	0	1 (3.7)	4 (14.8)
Rash	[redacted]	[redacted]	[redacted]	[redacted]	6 (22.2)	2 (7.4)
Fall	[redacted]	[redacted]	[redacted]	[redacted]	2 (7.4)	3 (11.1)
Conjunctivitis	[redacted]	[redacted]	[redacted]	[redacted]	2 (7.4)	3 (11.1)
Ear pain	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SAEs						
Patients with ≥ 1 SAE, n (%)[b]	0	3 (2.8)	0	2 (1.9)	3 (11.1)	1 (3.7)
SAEs occurring in ≥ 1 patient						
Necrotizing fasciitis	-	1 (0.9)	-	-	-	-
Concussion	-	1 (0.9)	-	-	-	-
Amino acid level, increased	-	1 (0.9)	-	-	-	-
Animal bite	-	-	-	1 (0.9)	-	-
Petit mal epilepsy	-	-	-	1 (0.9)	-	-
Gastroenteritis	-	-	-	-	1 (3.7)	-
Stomatitis	-	-	-	-	1 (3.7)	-

	PKU-016				SPARK	
	Randomized Treatment		Open-Label Treatment			
	SAP + Diet N = 98	Placebo + Diet N = 108	SAP + Diet N = 95	Placebo to SAP + Diet N = 104	SAP + Diet N = 27	Diet Alone N = 27
SAP overdose	-	-	-	-	1 (3.7)	-
Bronchopneumonia	-	-	-	-	-	1 (3.7)
WDAEs						
Patients with ≥ 1 WDAE, n (%)	1 (1.0)	0	0	0	0	0
WDAEs occurring in ≥ 1 patient						
Increased heart rate	1 (1.0)	-	-	-	-	-

AE = adverse event; NA = not applicable (i.e., did not occur in specified number or proportion of patients); SAE = serious adverse event; SAP = sapropterin; URTI = upper respiratory tract infection; WDAE = withdrawal due to adverse event.
Source: Clinical Study Report (CSR) PKU-016,[10] CSR SPARK study.[11]

4.7.2 Serious Adverse Events

There were few patients with SAEs in either trial, as shown in Table 7. No one type of SAE occurred more frequently than in one patient in either of the trials.

4.7.3 Withdrawals Due to Adverse Events

There was only one patient with WDAE in Study PKU-016 who withdrew due to increased heart rate.

4.7.4 Mortality

There were no deaths reported in either trial.

5. DISCUSSION

5.1 Summary of Available Evidence

5.1.1 Evidence Reviewed Previously

The original CDR systematic review of Kuvan included two double-blind RCTs: PKU-003 (n = 96) in patients aged eight years and older with PKU who were not adherent to a strict low-Phe diet and PKU-006 [Part 2] (n = 46) in children four to 12 years of age with PKU who were adherent to a low-Phe diet.[6] PKU-003 compared SAP 10 mg/kg/day with placebo over six weeks and PKU-006 [Part 2] compared SAP 20 mg/kg/day and placebo over 10 weeks. For both RCTs, eligible patients were required to demonstrate a 30% or greater reduction in Phe blood levels (Phe responders) following an eight-day challenge with SAP 10 mg/kg/day (PKU-003) or SAP 20 mg/kg/day (PKU-006). The primary outcome in PKU-003 was change in Phe blood level from baseline to week 6, and in PKU-006, the primary outcome was the amount of Phe supplement tolerated at week 10 while maintaining Phe blood levels at less than 360 µmol/L. Neither trial included validated measures of neuropsychological performance, quality of life, growth, or diet liberalization. In PKU-003, SAP-treated patients experienced statistically significantly greater mean reductions in Phe blood levels compared with placebo at six weeks: –235.9 µmol/L versus +2.9 µmol/L, respectively. In addition, a statistically significantly greater proportion of SAP-treated patients, compared with placebo-treated patients, achieved blood Phe levels of 600 µmol/L or less (54% versus 23%) and blood Phe levels of 360 µmol/L or less (32% versus 2%). In PKU-006 (Part 2), the mean Phe supplement tolerated was 21 mg/kg per day for SAP compared with 2.9 mg/kg per day for placebo. In general, AEs were mild, with the most frequent AEs observed in SAP-treated patients including headache, upper respiratory tract infection, and cough.

The original CDR review concluded that in the short term, in a subset of patients with HPA due to BH_4-responsive PKU who respond to treatment, SAP statistically significantly lowers Phe blood levels compared with placebo in patients on a Phe-restricted diet as well as in those non-adherent to diet. Despite this, uncertainty of the clinical relevance of the 30% responder threshold for reduction in Phe blood levels, short duration of the trials, and lack of reporting of clinically meaningful outcomes (e.g., growth, nutritional status, intelligence, neuropsychological performance, and quality of life) were identified as key limitations. Outstanding questions were which PKU patients would benefit most from SAP therapy, whether SAP-responder patients could safely increase dietary Phe intake and maintain blood Phe levels in acceptable ranges, and what are the long-term safety and cognitive outcomes associated with SAP? Although CEDAC found sufficient evidence that SAP lowers Phe blood levels in certain patients with PKU, following its review, it was concluded that the submission did not provide sufficient details of how to identify patients who benefit from SAP in a cost-effective manner.

5.1.2 New Evidence Identified in the Current Review

Two phase 3b RCTs met the selection criteria for inclusion in the current systematic review. PKU-016 (N = 206) was a double-blind, placebo-controlled, parallel-arm trial in patients aged eight years and older with PKU on a Phe-restricted diet. PKU-016 comprised two 13-week treatment periods: a double-blind randomized treatment period comparing SAP 20 mg/kg/day and placebo (both in conjunction with a Phe-restricted diet) and an open-label treatment period in which patients randomized to placebo in the first treatment period crossed over to open-label SAP, while maintaining a Phe-restricted diet. The SPARK study (n = 56) was an open-label, parallel-arm trial in patients younger than four years with PKU on a Phe-restricted diet who had Phe blood levels within the target range of 120 to 360 µmol/L. The SPARK study compared SAP 10 mg/kg/day plus a Phe-restricted diet with a Phe-restricted diet alone over 26 weeks. There were two co-primary outcomes in PKU-016, which were the change from baseline

to week 13 in the ADHD-RS or ASRS total score (and separate Inattention and Hyperactivity-Impulsivity subscale scores) in Phe responders with ADHD symptoms and the proportion of patients with a CGI-I scale rating score of 1 (very much improved) or 2 (much improved) at week 13 in Phe responders, with or without ADHD symptoms. The primary outcome in the SPARK trial was the change in dietary Phe tolerance from baseline at week 26. Key limitations are the lack of a matched placebo in the SPARK study, the small patient numbers and relatively short duration of the trials, lack of validation of measured outcomes in patients with PKU, lack of control for multiplicity in the statistical analyses, and the large placebo effect observed in both trials.

5.2 Interpretation of Results

5.2.1 Efficacy

The key reason that led to the negative recommendation by CEDAC following the original review of Kuvan was that patient details were insufficient to identify a subpopulation for which SAP may provide a significant clinical benefit that is cost-effective. Thus, the interpretation of the new clinical evidence provided in the resubmission requires careful consideration of whether or not the new information reduces uncertainty regarding which patients will ultimately benefit from SAP treatment and whether or not the outstanding questions arising from the original review have been addressed. Although sufficient evidence that SAP lowers blood Phe levels in certain patients with PKU was demonstrated in the original review of Kuvan, the clinical significance of the magnitude of the reduction and uncertainty regarding whether the reduction in Phe levels are associated with clinically meaningful outcomes are important considerations in interpreting the new clinical evidence.

Both the PKU-016 and SPARK trials included in the current review were associated with various limitations, and some common to both trials are the large placebo effect, short duration and small number of included patients, and the lack of validation of measured outcomes and MCIDs in patients with PKU. While the large placebo effect (or the effect of Phe-restricted diet alone) may, in part, be attributed to participation in a clinical trial and increased attention to dietary therapy, it underscores the fact that dietary therapy is an effective intervention for patients with PKU. In light of the variability in patient response to SAP and the uncertainty of the clinical significance of the magnitude of Phe blood level reductions provided by SAP, a Phe-restricted diet remains a key component of ongoing therapy for PKU. This is reiterated by the Health Canada–approved indication for Kuvan, which states that SAP should be used only in conjunction with a Phe-restricted diet.[5] The relatively short duration of the included trials does not address the outstanding question regarding long-term safety and cognitive outcomes in the original CDR review. It is noted, however, that additional new evidence provided in the resubmission did include two registry studies[15,18] and an ongoing, open-label, single-arm observational study[14] of long-term SAP treatment, which have been summarized in Appendix 6. Despite the limitations of these studies (discussed in Appendix 6), it appears that SAP is associated with favourable long-term safety and tolerability and that with uninterrupted use, it is effective in sustaining reductions in Phe blood levels and increasing dietary Phe tolerance in Phe responders. The small number of patients included in the efficacy analyses populations of the included trials in the current review affect the generalizability of the results to a broader patient population and preclude identification of a subpopulation that may benefit from SAP treatment, which was the main reason for the negative recommendation following the original CDR review. Although the primary objective of PKU-016 was to assess the efficacy and safety of SAP in patients with PKU and ADHD symptoms, only 38 patients (i.e., approximately 32% of the total Phe responders in the trial, who in turn represented only about 57% of the entire trial population) met this definition. According to the clinical expert consulted on this review, ADHD is not highly prevalent among patients with PKU, so it is unclear why the investigators would have considered this to be a potential target population. Therefore, the new clinical information does not clearly identify a subpopulation that

may benefit from SAP treatment. Overall, the interpretation of the clinical significance of the reported outcomes in both included trials is difficult, as the MCIDs for the reported outcomes are unknown and the majority of instruments used in the trials have not been validated in patients with PKU. Furthermore, neither included trial reported on outcomes that were identified as being most important to patients, were included in the review protocol, and which were also missing in the trials included in the original CDR review, namely quality of life and diet liberalization.

The first efficacy outcome in the current review protocol was quality of life; however, no quality of life outcomes were reported in either the PKU-016 trial or the SPARK study. According to the patient input received, patient expectations are that SAP will lower Phe levels and increase Phe tolerance so that patients are able to eat a more varied diet and be less isolated from their peers, both of which are expected to substantially affect patients' quality of life. As per Appendix 7 and as corroborated by the clinical expert, a Phe-restricted diet is expensive, unpalatable, and extremely difficult to maintain. There is evidence that adherence to the diet wanes over time and with increasing age.[19-21] Therefore, a clinically important treatment benefit of SAP could be its ability to facilitate diet liberalization by increasing dietary Phe tolerance. It follows that if SAP is able to maintain Phe blood levels within the desired range even when diet is not carefully restricted, this could have an important impact on patients' quality of life. Unfortunately, neither included trial has provided evidence that the reduction in Phe blood levels observed in Phe responders translates into improved quality of life. The only trial to specifically measure change in dietary Phe tolerance was the SPARK trial, which included children four years of age and younger. These children were already within the target Phe blood level range of 120 to 360 µmol/L with dietary therapy alone. Therefore, this patient population may not be the most appropriate in this context, as it is expected that their adherence to diet is very good due to vigilant control by parents and caregivers at this age. As noted previously, although there was a statistically significant increase in dietary Phe tolerance with SAP + diet compared with diet alone in the SPARK study, this was not linked with improved quality of life, or any measure of diet liberalization. Of note, the long-term studies summarized in Appendix 6 also do not provide evidence of an association of improved Phe tolerance with diet liberalization.

In both the PKU-016 (Phe responders) and SPARK trials, it was shown that mean Phe blood levels were reduced over the trial durations in both the SAP + diet arms and the placebo + diet or diet alone arms, respectively. In PKU-016, no statistical comparisons were made between treatment groups, but at week 13 (randomized double-blind treatment period), mean Phe blood levels decreased by approximately 30% in the SAP + diet arm, while remaining relatively unchanged in the placebo + diet arm. Following crossover to open-label SAP in the second 13-week open-label treatment period, mean Phe levels in patients who received placebo + diet in the first part of the trial also fell by approximately 30% after crossover with levels at week 26 being similar between both treatment groups. Of note, the mean Phe blood levels at week 26 were [redacted] (SAP + diet) and [redacted] (placebo + diet) which both exceed the upper limit of the blood Phe target range (360 µmol/L) as per the ACMG guidelines.[1] In the SPARK study, mean Phe levels were also reduced from baseline in both the SAP + diet and diet alone groups at week 26 (i.e., mean Phe levels were [redacted], respectively). The difference between groups was not statistically significant; however, both groups remained in the Phe blood target range of 120 to 360 µmol/L, as they were at baseline. In both included trials, diet was to remain unchanged throughout the duration of the trials which was verified by review of dietary report cards at clinic visits. Thus, there is no evidence from the included trials to show that increased diet liberalization occurred due to the addition of SAP to dietary therapy. Based on the results of the diet alone arm in the SPARK trial, it appears that diet alone is able to maintain patients within the target Phe blood level range.[1]

The results pertaining to the magnitude of reduction in Phe blood levels associated with SAP are difficult to interpret as the MCID for Phe blood level reduction (i.e., Phe response) is unknown, as shown in Appendix 5. In addition, the per cent reduction in Phe blood levels used to identify Phe responders (i.e., 20% or 30%) is arbitrary, as confirmed by the clinical expert. The ACMG guidelines state that clinical judgment is required to determine what constitutes a significant or beneficial decline in Phe blood levels from baseline in an individual and that 30% is often cited as evidence of effective Phe reduction, although some centres do consider 20% reduction as a measure of response.[1] In addition, although the ACMG guidelines recommend a target range of 120 to 360 µmol/L for mean Phe blood levels,[1] there is controversy over the threshold for treatment, as it has been suggested that the target range should be less than 360 µmol/L up to age 12 years and then less than 600 µmol/L thereafter.[8] There appears to be agreement that Phe blood levels of 360 µmol/L and less are not toxic to the brain and the clinical expert advised that in North America, the treatment aim is to maintain levels at below 360 µmol/L, although as shown in Appendix 8, it is suggested that in those aged 16 years and older, the target for Phe blood levels should be less than 600 µmol/L. Furthermore, as discussed in Appendix 5, there is also speculation that fluctuations in Phe levels are of potential clinical significance and that there may be up to 400% variation in day-to-day blood Phe levels in adults with well-controlled PKU, which may be influenced by age, genotype, rates of growth, dietary adherence, diet, and illness.[22] There is also evidence that Phe concentrations in the blood are higher than in the brain and that the disposition of Phe in the brain does not appear to parallel that in the blood.[22] Despite there being conflicting evidence regarding the effects of Phe fluctuations in patients with PKU,[22,23] it is still maintained that continual blood Phe concentrations should be monitored and controlled in childhood and that this should continue throughout the life of the patient with PKU.[22-24] It does not appear that there is yet widespread consensus regarding the threshold for initiating or maintaining treatment with SAP based on Phe blood levels; however, this is discussed in more detail below with regard to proposed reimbursement criteria.

One concern expressed by CEDAC in the original CDR review of Kuvan was that no neuropsychological outcomes were measured in the trials included in the review. In PKU-016, the effects of SAP treatment on neuropsychiatric and neurocognitive effects were investigated using various instruments (e.g., ADHD-RS/ASRS, HAM-D, HAM-A, BRIEF). As shown in Appendix 5, only the Inattention subscale of the ADHD-RS/ASRS appears to have been validated in children with PKU; however, no MCIDs for either the total score or either subscale score have been established.[25] The change from baseline to week 13 in the ADHD-RS/ASRS total score in Phe responders with ADHD symptoms was a co-primary end point in PKU-016 and results showed that while the change from baseline in the individual treatment arms were both statistically significant (i.e., indicating less severity of ADHD symptoms in both groups), the difference between the SAP + diet treated patients and placebo + diet patients was not statistically significant. At week 13, for the subscale score of Inattention, the difference between arms was statistically significant, whereas for the subscale score of Hyperactivity-Impulsivity, the difference between arms was not statistically significant. At week 26, the difference between arms was not statistically significantly different for the ADHD-RS/ASRS total score or either subscale score. These results are also difficult to interpret, as it appears that, in general, ADHD symptoms improved from baseline to week 13 in all treatment arms, but that there were no differences between arms with the exception of the Inattention subscale score. These results are also complicated by the fact that the MCIDs have not been established for this instrument in patients with PKU and that a correction factor was applied to the ASRS score, which has not been used previously or validated in PKU or other disease areas.[9] Other measures in PKU-016, such as the HAM-D and HAM-A, also showed a similar pattern as the ADHD-RS/ASRS total score, in that in individual treatment arms, the improvement from baseline to week 13 was statistically significant, but the difference between arms was not statistically significant.

Furthermore, in PKU-016 the results of the BRIEF assessment were not consistent. For the BRIEF-Parent results (i.e., patients younger than 18 years, for whom parents or caregivers completed the instrument), differences for the GEC and MI index scales were statistically significant between arms at week 13, but not for the BRI index scale. For the BRIEF-A (i.e., patients aged 18 years and older who self-completed the instrument), no statistically significant differences between arms were observed for any of the index subscales. It could be inferred from these findings that SAP does not appear to offer a significant additional benefit on improving neuropsychiatric outcomes compared with a well-controlled Phe-restricted diet. In the SPARK study, the effects of SAP treatment on four areas of neuromotor developmental milestones (i.e., fine motor, gross motor, language, and personal-social) were assessed. In all four areas, the majority of children were classified as being normal and there were no statistically significant differences between SAP + diet compared with diet alone reported in any area. This may be due to the children being within the Phe blood level target range at study entry or the short duration of the trial (26 weeks) precluding any significant change being observed in these assessments.

Another concern expressed by CEDAC in the original CDR review of Kuvan was that no information on growth parameters was provided in the included trials. The new clinical information did provide limited information on various growth parameters from the SPARK study where treatment differences were investigated for height, weight, BMI, and head circumference from baseline to week 26. [redacted]

The manufacturer has requested reimbursement criteria for Kuvan (as per section 1.3) that differ from the Exceptional Access Program (EAP) criteria (detailed in Appendix 8). In the manufacturer's suggested criteria, ongoing funding of SAP should be considered in patients who demonstrate increased Phe tolerance *or* improvement in neurobehavioural or neurocognitive function *or* improved quality of life, whereas the EAP criteria require that patients demonstrate increased Phe tolerance *and* improvement in neurobehavioural or neurocognitive function *and* improved quality of life. The full EAP criteria, which also encompass initial funding of Kuvan for six months (which is funded by the manufacturer), the requirements for a 72-hour Kuvan challenge to determine eligibility, and the management of pregnant patients are provided in Appendix 8. Also included in Appendix 8 are alternate EAP reimbursement criteria for Kuvan proposed by Ontario PKU experts, as in their view, the current EAP criteria are difficult, if not impossible, to meet due to contradictory requirements and lack of resources to perform serial neurobehavioural/neurocognitive testing, as required by the criteria. The rationale and background for the changes proposed by the Ontario specialists are also provided in Appendix 8; however, the key differences between the respective criteria are the Phe blood level at which treatment should be initiated, the length of time required to determine a response to Kuvan, and the therapeutic Phe blood level to be maintained in order to continue funding of Kuvan. To reiterate, the initial six months of funding for Kuvan is provided by the manufacturer and the key criteria for this initial coverage are compliance with a Phe-restrictive diet and a baseline Phe blood level greater than 360 µmol/L (despite dietary compliance) based on two levels over three to six months. Eligibility for initial funding is assessed by completion of a "72-hour Kuvan challenge" with SAP 20 mg/kg/day. Patients must demonstrate a reduction in Phe blood level of at least 30% and also have a baseline assessment of neurobehavioural or neurocognitive impairment and quality of life. In contrast, the Ontario specialists recommend initial funding of Kuvan in patients based on at least one Phe plasma level of 600 µmol/L or greater in non-pregnant patients and a four-week Kuvan challenge to determine response to SAP 20 mg/kg/day. Response to Kuvan is similarly defined as a 30% decrease in blood Phe levels from baseline. The

rationale for the higher threshold by the Ontario specialists is that there is a high degree of consensus that sustained Phe blood levels less than 360 µmol/L do not cause damage to the brain, but that levels 600 µmol/L and higher can cause significant damage.[1,26,27] The longer four-week challenge is proposed as it is widely accepted and can detect more patients who are responsive to SAP. For continued ongoing funding of Kuvan, the current EAP criteria require dietary compliance, Phe blood levels between 120 and 360 µmol/L, or sustained reductions of blood Phe of at least 30% if baseline levels are less than 1,200 µmol/L or at least 50% if baseline levels are greater than 1200 µmol/L (all based on two levels measured at least one month apart), demonstrated increase in dietary protein (Phe) tolerance, and clinically meaningful improvement in neurobehavioural/neurocognitive function or impairment as determined by valid quality of life instruments. In comparison, the Ontario specialists recommend ongoing funding with Kuvan if Phe blood levels are consistently in the control range, which is defined as less than 360 µmol/L in more than 50% of monitoring samples and less than 600 µmol/L in more than 80% of monitoring samples in patients younger than 16 years of age, and as less than 600 µmol/L in more than 80% of monitoring samples in patients aged 16 years or older (i.e., monitoring samples should be collected at least monthly). The Ontario specialists also advocate removal of the requirements for dietary compliance, demonstration of increased dietary Phe tolerance, and the need to demonstrate neurobehavioural/neurocognitive or quality of life improvements. For recommendations pertaining to pregnant patients, patients planning pregnancy, or neonatal patients, please see Appendix 8. Based on the new clinical evidence provided in this review, the included trials both employed four-week time frames to determine Phe responders and the Phe blood level target range was identified as 120 to 360 µmol/L for the patient populations in both trials. Although PKU-016 did show that treatment arms had improvements from baseline in some neurobehavioural/neurocognitive outcomes at week 13, for the most part, the differences between arms were not statistically significant or were inconsistent.

5.2.2 Harms

Although most patients in both trials, regardless of treatment arm, experienced treatment-emergent AEs, the frequencies of AEs were similar between treatment arms and most were mild or moderate in severity. The most frequent AEs were headache, nasopharyngitis, and vomiting in Study PKU-016 and pyrexia, cough, decreased amino acid level, and vomiting in the SPARK study. There were no deaths in either trial, few patients with SAEs, and only one patient with a WDAE in PKU-016. No safety concerns were identified in the original review of Kuvan. Thus, the findings from the new clinical evidence are consistent with those in the original CDR review and support that SAP is generally safe and well tolerated. In addition, no new safety issues were identified in the new clinical evidence.

5.3 Potential Place in Therapy

The information in this section is based on information provided in draft form by two clinical experts consulted by CDR reviewers for the purpose of this review.

Treatment of PKU is aimed at maintaining peripheral plasma Phe concentration within a therapeutic range, which, according to the ACMG guidelines,[1] is 120 to 360 µmol/L in patients of all ages, while European guidelines[8] recommend that levels be maintained at < 360 µmol/L up to an age of 12 years and < 600 µmol/L thereafter. The target of < 360 µmol/L recommended by the ACMG is, in practice, unachievable for most adult patients due to the dietary burden. For adults, clinicians often individualize the target Phe levels to between 360 and 600 µmol/L, depending on neurocognitive and psychiatric function and ability to adhere to diet.

A low-Phe diet is the current standard of care used to maintain Phe levels within an appropriate range: patients are not allowed meat, fish, fowl, bread, milk, cheese, nuts, or certain legumes (i.e., they must

follow a "vegan-vegetarian" diet supplemented with synthetic amino-acid "medical foods" to provide protein, vitamins, and trace elements). The low-Phe diet required to maintain low blood Phe levels is very restrictive and, in addition to being unpalatable, is very burdensome and many patients cannot maintain this control.

While newborn screening and dietary treatment have been very successful in preventing severe neurological deficits in PKU patients, there remains an unmet need. The unpalatable and burdensome dietary requirements mean that dietary compliance becomes increasingly poor after the toddler age, and by adolescence and adulthood, a lack of dietary compliance means that desirable blood levels are attained in less than 20% of patients.[19-21] This can result in executive function abnormalities and psychiatric symptoms that can prevent many patients from achieving their full potential in terms of education achievement and social functioning. Absence of life-threatening consequences that occur in other metabolic diseases hides the true socioeconomic burden of PKU.

The aforementioned unmet needs of PKU patients could be met by improving access to psychological and social services, dietary products, and innovative treatments, such as sapropterin, in a coordinated clinical setting. Sapropterin can improve Phe tolerance and stabilize Phe levels in 50% to 60% of individuals with PKU. Increased Phe tolerance is demonstrated by an ability to tolerate increased natural protein in diet.

Patients who respond to SAP are able to ingest up to 50% more natural protein versus synthetic protein and still maintain target blood levels.[14,28,29] Only patients with mild PKU (untreated Phe levels < 600 µmol/L) can be treated with SAP alone; all other responders will require continued dietary restriction, albeit with increased flexibility that would likely increase dietary compliance, which in turn might be expected to improved patients' quality of life.

Identification of patients who would benefit from SAP is based on determining which patients respond to treatment. Thus, "testing" for SAP responders involves administering a loading dose of SAP (e.g., 20 mg/kg/day) and assessing changes in Phe levels. Responders are usually identifiable by a decrease in Phe levels by 48 hours, although a few "slow responders" take up to one month for Phe levels to decrease. Therefore, a one-month loading process is recommended to identify responders. A responder is generally accepted as someone who has a 30% or greater reduction in Phe in response to SAP while on a stable diet. No other specialized tests are needed. For responders, the effective maintenance dosage of SAP varies from 5 mg/kg/day to 20 mg/kg/day.

Clinical experience in Canada suggests that approximately 30% of PKU patients will respond to SAP and continue on the medication; others will either not respond, or will respond but discontinue treatment due to the additional burden of having to take the medication. There is limited evidence to support improvement of executive or psychological functioning without a demonstrable change in Phe tolerance. If SAP is prescribed for improvement in function alone, continuation of treatment should be supported by appropriate objective measures as suggested by the information provided by the Ontario physicians for the purpose of this review (see Appendix 8). However, reimbursement criteria that include use of placebos and/or formal neurocognitive testing could be a barrier to access due to limited resources.

6. CONCLUSIONS

Two randomized controlled phase 3b trials met the selection criteria for inclusion in the systematic review: PKU-016 and the SPARK study. In a proportion of patients identified as Phe responders, SAP + a Phe-restricted diet either reduced (PKU-016) or maintained (SPARK) Phe blood levels within the target range. The magnitude of the reduction in Phe levels with SAP + diet was similar to that of placebo + diet, or diet alone, respectively, in the included trials. In the SPARK study, diet alone was able to maintain patients within the target Phe blood level range of 120 to 360 µmol/L, although SAP + diet was associated with a statistically significant increase in dietary Phe tolerance compared with diet alone. In neither trial was a reduction in Phe blood levels or an increase in Phe tolerance associated with diet liberalization or improved quality of life, both of which are outcomes that are important to patients. In PKU-016, improvements from baseline in various neuropsychiatric/neurocognitive outcomes were observed with SAP + diet, including ADHD symptoms in a subpopulation of Phe responders, but the differences between treatment arms were either not statistically significant or inconsistent. Interpretation of the results from the included trials is complicated by a large placebo effect and lack of validation of measured outcomes or identification of MCIDs in patients with PKU. The safety and tolerability profile of SAP was consistent with that reported in the original CDR review, which suggests that SAP is generally safe and well tolerated. Overall, the new clinical evidence does not appear to reduce the uncertainty identified in the original CDR review regarding identification of a subpopulation of patients with PKU who will ultimately benefit from SAP treatment, what is considered to be a clinically meaningful response to SAP, and whether the changes observed in Phe levels are correlated with clinically important outcomes.

APPENDIX 1: PATIENT INPUT SUMMARY

This section was prepared by CADTH staff based on the input provided by patient groups.

1. Brief Description of Patient Group Supplying Input
Canadian PKU and Allied Disorders Inc. (CanPKU) is a non-profit organization dedicated to helping families and professionals dealing with phenylketonuria (PKU) and other inherited metabolic disorders by providing information, outreach, education, advocacy, and support. CanPKU has received unrestricted educational grants in 2015 and 2016 from BioMarin, Cambrooke Therapeutics, Innomar, Nutricia, Rx&D, and Vitaflo.

No conflicts of interest were declared with regard to this patient group submission.

2. Condition-Related Information
CanPKU conducted two separate online surveys between January 7 and 27, 2016 to gather information (the surveys contained free-form commentary, scoring options, and limited closed questions). Patients with PKU and their caregivers were contacted using membership databases, social media, and a partner patient group in the United States. A total of 297 of those contacted responded. Physicians and dieticians who treat PKU (n = 8) were also contacted and provided input regarding prescribing information, key factors for treatment choices, and obstacles to positive outcomes.

PKU is a metabolic disorder whereby the patients affected cannot break down phenylalanine (Phe). This subsequently leads to an increase in the amount of blood Phe to neurotoxic levels and can cause moderate to severe neurocognitive, neuropsychiatric, and developmental problems. Therefore, patients diagnosed as having PKU are required to follow a strict Phe-restricted diet, which involves controlling natural protein intake and drinking a synthetic Phe-free protein formula. Not only is this diet expensive with limited food choices, it is also laborious, complex, and unpalatable. Most patients surveyed have difficulties maintaining their strict diet and either they or their caregivers must spend a disproportionate amount of their time maintaining the diet. This includes undertaking things like cooking for the Phe-restricted diet, supervising protein intake, planning daily Phe intake, weighing foods, keeping records, ordering low proteins and amino acids, and preparing for social events. In addition, time is also consumed by clinic appointments and going for blood testing.

While each patient's experience with PKU is unique, the management of this disorder places a large burden on both patients and their families. Patients with PKU can suffer from behavioural, mood, emotional, and social problems, psychiatric disorders, intellectual development delays, and neurological deficits. Manifestations from the aforementioned include difficulty focusing and concentrating, the inability to retain written or verbal communications, problems with short-term memory, hyperactivity, anxiety, panic attacks, depression, attention-deficit/hyperactivity disorder, developmental delays and learning disabilities, lower IQs, the inability to perform executive functioning, shaking, and seizures. In addition, the patient with PKU may also experience poor nutritional status, which subsequently can lead to poor bone strength or other nutritional issues (e.g., osteoporosis, thyroid issues, maintaining weight and muscle mass).

Isolation is experienced by both the patients and their caregivers. Children may become socially isolated at school due to dietary restrictions or behavioural issues arising from high Phe levels, while families may not attend social functions due to the time or inability to prepare appropriate foods. Travel is often

decreased or halted altogether. If travel does occur, families have confessed to travelling with one additional suitcase full of medical food.

The caregivers are often affected as much as the patient with PKU. Parents often feel that they cannot be separated from their child for extensive periods of time due to fears that the child will not receive the proper food or will eat inappropriate foods, or others will not be able to appropriately control, log, weigh, or calculate everything. In addition, there is no "break" from the Phe-restricted diet. Many parents have quit employment to attend to their child's needs, which can adversely affect the person staying home, along with the financial viability of the family. Isolation and lack of time for oneself can also be substantial, as the caregiver is often responsible for preparing, controlling, counting, logging, and calculating everything with regard to their child's diet. Furthermore, the ordering of the food and amino acids and the trips to the clinic including for regular blood draws take additional time from their lives. As one person stated, *"We don't manage PKU around our lives; we manage our lives around PKU."* The true sacrifice that parents are often faced with is observed through the following statement, *"I was very strict with her diet and it was all consuming. At McDonald's birthday parties, I would cut fries into 2 inch pieces and count out 20 of them, knowing that was her lunch allotment. If she didn't eat her school lunch, I would sit in the car with her in the parking lot until she finished it all. I had to run her formula to her if plans changed at school or activities because she could not wait till she got home to drink it or she would start shaking."* Stress also remains a constant factor, as caregivers are often consumed with the potential for brain damage to occur, dealing with behavioural issues or learning issues, and potentially not spending enough time with their other children. In addition, financial difficulties are often associated with the cost of the Phe-restricted diet. This, in turn, can lead to increased frustration and family strife.

3. Current Therapy-Related Information

The gold standard for PKU treatment (and, up until Kuvan, the only treatment) is the Phe-restricted diet for life. This diet is extremely limited in natural protein and requires supplementation with a Phe-free amino acid synthetic formula. This Phe-restricted treatment is one of the most restricted diets known and is unpalatable, complicated (requiring painstaking efforts in preparation, record-keeping, and blood testing), expensive, and is an extreme burden on both the patient and the caregiver. Problems with adherence in the child and adolescent populations are of particular importance as patients with PKU can suffer from isolation due to such a restrictive diet at this age. The aforementioned reasons are causes for discontinuation of the Phe-restricted diet: thus, noncompliance remains a serious disease management issue. The heavy burden of the diet treatment was confirmed by the treating clinicians in their separate survey.

There is a large unmet need for the introduction of any other form of treatment that will allow patients with PKU the chance at a more normal life. Patients and their caregivers seek new treatments that will allow the patient the ability to consume other more natural foods while keeping the Phe levels in a normal range. By simply having only one treatment option (which is hard to adhere to and is very restrictive and unpalatable), the patient is put into a position that leaves them vulnerable to neurologic, neuropsychiatric, behavioural, and developmental issues.

4. Expectations About the Drug Being Reviewed

Patients with PKU and their caregivers who do not have experience with Kuvan believe that it will help lower blood Phe levels and increase Phe tolerance so that patients will be able to eat a more varied diet and be less isolated from their peers. However, a lot of families worry about the cost of Kuvan as a treatment as, without coverage, it is financially unfeasible. This was observed from the following,

"We are very worried and saddened to think that in six months, we may not be able to afford Kuvan and our daughter will be taken off a drug that has potentially opened up her life to success in so many ways; mentally, physically, and emotionally. It is devastating to know that there is a drug available like this and we can't have access to it; then to watch your child regress to the shell she was without it. There are no words." Some families who were offered the trial were too afraid to enroll as they knew, once the trial was completed, they would not be able to afford the treatment.

In patients who have experience with Kuvan, Kuvan is seen as a positive, life-altering treatment that not only leads to a decrease and sometimes a stabilization in blood Phe levels, but also allows the patient to eat a more normal diet and avoid the adverse consequences associated with higher blood Phe levels. Patients on Kuvan have reported beneficial cognitive impacts, improved academic performance, improved executive functioning and focus, improved energy, and increased confidence in their futures. Some patients have reported it as their "cure," indicating that their levels of Phe became so low that the specialized formula and severe dietary restrictions previously required were no longer a requirement. There was also an increase in the tolerance of Phe in a lot of patients who were taking Kuvan, which subsequently allowed them to eat more natural foods. In addition, patients were observed to show developmental and behavioural improvement, which was observed in one patient as improved *"ability to focus and she is able to absorb and process new information. Drastically lowered anxiety, very little shaking, and no irritability."* In essence, some children on Kuvan began experiencing a normal life, as indicated in the following quote, *"He is now on a normal diet. Yes, that means he can eat whatever he wants and doesn't have to count Phes (he can now eat anything including meat, seafood, pizza, anything). He doesn't have to drink his formula, either. He is now thinking better and growing better and is feeling more like everyone else. We only have to go for blood work once a month and his blood Phe numbers are always right on and not fluctuating up and down."* In many cases, the thought of returning to the strict Phe-restricted diet should Kuvan be unattainable is devastating. Finally, decreased costs associated with the Phe-restricted diet has relieved some of the financial burden associated with PKU treatment.

Adverse events noted by patients on Kuvan included heartburn, gastroesophageal reflux disease (GERD), or acid reflux or stomach discomfort, mild injection-site reactions, sleep disturbances, hyperactivity, stomach pain, headache, nausea, diarrhea, agitation, nasal congestion, cough, vomiting, joint pain, and dizziness; however, these were reported to be mild and often disappeared with use.

APPENDIX 2: LITERATURE SEARCH STRATEGY

OVERVIEW	
Interface:	Ovid
Databases:	Embase 1974 to present MEDLINE Daily and MEDLINE 1946 to present MEDLINE In-Process & Other Non-Indexed Citations **Note:** Subject headings have been customized for each database. Duplicates between databases were removed in Ovid.
Date of Search:	March 1, 2016
Alerts:	Weekly search updates until July 20, 2016
Study Types:	No search filters were applied
Limits:	Date limits: 2011 – 2016 (to update a previous Kuvan resubmission search performed in June 2011) Human filter was applied Conference abstracts were excluded

SYNTAX GUIDE	
/	At the end of a phrase, searches the phrase as a subject heading
.sh	At the end of a phrase, searches the phrase as a subject heading
MeSH	Medical Subject Heading
fs	Floating subheading
exp	Explode a subject heading
*	Before a word, indicates that the marked subject heading is a primary topic; or, after a word, a truncation symbol (wildcard) to retrieve plurals or varying endings
adj	Requires words are adjacent to each other (in any order)
.ti	Title
.ab	Abstract
.ot	Original title
.hw	Heading word; usually includes subject headings and controlled vocabulary
.kf	Author keyword heading word (MEDLINE)
.kw	Author keyword (Embase)
.pt	Publication type
.po	Population group [PsycInfo only]
.rn	CAS registry number
.nm	Name of substance word
pmez	Ovid database code; MEDLINE In-Process & Other Non-Indexed Citations, MEDLINE Daily and Ovid MEDLINE 1946 to Present
oemezd	Ovid database code; Embase 1974 to present, updated daily

MULTI-DATABASE STRATEGY	
Embase, Ovid MEDLINE(R)	
#	
1	(27070-47-9 or 69056-38-8).rn,nm.
2	(sapropterin* or kuvan* or 6R-BH4* or phenoptin* or SUN-0588 or sun0588 or biobuden* or biopten* or r-thbp or bipten* or RG277LFSB3).ti,ab,ot,kf,rn,hw,nm.
3	2 use pmez
4	(2011* or 2012* or 2013* or 2014* or 2015* or 2016*).ed,dp,dc,ep.
5	3 and 4
6	*sapropterin/
7	(sapropterin* or kuvan* or 6R-BH4* or phenoptin* or SUN-0588 or sun0588 or biobuden* or biopten* or r-thbp or bipten*).ti,ab,kw.
8	6 or 7
9	8 use oemezd
10	(2011* or 2012* or 2013* or 2014* or 2015* or 2016*).dd.
11	9 and 10
12	5 or 11
13	conference abstract.pt.
14	12 not 13
15	exp animals/
16	exp animal experimentation/ or exp animal experiment/
17	exp models animal/
18	nonhuman/
19	exp vertebrate/ or exp vertebrates/
20	or/15-19
21	exp humans/
22	exp human experimentation/ or exp human experiment/
23	or/21-22
24	20 not 23
25	14 not 24
26	remove duplicates from 25

OTHER DATABASES	
PubMed	A limited PubMed search was performed to capture records not found in MEDLINE. Same MeSH, keywords, limits, and study types used as per MEDLINE search, with appropriate syntax used.
Trial registries (Clinicaltrials.gov and others)	Same keywords, limits used as per MEDLINE search.

Grey Literature

Dates for Search:	February 26, 2016
Keywords:	Kuvan (sapropterin dihydrochloride), Phenylketonuria
Limits:	Date limits: 2011 – 2016; No language limits used

Relevant websites from the following sections of the CADTH grey literature checklist, "Grey matters: a practical tool for searching health-related grey literature" (www.cadth.ca/grey-matters) were searched:

- Health Technology Assessment Agencies
- Health Economics
- Clinical Practice Guidelines
- Drug and Device Regulatory Approvals
- Advisories and Warnings
- Drug Class Reviews
- Databases (free)
- Internet Search.

APPENDIX 3: EXCLUDED STUDIES

There were no excluded studies, as all potentially relevant studies were included in the systematic review.

APPENDIX 4: DETAILED OUTCOME DATA

TABLE 8: STUDY PKU-016: ADHD-RS/ASRS TOTAL SCORE (PHE RESPONDERS WITH ADHD SYMPTOMS)

PKU-016		
	SAP + Diet (n = 19)	Placebo + Diet (n = 19)
Change From Baseline to Week 13		
Baseline		
LS mean (SE)[a] 95% CI of LS mean	28.9 (2.4) (24.1 to 33.7)	31.2 (2.2) (26.8 to 35.6)
LS mean difference from placebo (SE)[a] 95% CI of LS mean difference	[redacted]	
Week 13		
LS mean (SE)[b] 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE)[b] 95% CI of LS mean difference	[redacted]	
Week 13 change from baseline		
LS mean (SE)[c] 95% CI of LS mean; *P* value[d]	–9.1 (2.2) (–13.5 to –4.7); [redacted]	–4.9 (2.0) (–8.9 to -0.9); [redacted]
LS mean difference from placebo (SE)[c] 95% CI of LS mean difference; *P* value[e]	–4.2 (2.3) (–8.9 to 0.6); 0.085	
Change from Week 13 to Week 26		
Week 13		
LS mean (SE)[a] 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE)[a] 95% CI of LS mean difference	[redacted]	
Week 26		
LS mean (SE)[b] 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE)[b] 95% CI of LS mean difference	[redacted]	
Week 26 change from week 13		
LS mean (SE)[c] 95% CI of LS mean; *P* value[d]	[redacted]	[redacted]
LS mean difference from placebo (SE)[c] 95% CI of LS mean difference; *P* value[e]	[redacted]	

ADHD = attention-deficit/hyperactivity disorder; ADHD-RS = Attention Deficit Hyperactivity Disorder Rating Scale; ANCOVA = analysis of covariance; ASRS = Adult Attention Deficit Hyperactivity Disorder Self-Report Scale; CI = confidence interval; LS = least squares; Phe = phenylalanine; SAP = sapropterin; SD = standard deviation; SE = standard error.
Note: A higher score on either the ADHD-RS or ASRS corresponds to greater severity of ADHD symptoms. A correction factor of 0.75 was applied to the ASRS scores to allow numerical combination of the scores.
[a] Based on ANCOVA of baseline ADHD-RS/ASRS total score, with model terms treatment, age group, and ADHD medication.
[b] Based on ANCOVA of week 13 ADHD-RS/ASRS total score, with model terms treatment, age group, and ADHD medication.
[c] ANCOVA of change from baseline with terms treatment, baseline ADHD-RS/ASRS total score, age group, and ADHD medication.
[d] *P* value determined by t test testing mean change from Baseline at week 13 equals 0.
[e] *P* value determined by t test versus placebo from ANCOVA model.
Source: Clinical Study Report PKU-01.[10]

TABLE 9: STUDY PKU-016: ADHD-RS/ASRS INATTENTION SUBSCALE SCORE (PHE RESPONDERS WITH ADHD SYMPTOMS)

PKU-016		
	SAP + Diet (n = 19)	Placebo + Diet (n = 19)
Change From Baseline to Week 13		
Baseline		
LS mean (SE)[a] 95% CI of LS mean	18.0 (1.3) [redacted]	19.2 (1.2) [redacted]
LS mean difference from placebo (SE)[a] 95% CI of LS mean difference	[redacted]	
Week 13		
LS mean (SE)[b] 95% CI of LS mean	12.4 (1.7) [redacted]	16.6 (1.5) [redacted]
LS mean difference from placebo (SE)[b] 95% CI of LS mean difference	[redacted]	
Week 13 change from baseline		
LS mean (SE)[c] 95% CI of LS mean; *P* value[d]	–5.9 (1.4) (–8.9 to –3.0); [redacted]	–2.5 (1.3) (–5.2 to 0.1); [redacted]
LS mean difference from placebo (SE)[c] 95% CI of LS mean difference; *P* value[e]	–3.4 (1.6) (–6.6 to –0.2); 0.036	
Change From Week 13 to Week 26		
Week 13		
LS mean (SE)[a] 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE)[a] 95% CI of LS mean difference	[redacted]	
Week 26		
LS mean (SE)[b] 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE)[b] 95% CI of LS mean difference	[redacted]	
Week 26 change from week 13		
LS mean (SE)[c] 95% CI of LS mean; *P* value[d]	[redacted]	[redacted]
LS mean difference from placebo (SE)[c] 95% CI of LS mean difference; *P* value[e]	[redacted]	

ADHD = attention-deficit/hyperactivity disorder; ADHD-RS = Attention Deficit Hyperactivity Disorder Rating Scale; ANCOVA = analysis of covariance; ASRS = Adult Attention Deficit Hyperactivity Disorder Self-Report Scale; CI = confidence interval; LS = least squares; Phe = phenylalanine; SAP = sapropterin; SD = standard deviation; SE = standard error.
Note: The ADHD-RS or ASRS Inattention subscale score comprises 9 questions out of the total score for Inattention. The maximum score for a subscale in the ADHD-RS is 27 and in the ASRS is 36. A correction factor of 0.75 was applied to the ASRS scores to allow numerical combination of the scores.
[a] Based on ANCOVA of baseline ADHD-RS/ASRS Inattention subscale score, with model terms treatment, age group, and ADHD medication.
[b] Based on ANCOVA of week 13 ADHD-RS/ASRS Inattention subscale score, with model terms treatment, age group, and ADHD medication.
[c] ANCOVA of change from baseline with terms treatment, baseline ADHD-RS/ASRS Inattention Subscale score, age group, and ADHD medication.
[d] *P* value determined by t test testing mean change from Baseline at week 13 equals 0.
[e] *P* value determined by t test versus placebo from ANCOVA .Source: Clinical Study Report PKU-016.[10]

TABLE 10: STUDY PKU-016: ADHD-RS/ASRS HYPERACTIVITY-IMPULSIVITY SUBSCALE SCORE (PHE RESPONDERS WITH ADHD SYMPTOMS)

PKU-016		
	SAP + Diet (n = 19)	Placebo + Diet (n = 19)
Change from Baseline to Week 13		
Baseline		
LS mean (SE)[a] 95% CI of LS mean	10.9 (2.0) [redacted]	12.0 (1.8) [redacted]
LS mean difference from placebo (SE)[a] 95% CI of LS mean difference	[redacted]	
Week 13		
LS mean (SE)[b] 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE)[b] 95% CI of LS mean difference	[redacted]	
Week 13 change from baseline		
LS mean (SE)[c] 95% CI of LS mean; *P* value[d]	–3.3 (1.1) (–5.6 to –1.1); [redacted]	–2.3 (1.0) (–4.3 to –0.3); [redacted]
LS mean difference from placebo (SE)[c] 95% CI of LS mean difference; *P* value[e]	–1.0 (1.2) (–3.4 to 1.4); 0.396	
Change from Week 13 to Week 26		
Week 13		
LS mean (SE)[a] 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE)[a] 95% CI of LS mean difference	[redacted]	
Week 26		
LS mean (SE)[b] 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE)[b] 95% CI of LS mean difference	[redacted]	
Week 26 change from week 13		
LS mean (SE)[c] 95% CI of LS mean; *P* value[d]	[redacted]	[redacted]
LS mean difference from placebo (SE)[c] 95% CI of LS mean difference; *P* value[e]	[redacted]	

ADHD = attention-deficit/hyperactivity disorder; ADHD-RS = Attention Deficit Hyperactivity Disorder Rating Scale; ANCOVA = analysis of covariance; ASRS = Adult Attention Deficit Hyperactivity Disorder Self-Report Scale; CI = confidence interval; LS = least squares; Phe = phenylalanine; SAP = sapropterin; SD = standard deviation; SE = standard error.
Note: The ADHD-RS or ASRS Hyperactivity-Impulsivity subscale score comprises 9 questions out of the total score for Hyperactivity-Impulsivity. The maximum score for a subscale in the ADHD-RS is 27 and in the ASRS is 36. A correction factor of 0.75 was applied to the ASRS scores to allow numerical combination of the scores.
[a] Based on ANCOVA of baseline ADHD-RS/ASRS Hyperactivity-Impulsivity subscale score, with model terms treatment, age group, and ADHD medication.
[b] Based on ANCOVA of week 13 ADHD-RS/ASRS Hyperactivity-Impulsivity subscale score, with model terms treatment, age group, and ADHD medication.
[c] ANCOVA of change from baseline with terms treatment, baseline ADHD-RS/ASRS Hyperactivity-Impulsivity subscale score, age group, and ADHD medication.
[d] *P* value determined by t test testing mean change from Baseline at week 13 equals 0.
[e] *P* value determined by t test versus placebo from ANCOVA.
Source: Clinical Study Report PKU-016.[10]

TABLE 11: STUDY PKU-016: GLOBAL FUNCTION EVALUATION (CGI-I) IN PHE RESPONDERS

PKU-016		
	SAP + Diet (n = 61)	Placebo + Diet (n = 57)
All Phe Responders		
Week 13		
N	60	57
Patients with CGI-I rating of 1 or 2, n (%)	13 (21.7)	15 (26.3)
Relative risk[a] 95% CI of relative risk; *P* value	0.87 (0.46 to 1.64); 0.670	
Week 26		
N	[redacted]	[redacted]
Patients with CGI-I Rating of 1 or 2, n (%)	[redacted]	[redacted]
Relative risk[a] 95% CI of relative risk; *P* value	[redacted]	
Phe Responders With ADHD Symptoms		
Week 13		
N	[redacted]	[redacted]
Patients with CGI-I Rating of 1 or 2, n (%)	[redacted]	[redacted]
Relative risk[a] 95% CI of relative risk; *P* value	[redacted]	
Week 26		
N	[redacted]	[redacted]
Patients with CGI-I Rating of 1 or 2, n (%)	[redacted]	[redacted]
Relative risk[a] 95% CI of relative risk; *P* value	[redacted]	

ADHD = attention-deficit/hyperactivity disorder; CGI-I = Clinical Global Impression–Improvement; CI = confidence interval; SAP = sapropterin.

Notes:

- Cochran–Mantel–Haenszel analysis at week 13 or week 26.
- The CGI-I is a 7-point scale that assesses improvement in global impression of behaviour, symptoms, and functioning relative to the baseline state at the beginning of the intervention. Lower scores indicate improvement. A rating of 1= very much improved and 2= much improved.

[a] Adjusted for age group, ADHD symptoms and ADHD medication.

Source: Clinical Study Report PKU-016.[10]

TABLE 12: STUDY PKU-016: GLOBAL FUNCTION EVALUATION (CGI-S) IN PHE RESPONDERS

PKU-016		
	SAP + Diet (n = 61)	Placebo + Diet (n = 57)
Change From Baseline to Week 13		
Baseline		
LS mean (SE)[a] 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE)[a] 95% CI of LS mean difference	[redacted]	

PKU-016		
	SAP + Diet (n = 61)	Placebo + Diet (n = 57)
Week 13		
LS mean (SE)[b] 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE)[b] 95% CI of LS mean difference	[redacted]	
Week 13 change from baseline		
LS mean (SE)[c] 95% CI of LS mean; *P* value[d]	–0.6 (0.2) (–0.9 to –0.3); [redacted]	–0.5 (0.2) (–0.8 to –0.2); [redacted]
LS mean difference from placebo (SE)[c] 95% CI of LS mean difference; *P* value[e]	–0.1 (0.2) (–0.4 to 0.2); 0.531	
Change From Week 13 to Week 26		
Week 13		
LS mean (SE)[a] 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE)[a] 95% CI of LS mean difference	[redacted]	
Week 26		
LS mean (SE)[b] 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE)[b] 95% CI of LS mean difference	[redacted]	
Week 26 change from week 13		
LS mean (SE)[c] 95% CI of LS mean; *P* value[d]	0.1 (0.2) (–0.2 to 0.4); [redacted]	(–0.0 (0.2) (–0.3 to 0.3); [redacted]
LS mean difference from placebo (SE)[c] 95% CI of LS mean difference; *P* value[e]	0.1 (0.2) (–0.2 to 0.5); 0.510	

ADHD = attention-deficit/hyperactivity disorder; ANCOVA = analysis of covariance; CGI-S = Clinical Global Impression–Severity; CI = confidence interval; LS = least squares; SAP = sapropterin; SE = standard error.
Note: The CGI-S is a 7-point scale that requires the clinician to rate the severity of a patient's illness at the time of assessment relative to the clinician's past experience with patients who have the same diagnosis. Lower scores indicate improvement.
[a] Based on ANCOVA of baseline CGI-S response with model terms treatment, age group, and ADHD medication.
[b] Based on ANCOVA of week 13/week 26 CGI-S response with model terms treatment, age group, and ADHD medication.
[c] ANCOVA of change from baseline/week 13 with terms treatment, baseline CGI-S response, age group, and ADHD medication.
[d] *P* value determined by t test testing mean change from baseline/week 13 at week 13/week 26 equals 0.
[e] *P* value determined by t test versus placebo from ANCOVA model.
Source: Clinical Study Report PKU-016.[10]

TABLE 13: STUDY PKU-016: HAMILTON ANXIETY RATING SCALE TOTAL SCORE IN PHE RESPONDERS

PKU-016		
	SAP + Diet (n = 61)	Placebo + Diet (n = 57)
Change From Baseline to Week 13		
Baseline		
LS mean (SE)[a] 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE)[a] 95% CI of LS mean difference	[redacted]	

PKU-016		
	SAP + Diet (n = 61)	**Placebo + Diet (n = 57)**
Week 13		
LS mean (SE)[b] 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE)[b] 95% CI of LS mean difference	[redacted]	
Week 13 change from baseline		
LS mean (SE)[c] 95% CI of LS mean; *P* value[d]	–3.2 (0.9) (–5.1 to –1.4); < 0.001	–3.6 (0.9) (–5.4 to –1.9); < 0.001
LS mean difference from placebo (SE)[c] 95% CI of LS mean difference; *P* value[e]	0.4 (0.9) (–1.5 to 2.3); 0.669	
Change From Week 13 to Week 26		
Week 13		
LS mean (SE)[a] 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE)[a] 95% CI of LS mean difference	[redacted]	
Week 26		
LS mean (SE)[b] 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE)[b] 95% CI of LS mean difference	[redacted]	
Week 26 change from week 13		
LS mean (SE)[c] 95% CI of LS mean; *P* value[d]	–0.5 (0.9) (–2.3 to 1.4); 0.328	0.1 (0.9) (–1.7 to 1.8); 0.947
LS mean difference from placebo (SE)[c] 95% CI of LS mean difference; *P* value[e]	–0.5 (1.0) (–2.4 to 1.4); 0.590	

ADHD = attention-deficit/hyperactivity disorder; ANCOVA = analysis of covariance; CI = confidence interval; HAM-A = Hamilton Anxiety Rating Scale; LS = least squares; SAP = sapropterin; SE = standard error.
Note: Decline in mean score in the HAM-A represents an improvement in symptoms.
[a] Based on ANCOVA of baseline HAM-A response with model terms treatment, age group, ADHD symptoms, and ADHD medication.
[b] Based on ANCOVA of week 13/week 26 HAM-A response with model terms treatment, age group, ADHD symptoms and ADHD medication.
[c] ANCOVA of change from baseline/week 13 with terms treatment, baseline HAM-A response, age group, ADHD symptoms and ADHD medication.
[d] *P* value determined by t test testing mean change from baseline/week 13 at week 13/week 26 equals 0.
[e] *P* value determined by t test versus placebo from ANCOVA model.
Source: Clinical Study Report PKU-016.[10]

TABLE 14: STUDY PKU-016: HAMILTON DEPRESSION RATING SCALE TOTAL SCORE IN PHE RESPONDERS

PKU-016		
	SAP + Diet (n = 61)	Placebo + Diet (n = 57)
Change From Baseline to Week 13		
Baseline		
LS mean (SE)[a] 95% CI of LS mean	8.7 (1.1) (6.5 to 10.9)	8.2 (1.0) (6.1 to 10.3)
LS mean difference from placebo (SE)[a] 95% CI of LS mean difference	0.5 (1.1) (–1.7 to 2.7)	
Week 13		
LS mean (SE)[b] 95% CI of LS mean	6.3 (0.9) (4.5 to 8.0)	5.6 (0.8) (4.0 to 7.3)
LS mean difference from placebo (SE)[b] 95% CI of LS mean difference	0.6 (0.9) (–1.2 to 2.4)	
Week 13 change from baseline		
LS mean (SE)[c] 95% CI of LS mean; *P* value[d]	–2.1 (0.7) (–3.6 to –0.6); < 0.001	–2.5 (0.7) (–3.9 to –1.1); < 0.001
LS mean difference from placebo (SE)[c] 95% CI of LS mean difference; *P* value[e]	0.4 (0.8) (–1.1 to 1.9); 0.588	
Change From Week 13 to Week 26		
Week 13		
LS mean (SE)[a] 95% CI of LS mean	6.3 (0.9) (4.5 to 8.0)	5.6 (0.8) (4.0 to 7.3)
LS mean difference from placebo (SE)[a] 95% CI of LS mean difference	0.6 (0.9) (–1.2 to 2.4)	
Week 26		
LS mean (SE)[b] 95% CI of LS mean	6.4 (1.0) (4.5 to 8.4)	5.5 (0.9) (3.7 to 7.4)
LS mean difference from placebo (SE)[b] 95% CI of LS mean difference	0.9 (1.0) (–1.1 to 2.9)	
Week 26 change from week 13		
LS mean (SE)[c] 95% CI of LS mean; *P* value[d]	0.5 (0.8) (–1.1 to 2.0); 0.873	0.1 (0.7) (–1.3 to 1.5); 0.734
LS mean difference from placebo (SE)[c] 95% CI of LS mean difference; *P* value[e]	0.4 (0.8) (–1.2 to 1.9); 0.636	

CI = confidence interval; HAM-D = Hamilton Depression Rating Scale; LS = least squares; SAP = sapropterin; SE = standard error.
Note: Decline in mean score in the Hamilton Depression rating scale represents an improvement in symptoms.
[a] Based on ANCOVA of baseline HAM-D response with model terms treatment, age group, ADHD symptoms, and ADHD medication.
[b] Based on ANCOVA of week 13/week 26 HAM-D response with model terms treatment, age group, ADHD symptoms, and ADHD medication.
[c] ANCOVA of change from baseline/week 13 with terms treatment, baseline HAM-D response, age group, ADHD symptoms, and ADHD medication.
[d] *P* value determined by t test testing mean change from baseline/week 13 at week 13/week 26 equals 0.
[e] *P* value determined by t test versus placebo from ANCOVA model.
Source: Clinical Source Report PKU-016.[10]

TABLE 15: STUDY PKU-016: BRIEF-PARENT T SCORES IN PHE RESPONDERS < 18 YEARS OF AGE

PKU-016		
	SAP + Diet (n = 36)	Placebo + Diet (n = 29)
GEC T Score		
Baseline		
LS mean (SE) 95% CI of LS mean	63.9 (1.9) [redacted]	63.7 (2.0) [redacted]
LS mean difference from placebo (SE) 95% CI of LS mean difference	[redacted]	
Week 13		
LS mean (SE) 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE) 95% CI of LS mean difference	[redacted]	
Week 13 change from baseline		
LS mean (SE) 95% CI of LS mean	–4.8 (1.6) (–8.0 to –1.6)	–0.7 (1.7) (–4.0 to 2.7)
LS mean difference from placebo (SE) 95% CI of LS mean difference; *P* value	–4.1 (1.9) (–7.9 to –0.3); 0.034	
MI T Score		
Baseline		
LS mean (SE) 95% CI of LS mean	64.9 (1.9) [redacted]	66.6 (2.0) [redacted]
LS mean difference from placebo (SE) 95% CI of LS mean difference	[redacted]	
Week 13		
LS mean (SE)[b] 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE) 95% CI of LS mean difference	[redacted]	
Week 13 change from baseline		
LS mean (SE) 95% CI of LS mean	–4.1 (1.7) (–7.5 to –0.6)	0.3 (1.9) (–3.4 to 4.0)
LS mean difference from placebo (SE) 95% CI of LS mean difference; *P* value	–4.4 (2.1) (–8.5 to –0.2); 0.038	
BRI T Score		
Baseline		
LS mean (SE) 95% CI of LS mean	59.6 (2.5) (54.6 to 64.6)	56.9 (2.6) (51.7 to 62.1)
LS mean difference from placebo (SE) 95% CI of LS mean difference	2.7 (3.0) (–3.4 to 8.8)	
Week 13		
LS mean (SE)[b] 95% CI of LS mean	55.5 (2.7) (49.5 to 60.5)	57.1 (2.8) (51.5 to 62.8)
LS mean difference from placebo (SE) 95% CI of LS mean difference	–2.1 (3.3) (–8.7 to 4.5)	

PKU-016		
	SAP + Diet (n = 36)	Placebo + Diet (n = 29)
Week 13 change from baseline		
LS mean (SE) 95% CI of LS mean	–4.3 (1.4) (–7.1 to –1.4)	–0.9 (1.5) (–3.8 to 2.1)
LS mean difference from placebo (SE) 95% CI of LS mean difference; *P* value	–3.4 (1.7) (–6.8 to 0.0); 0.053	

ANCOVA = analysis of covariance; BRI = Behavioral Regulation Index; BRIEF = Behavior Rating Inventory of Executive Function; CI = confidence interval; GEC = Global Executive Composite; LS = least squares; MI = Metacognition Index; SAP = sapropterin; SE = standard error.

Notes:

- The BRIEF-Parent contains 86 items within 8 clinical scales that measure different aspects in 3 index scales (GEC, MI, and BRI). The BRIEF-Parent was used in patients < 18 years of age.
- ANCOVA model included baseline BRIEF T score, treatment, age group, ADHD symptom, and ADHD medication.
- *P* value determined by t test versus placebo from ANCOVA model.

Source: Clinical Study Report PKU-016.[10]

TABLE 16: STUDY PKU-016: BRIEF-A T SCORES IN PHE RESPONDERS ≥ 18 YEARS OF AGE

PKU-016		
	SAP + Diet (n = 25)	Placebo + Diet (n = 28)
GEC T Score		
Baseline		
LS mean (SE) 95% CI of LS mean	55.4 (3.8) [redacted]	59.2 (3.1) [redacted]
LS mean difference from placebo (SE) 95% CI of LS mean difference	[redacted]	
Week 13		
LS mean (SE) 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE) 95% CI of LS mean difference	[redacted]	
Week 13 change from baseline		
LS mean (SE) 95% CI of LS mean	–9.1 (2.7) (–14.6 to –3.5)	–8.1 (2.2) (–12.6 to –3.6)
LS mean difference from placebo (SE) 95% CI of LS mean difference; *P* value	–1.0 (2.2) (–5.5 to 3.6); 0.661	
MI T Score		
Baseline		
LS mean (SE) 95% CI of LS mean	54.2 (4.0) [redacted]	60.2 (3.3) [redacted]
LS mean difference from placebo (SE) 95% CI of LS mean difference	[redacted]	
Week 13		
LS mean (SE)[b] 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE) 95% CI of LS mean difference	[redacted]	

PKU-016		
	SAP + Diet (n = 25)	**Placebo + Diet (n = 28)**
Week 13 change from baseline		
LS mean (SE) 95% CI of LS mean	–7.9 (2.8) (–13.5; –2.2)	–7.3 (2.3) (–11.9; –2.8)
LS mean difference from placebo (SE) 95% CI of LS mean difference; *P* value	–0.5 (2.3) (–5.3; 4.2); 0.824	
BRI T Score		
Baseline		
LS mean (SE) 95% CI of LS mean	56.3 (3.6) [redacted]	56.3 (2.9) [redacted]
LS mean difference from placebo (SE) 95% CI of LS mean difference	[redacted]	
Week 13		
LS mean (SE)[b] 95% CI of LS mean	[redacted]	[redacted]
LS mean difference from placebo (SE) 95% CI of LS mean difference	[redacted]	
Week 13 change from baseline		
LS mean (SE) 95% CI of LS mean	–8.9 (2.5) (–14.0 to –3.9)	–7.2 (2.0) (–11.3 to –3.1)
LS mean difference from placebo (SE) 95% CI of LS mean difference; *P* value	–1.7 (2.0) (–5.8 to 2.3); 0.396	

ADHD = attention-deficit/hyperactivity disorder; ANCOVA = analysis of covariance; BRI = Behavioral Regulation Index; BRIEF-A = Behavior Rating Inventory of Executive Function–Adult; CI = confidence interval; GEC = Global Executive Composite; LS = least squares; MI = Metacognition Index; SAP = sapropterin; SE = standard error.

Notes:

- The BRIEF-A contains 75 items within 9 clinical scales distributed among 3 index scales (GEC, MI, and BRI). The BRIEF-A was used in patients ≥ 18 years of age.
- ANCOVA model included baseline BRIEF T score, treatment, age group, ADHD symptom, and ADHD medication.
- *P* value determined by t test versus placebo from ANCOVA model.

Source: Clinical Study Report PKU-016.[10]

TABLE 17: STUDY PKU-016: MEAN PHE BLOOD LEVELS IN PHE RESPONDERS

PKU-016		
	SAP + Diet (n = 61)	**Placebo + Diet (n = 57)**
Phe Blood Level (µmol/L)		
Baseline		
n Mean (SD) 95% CI	61 680.2 (435.44) [redacted]	57 789.5 (464.97) [redacted]
Week 4		
n Mean (SD) 95% CI	[redacted]	[redacted]
Week 8		
n Mean (SD) 95% CI	[redacted]	[redacted]

PKU-016		
	SAP + Diet (n = 61)	Placebo + Diet (n = 57)
Week 13		
n Mean (SD) 95% CI	[redacted]	[redacted]
Week 17		
n Mean (SD) 95% CI	[redacted]	[redacted]
Week 21		
n Mean (SD) 95% CI	[redacted]	[redacted]
Week 26		
n Mean (SD) 95% CI	[redacted]	[redacted]

Phe = phenylalanine; SAP = sapropterin; SD = standard deviation.
Note: No statistical comparisons were conducted between treatment groups.
Source: Clinical Study Report PKU-016.[10]

TABLE 18: SPARK STUDY: MEAN PHE BLOOD LEVELS (INTENTION-TO-TREAT POPULATION)

SPARK		
	SAP + Diet (n = 27)	Diet Alone (n = 29)
Time Point		
Week 2		
Baseline n Mean (SD)	[redacted]	[redacted]
Week 2		
n Mean (SD)	[redacted]	[redacted]
Week 4		
n Mean (SD)	[redacted]	[redacted]
Week 6		
n Mean (SD)	[redacted]	[redacted]
Week 8		
n Mean (SD)	[redacted]	[redacted]
Week 10		
n Mean (SD)	[redacted]	[redacted]
Week 12		
n Mean (SD)	[redacted]	[redacted]

SPARK		
	SAP + Diet (n = 27)	Diet Alone (n = 29)
Week 14		
n Mean (SD)	[redacted]	[redacted]
Week 16		
n Mean (SD)	[redacted]	[redacted]
Week 18		
n Mean (SD)	[redacted]	[redacted]
Week 20		
n Mean (SD)	[redacted]	[redacted]
Week 22		
n Mean (SD)	[redacted]	[redacted]
Week 24		
n Mean (SD)	[redacted]	[redacted]
Week 26		
n Mean (SD)	[redacted]	[redacted]
Change From Baseline and Treatment Difference at Week 26		
Week 26		
Adjusted mean (SE) (95% CI)	[redacted]	[redacted]
Adjusted difference between groups (SE) (95% CI); *P* value	[redacted]	

ANCOVA = analysis of covariance; CI = confidence interval; ITT = intention-to-treat; Phe = phenylalanine; SAP = sapropterin; SD = standard deviation; SE = standard error.
Note: Phe blood levels are assessed from the mean of filter-paper Phe blood levels.
[a] Adjusted means estimates, SEs, and 95% CIs were based on the repeated-measures ANCOVA, which included the fixed categorical effects of treatment, age group, visit, treatment x visit as well as the continuous fixed covariates of baseline dietary Phe tolerance and baseline mean filter-paper blood Phe level, with a compound symmetry matrix for the within-subject error variance-covariance.
Source: Clinical Study Report SPARK study.[11]

TABLE 19: SPARK STUDY: PROPORTION OF PATIENTS MAINTAINING PHE BLOOD LEVELS WITHIN THERAPEUTIC TARGET (INTENTION-TO-TREAT POPULATION)

SPARK		
	SAP + Diet (n = 27)	Diet Alone (n = 29)
All Phe Blood Levels ≤ 360 µmol/L, n (%)	[redacted]	[redacted]
All Phe Blood Levels ≥ 120 and ≤ 360 µmol/L, n (%)	9 (33.3)	3 (10.3)

Phe = phenylalanine; SAP = sapropterin.
Note: Phe blood levels are assessed from the mean of filter-paper Phe blood levels.
Source: Clinical Study Report SPARK study.[11]

TABLE 20: SPARK STUDY: ADJUSTED MEAN TREATMENT DIFFERENCE IN DIETARY PHE TOLERANCE BASED ON PRESCRIBED PHE (MG/KG/DAY) (INTENTION-TO-TREAT POPULATION)

SPARK		
	SAP + Diet (n = 27)	Diet Alone (n = 27)
Time Point		
Week 2		
Adjusted mean (SE) (95% CI)	35.5 (3.8) (27.9 to 43.1)	42.8 (4.1) (34.6 to 51.0)
Week 4		
Adjusted mean (SE) (95% CI)	40.2 (3.9) (32.4 to 47.9)	39.4 (4.2) (31.1 to 47.6)
Week 6		
Adjusted mean (SE) (95% CI)	51.7 (3.9) (44.0 to 59.5)	42.3 (4.0) (34.3 to 50.2)
Week 8		
Adjusted mean (SE) (95% CI)	58.3 (4.0) (50.4 to 66.2)	43.2 (4.0) (35.2 to 51.2)
Week 10		
Adjusted mean (SE) (95% CI)	60.2 (4.1) (52.1 to 68.2)	44.9 (4.0) (36.9 to 52.8)
Week 12		
Adjusted mean (SE) (95% CI)	65.1 (4.2) (56.8 to 73.5)	45.0 (4.3) (36.5 to 53.5)
Week 14		
Adjusted mean (SE) (95% CI)	64.6 (4.0) (56.6 to 72.5)	47.0 (4.0) (39.1 to 54.9)
Week 16		
Adjusted mean (SE) (95% CI)	69.2 (4.1) (61.1 to 77.2)	49.7 (4.1) (41.5 to 57.8)
Week 18		
Adjusted mean (SE) (95% CI)	70.5 (4.1) (62.4 to 78.5)	48.7 (4.1) (40.5 to 56.8)
Week 20		
Adjusted mean (SE) (95% CI)	76.0 (4.0) (68.0 to 84.0)	50.0 (4.1) (41.8 to 58.1)
Week 22		
Adjusted mean (SE) (95% CI)	74.4 (4.0) (66.5 to 82.3)	50.2 (4.1) (42.0 to 58.3)
Week 24		
Adjusted mean (SE) (95% CI)	75.6 (3.9) (67.8 to 83.4)	51.2 (4.3) (42.7 to 59.7)

SPARK		
	SAP + Diet (n = 27)	Diet Alone (n = 27)
Week 26		
Adjusted mean (SE) (95% CI)	80.6 (4.2) (72.3 to 88.8)	50.1 (4.3) (41.6 to 58.6)
Adjusted difference between groups (SE) (95% CI); *P* value	30.5 (6.0) (18.7 to 42.3); < 0.001	

ANCOVA = analysis of covariance; CI = confidence interval; Phe = phenylalanine; SAP = sapropterin; SE = standard error.
Notes:
- Phe tolerance was defined as the prescribed amount of dietary Phe (mg/kg/day) while maintaining the mean filter-paper blood Phe levels within the target ranges (≥ 120 to < 360 µmol/L).
- Adjusted means estimates, SEs, and 95% CIs were based on the repeated-measures ANCOVA, which included the fixed categorical effects of treatment, age group, visit, treatment x visit, as well as the continuous fixed covariates of baseline dietary Phe tolerance and baseline mean filter-paper blood Phe level, with a compound symmetry matrix for the within-subject error variance-covariance.

Source: Clinical Study Report, SPARK study.[11]

Table 21: SPARK Study — Mean Change in Dietary Phe Tolerance (mg/kg/day) From Baseline to End of Treatment (Intention-To-Treat Population)

SPARK		
	SAP + Diet (n = 27)	Diet Alone (n = 29)
Baseline		
n Mean (SD)	27 37.1 (17.3)	29 35.8 (20.9)
Week 26 LOCF		
n Mean (SD)	27 74.0 (38.6)	27 49.8 (24.2)
Week 26 LOCF — Baseline		
n Mean (SD) (95% CI); *P* value	27 36.9 (27.3) (26.1 to 47.7); < 0.001	27 13.1 (19.6) (5.4 to 20.9); 0.002

CI = confidence interval; LOCF = last observation carried forward; Phe = phenylalanine; SAP = sapropterin; SD = standard deviation.
Note: *P* value was derived by paired t test.
Source: Clinical Study Report SPARK study.[11]

Table 22: SPARK Study — Adjusted Means Over Time and Treatment Difference at Week 26 in Change from Baseline in Growth Parameters (Intention-to-Treat Population)

SPARK		
	SAP + Diet (n = 27)	Diet Alone (n = 29)
Number of patients included in model	**26**	**27**
Height SDS		
Week 26		
Adjusted mean (SE) (95% CI)	[redacted]	[redacted]
Adjusted difference (SE)	[redacted]	

SPARK		
	SAP + Diet (n = 27)	Diet Alone (n = 29)
(95% CI); *P* value	[redacted]	
Weight SDS		
Week 26		
Adjusted mean (SE) (95% CI)	[redacted]	[redacted]
Adjusted difference (SE) (95% CI); *P* value	[redacted]	
BMI SDS		
Week 26		
Adjusted mean (SE) (95% CI)	[redacted]	[redacted]
Adjusted difference (SE) (95% CI); *P* value	[redacted]	
Max Occipital-Frontal Head Circumference SDS		
Week 26		
Adjusted mean (SE) (95% CI)	[redacted]	[redacted]
Adjusted difference (SE) (95% CI); *P* value	[redacted]	

BMI = body mass index; CI = confidence interval; SAP = sapropterin; SDS = standard deviation score; SE = standard error.
Note: Adjusted means estimates, SEs, and 95% CIs were based on the repeated-measures ANCOVA, which included the fixed categorical effects of treatment, age group, visit, treatment x visit as well as the continuous fixed covariates of baseline dietary Phe tolerance and baseline mean filter-paper blood Phe level, with a compound symmetry matrix for the within-subject error variance-covariance.
Source: Clinical Study Report, SPARK study.[11]

TABLE 23: SPARK STUDY: NEUROMOTOR DEVELOPMENTAL MILESTONES (INTENTION-TO-TREAT POPULATION)

SPARK		
	SAP + Diet (n = 27)	Diet Alone (n = 29)
Area of Assessment: Fine Motor		
Baseline, n (%)		
n Normal Abnormal	[redacted]	[redacted]
Week 12, n (%)		
n Normal Abnormal	[redacted]	[redacted]
Week 26, n (%)		
n Normal Abnormal	[redacted]	[redacted]
P value[a]	[redacted]	
Area of Assessment: Gross Motor		
Baseline		
n	[redacted]	[redacted]

SPARK		
	SAP + Diet (n = 27)	Diet Alone (n = 29)
Normal Abnormal	[redacted]	[redacted]
Week 12		
n Normal Abnormal	[redacted]	[redacted]
Week 26		
n Normal Abnormal	[redacted]	[redacted]
P value[a]	[redacted]	
Area of Assessment: Language		
Baseline		
n Normal Abnormal	[redacted]	[redacted]
Week 12		
n Normal Abnormal	[redacted]	[redacted]
Week 26		
n Normal Abnormal	[redacted]	[redacted]
P value[a]	[redacted]	
Area of Assessment: Personal — Social		
Baseline		
n Normal Abnormal	[redacted]	[redacted]
Week 12		
n Normal Abnormal	[redacted]	[redacted]
Week 26		
n Normal Abnormal	[redacted]	[redacted]
P value[a]	[redacted]	

ITT = intention-to-treat; SAP = sapropterin.
Note: Neuromotor developmental milestones were assessed by standardized developmental milestones using a parent/guardian report form (Denver Developmental Screening Test). Neurodevelopmental status assessment was performed using the standardized Bayley Scales of Infant and Toddler Development, Third Edition (Bayley-III) for patients aged < 3.5 years and the Wechsler Preschool and Primary Scale of Intelligence for patients aged ≥ 3.5 and < 4 years.
[a] *P* value for comparison of treatment groups was derived from Chi square statistics.
Source: Clinical Study Report, SPARK study.[11]

APPENDIX 5: VALIDITY OF OUTCOMES

Aim

To summarize the validity of the following outcome measures:

- Definition of phenylalanine (Phe) response
- Attention Deficit Hyperactivity Disorder Rating Scale (ADHD-RS) in Attentive subscale
- Adult Attention Deficit Hyperactivity Disorder Self-Report Scale (ASRS) Inattention subscale
- Behavior Rating Inventory of Executive Function (BRIEF).

Findings

Phe Response

It is well known that deleterious effects on neurocognition, intelligence, and executive functioning are associated with high blood and brain concentrations of Phe in patients diagnosed with phenylketonuria (PKU).[22] The purpose of dietary Phe restrictions is to lower blood and brain Phe levels, thereby reducing the risk of damage to the brain.

There is speculation that fluctuations in Phe levels are of potential significance in their relation to intelligence and neurocognition; however, there remains no current definition regarding Phe fluctuations and these have been measured in many different ways (e.g., standard deviation [SD], regression analysis of Phe concentrations, standard error of the estimate [SEE], and mean and its accompanying SD of index dietary control [IDC] measured by six-month mean Phe values).[22] In the review by Cleary et al.,[22] it was noted that a number of studies have reported that the highest concentrations of Phe were seen in the morning in patients with PKU when observing diurnal variation of blood Phe levels. Other studies identified in this review have also reported that there may be up to 400% variation in day-to-day blood Phe levels in adults with well-controlled PKU, that blood Phe levels fluctuate to a larger extent in patients with PKU when compared with healthy patients, that blood Phe concentrations increase with age while there is uncertainty regarding whether fluctuations decrease with age, and that fluctuations may be influenced by genotype, rates of growth, dietary adherence, diet, and illness.[22]

With regard to the impact of blood Phe fluctuations on the brain, Cleary et al.[22] reported that healthy controls have approximately equal concentrations of blood and brain Phe concentrations, while increases in Phe concentrations in the blood are higher than increases in the brain in patients with PKU. In addition, they noted that peaks of Phe levels last longer, are not as steep, and occur later in the brain when compared with blood in these patients.[22] There is contradicting evidence regarding the effects on neurocognition, measures of brain activity, and Phe fluctuations in patients with PKU, with some studies showing correlations between Phe fluctuations and deficits in executive functioning, cognition, and intelligence while others finding no associations.[22] Hood et al.[23] reported that Phe variability was a better predictor of cognitive performance when compared with the various other aforementioned indices of Phe control, along with being a better predictor of executive functioning in children aged five years and older when compared with patients younger than five years of age. Hood et al.[24] retrospectively analyzed the microstructural white matter integrity using mean diffusivity from diffusion tensor imaging and various Phe indices to measure blood Phe concentrations in early and continuously treated patients with PKU in order to determine if prolonged exposure to both high and variable levels of blood Phe correlated with white matter compromise. The authors reported that microstructural white matter integrity compromise was correlated with mean Phe, the IDC, mean exposure, and the SD of exposure, indicating that high and variable blood Phe concentrations were predictors of white matter

compromise in this population.[24] Viau et al.[30] suggested that there was an association between measures of intelligence and the quality of metabolic blood Phe control during certain developmental periods. Perceptual reasoning appeared to be strongly associated with proper blood Phe control during their zero-to-six and seven-to-12 years of age periods, with specific areas of verbal comprehension being affected by increases of blood Phe levels in children aged zero to six years.[30] However, the evidence obtained by Viau et al.[30] was not supportive that blood Phe level variability was good indicator of intelligence. That being said, all of these results indicate that continual blood Phe concentrations should be monitored and controlled in childhood and that this should continue throughout the life of the patient with PKU.[22-24]

ADHD-RS-IV and ASRS

The ADHD-RS-IV scale is based on Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) criteria used to assess ADHD symptoms in children and is completed by a parent or guardian.[25] The full scale comprises 18 items that are separated into two subscales: an Inattention subscale and a Hyperactivity-Impulsivity subscale. Each of the subscales is composed of nine items that assess the frequency of ADHD symptoms, with each item rated on a four-point Likert frequency scale (0 = never or rarely; 1 = sometimes; 2 = often; and 3 = very often).[25] The recall period is one month. A higher score corresponds with worse severity of ADHD. This scale has been observed to be reliable and valid in children with ADHD.[25] In the Inattention subscale, scores range from 0 to 27, with greater inattentive severity measured with higher scores.[25]

The full ASRS scale is also an 18-item instrument that was developed to assess ADHD symptom and behaviour frequency in adults.[25] It is a self-rated instrument with five response options for each of the 18 items (0 = never; 1 = rarely; 2 = sometimes; 3 = often; and 4 = very often), which are subdivided into two subsections: Part A, which measures inattention symptoms, and Part B, which measures hyperactivity and impulsivity symptoms.[25] It has a recall period of one month. Scores range from 0 to 36 in the ASRS Inattention score, with higher scores indicating a greater amount of inattentive severity.[25]

Wyrwich et al.[25] used only the summed nine-item Inattention subscale in both of these scales in order to ascertain its validity, reliability, and responsiveness in patients with PKU from PKU-016 who were undergoing treatment with SAP. The Inattention subscales of both the ADHD-RS-IV and ASRS were observed to be valid, reliable, responsive, and robust measures of inattention in children with PKU aged eight years or older.[25] No MCIDs for children with PKU were available for either of the subscale scores.

BRIEF

The BRIEF was developed to systematically and quantitatively assess everyday behaviour, by recording common descriptors of executive function-related behaviours. The BRIEF–Adult Version (BRIEF-A) is a self-reported questionnaire for patients aged 18 years or older, whereas the Parent Form of the BRIEF (BRIEF-Parent) is a parent- or legal guardian–reported questionnaire for patients younger than 18 years. The BRIEF-Parent contains 86 items within eight clinical scales that measure different aspects in three index scales including Behavioral Regulation Index (BRI), Metacognition Index (MI), and Global Executive Composite (GEC). The BRIEF includes three-point Likert scale items that indicate the extent to which the child's behaviour never occurred (1), sometimes occurred (2), or occurred often (3). The raw total score is calculated by totalling the scores of 72 items constituting eight subscales. The BRIEF-A is composed of 75 items within nine clinical scales and three index scales. The raw total score is calculated by totalling the scores of 70 items constituting nine subscales.

No relevant information on the validity or reliability of the BRIEF instrument could be located following a focused medical literature search.

TABLE 24: INSTRUMENT DESCRIPTIONS

Instrument	Type	Validated	MCID	References
Phe reference	Can be measured as: – SD – regression analysis of Phe concentration – SEE – mean and its accompanying SD IDC measured by 6-month mean Phe values	Yes	Unknown	[22-24]
ADHD-RS-IV	Composed of 2 subscales: – Inattention – Hyperactivity-Impulsivity Measured on 5-point Likert frequency scale: – 0 = never – 1 = rarely – 2 = sometimes – 3 = often – 4 = very often	Yes	Unknown (for children with PKU)	[25]
ASRS	Composed of 2 subscales: – Inattention – Hyperactivity-Impulsivity Measured on 4-point Likert frequency scale: – 0 = never or rarely – 1 = sometimes – 2 = often – 3 = very often	Yes	Unknown (for adults with PKU)	[25]
BRIEF	Composed of 3 index scales: – BRI – MI – GEC Measured on 3-point Likert frequency scale: – 1 = never occurred – 2 = sometimes occurred – 3 = often occurred	Unknown	Unknown	None

ADHD = attention-deficit/hyperactivity disorder; ADHD-RS-IV = Attention Deficit Hyperactivity Disorder Rating Scale-IV; ASRS = Adult Attention Deficit Hyperactivity Disorder Self-Report Scale; BRI = Behavior Regulation Index; BRIEF = Behavior Rating Inventory of Executive Function; GEC = Global Executive Composite; IDC = index dietary control; MI = Metacognition Index; Phe = phenylalanine; PKU = phenylketonuria; SD = standard deviation; SEE = standard error of the estimate.

APPENDIX 6: SUMMARY OF OTHER STUDIES

Aim

To summarize the results from the PKUDOS[15] and KAMPER[18] registry studies and the PKU-015 trial.[14] The PKUDOS registry study and PKU-015 trial were submitted as new evidence by the manufacturer but did not meet the selection criteria for inclusion in the systematic review. The following summary is based on published data from the corresponding journal articles and any accompanying supplemental information.

a) The Phenylketonuria Demographics, Outcomes and Safety (PKUDOS) Registry[15]

The PKUDOS registry was a voluntary, phase 4, observational study designed to capture long-term (up to 15 years) effectiveness and safety data in patients who have been diagnosed with PKU and who have been treated with, are currently on, or plan to receive treatment with sapropterin (SAP) within 90 days of enrolment. The statistical analysis focused on two distinct populations of SAP treatment: patients with uninterrupted use (who have had continuous treatment) and patients with short-term use (SAP treatment for three months or less; with dose gaps allowed within the three months of exposure). Blood Phe levels, prescribed and actual dietary Phe amounts, blood tyrosine levels, and adverse events (AEs) were obtained and reported in this registry. Disease severity and the rationale for prescribing SAP were not included. In addition, information pertaining to missing SAP dosages or outliers of the 5 to 25 mg/kg /day range and outliers in blood Phe levels at birth or other ages (greater than 5,400 µmol/L and greater than 3,000 µmol/L, respectively) was excluded.

Findings

The PKUDOS registry contained 1,224 patients from 52 active sites in the United States as of June 2013; this analysis included information on 1,189 of these patients. Patients were predominantly white (89%), 8% were younger than four years of age, and 52% were female. Upon registry entry, a small proportion of patients were considered disabled (2%), about half of the population was in school, and about 22% were employed in some capacity. Detailed patient characteristics and demographics are provided in Table 25.

TABLE 25: PATIENT CHARACTERISTICS AND DEMOGRAPHICS

	Uninterrupted Use N = 504	Short-Term Use N = 211
Age at First SAP Dose, n (%)		
Birth to < 4 years	48 (9.5)	10 (4.7)
4 to < 13 years	206 (40.9)	60 (28.4)
13 to < 25 years	144 (28.6)	86 (40.8)
≥ 25 years	105 (20.8)	55 (26.1)
Missing	1 (0.2)	0
Gender, n (%)		
Males	253 (50.2)	94 (44.5)
Race,[a] n (%)		
White	437 (86.7)	191 (90.5)
Black/African-American	1 (0.2)	6 (2.8)
Asian	8 (1.6)	2 (0.9)
Other	10 (2.0)	2 (0.9)
Education at Study Entry, n (%)		
Kindergarten	223 (44.2)	109 (51.7)

	Uninterrupted Use N = 504	Short-Term Use N = 211
Primary: grades 1 to 2	39 (7.7)	21 (10.0)
Elementary: grades 3 to 5	62 (12.3)	10 (4.7)
Intermediate: grades 6 to 8	45 (8.9)	15 (7.1)
Secondary school: grades 9 to 12	66 (13.1)	25 (11.8)
College	39 (7.7)	17 (8.1)
Vocational	4 (0.8)	3 (1.4)
Master's	1 (0.2)	2 (0.9)
Doctoral	2 (0.4)	0
Missing	23 (4.6)	9 (4.3)
Employment at Registry Entry, n (%)		
Unemployed	316 (62.7)	125 (59.2)
Disabled	11 (2.2)	4 (1.9)
Employed < 20 hours/week	14 (2.8)	8 (3.8)
Employed 20 to 39 hours/week	26 (5.2)	13 (6.2)
Employed > 39 hours/week	66 (13.1)	36 (17.1)
Missing	71 (14.1)	25 (11.8)

SAP = sapropterin.
[a] Categories are not mutually exclusive.

Clinical Effectiveness

Total Cohort (1,189 Patients)

With regard to SAP use, 42% of the cohort had continuous use (uninterrupted use) and 18% had short-term use. In the uninterrupted use population, the duration of SAP use ranged from 0.9 to 7.2 years, with a first recorded median dose of 20 mg/kg/day and a median duration of four years. In the short-term use population, the duration of SAP use ranged from 0.03 to 3 months, with a first recorded median dose of 20 mg/kg/day. The SAP dose did not change throughout the time period analyzed in either population.

Blood Phe Levels

With regard to baseline peak Phe blood levels, 56% of all patients had peak levels less than or equal to 1,200 µmol/L, 34% had peak levels between 600 and 1,200 µmol/L, and the remainder (9%) had peak levels less than 600 µmol/L. For patients on uninterrupted use, 43%, 42%, and 15% had peak levels less than or equal to 1,200 µmol/L, between 600 and 1,200 µmol/L, and less than 600 µmol/L, respectively. For patients on short-term use, 82%, 16%, and 3% had peak levels less than or equal to 1,200 µmol/L, between 600 and 1,200 µmol/L, and less than 600 µmol/L, respectively.

When compared with baseline, statistically significant and sustained decreases in blood Phe levels ranging from –25% to –34% were observed in patients with uninterrupted SAP use over the time period of one to six years. In patients on short-term SAP use, decreases in blood Phe levels were smaller and ranged from –1% to –9% over the one- to six-year time period and none were statistically significantly different from baseline. In the uninterrupted use population, 71% of patients were SAP-responsive compared with 27% of short-term use patients; this resulted in a statistically significant difference (P = 0.001). Details of Phe blood levels over the six years of SAP exposure are provided in Table 26.

TABLE 26: BLOOD PHE LEVELS: PRE-SAPROPTERIN TO SIX YEARS OF EXPOSURE

	Time Period on Sapropterin[a]						
	Baseline[b]	0 to ≤ 1 Year	> 1 to ≤ 2 Years	> 2 to ≤ 3 Years	> 3 to ≤ 4 Years	> 4 to ≤ 5 Years	> 5 to ≤ 6 Years
Uninterrupted Use							
N[c]	128	318	333	312	237	161	48
Mean (SD), µmol/L	591 (382)	418 (333)	415 (299)	432 (298)	441 (288)	421 (265)	392 (239)
Change in Phe levels, µmol/L (%)[d]	-	−173 (29)	−176 (30)	−160 (27)	−150 (25)	−170 (29)	−199 (34)
P value	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0009
Short-Term Use[e]							
N[c]	66	129	143	135	106	64	30
Mean (SD), µmol/L	830 (476)	792 (461)	752 (413)	800 (412)	817 (380)	823 (380)	759 (366)
Change in Phe levels, µmol/L (%)[d]	-	−38 (5)	−78 (9)	−30 (4)	−13 (2)	−7 (1)	–71 (9)
P value	-	NS	NS	NS	NS	NS	NS
Comparison Between Uninterrupted Use and Short-Term Use Populations							
Mean difference, µmol/L[f]	239	374	338	369	376	402	366
P value	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001

NS = not significant; Phe = phenylalanine; SD = standard deviation.
[a] Median used if the patient had more than one value during a time period. These are time periods of sapropterin exposure.
[b] Baseline: pre-sapropterin exposure.
[c] Number of patients in the time period who had both blood Phe and dietary Phe values in the PKUDOS registry.
[d] Percentage change compared with baseline pre-sapropterin mean value.
[e] Short-term use is the use of sapropterin for ≤ 3 months, with dose gaps allowed within 3 months of exposure.
[f] Patients may have multiple blood analyses but are only counted once in the total; therefore, the individual columns may not sum to the all patients column.

Dietary Phe Intake

With regard to increases in actual dietary Phe intake, in the uninterrupted use population, average actual increases from 1,000 mg/day (baseline) to 1,539 mg/day (six years) were observed. For the short-term use population, average actual decreases from 815 mg/day to 725 mg/day were observed during this same time period. Details of the prescribed, actual, and per cent change from baseline results are provided in Table 27.

TABLE 27: DIETARY PHE: PRESCRIBED, ACTUAL, AND PERCENT CHANGE IN ACTUAL DIETARY PHE INTAKE

	Time Period on Sapropterin							
	Baseline[a]	0 to ≤ 1 Year	> 1 to ≤ 2 Years	> 2 to ≤ 3 Years	> 3 to ≤ 4 Years	> 4 to ≤ 5 Years	> 5 to ≤ 6 Years	> 6 to ≤ 7 Years
Uninterrupted Use								
Prescribed dietary Phe								
N	62	171	151	114	93	62	15	3
Mean (SD)	565 (620)	772 (614)	855 (682)	811 (677)	814 (488)	885 (649)	1,112 (771)	1,198 (232)
Actual dietary Phe intake								
N	79	206	164	139	107	63	19	5
Mean (SD)	1,000 (959)	1,172 (894)	1,153 (848)	1,311 (924)	1,153 (860)	1,089 (670)	1,197 (667)	1,539 (840)
Actual dietary Phe								
Per cent change from baseline	-	17	15	31	15	9	20	54
Short-Term Use								
Prescribed dietary Phe								
N	39	76	77	73	52	34	9	3
Mean (SD)	403 (383)	365 (156)	366 (186)	341 (152)	387 (267)	434 (607)	392 (183)	370 (113)
Actual dietary Phe intake								
N	42	56	69	60	48	29	9	3
Mean (SD)	815 (739)	722 (654)	614 (543)	586 (570)	743 (876)	586 (746)	636 (529)	725 (855)
Actual dietary Phe								
Per cent change from baseline	-	−11	−25	−25	−9	−28	−22	−11

Phe = phenylalanine; SD = standard deviation.
Notes:
- Not all patients have data at each time period.
- Median used if patient had more than one value during a specific time period.

[a] Baseline: pre-sapropterin exposure.

Children Aged Four Years or Younger

For children aged four years or younger, the mean ages were similar between the uninterrupted use and short-term use groups at 1.8 years (standard deviation [SD] of 1.2 years) and 1.9 years (SD of 1.2 years), respectively.

Blood Phe Levels

In the uninterrupted use group, the mean of median blood Phe levels decreased by 22% after one year, 26% after two years, and 42% after three years from a baseline pre-SAP level of 259 µmol/L. In the short-term use group, the mean of median blood Phe levels decreased by 13% after one year, 50% after two years, and 39% after three years from a baseline pre-SAP level of 369 µmol/L. Details of the blood Phe levels in these populations are provided in Table 28.

TABLE 28: BLOOD PHE LEVELS IN PATIENTS AGED ZERO TO FOUR YEARS

	Time Period on Sapropterin			
	Baseline[a]	0 to ≤ 1 Year	>1 to ≤ 2 Years	>2 to ≤ 3 Years
Uninterrupted Use (Blood Phe Levels, µmol/L[b])				
Venipuncture method				
N[c] Mean (SD) Median	5 283 (143) 227	8 278 (182) 197	8 232 (186) 162	4 174 (79) 156
Filter-paper method				
N Mean (SD) Median	15 335 (177) 290	25 227 (103) 205	10 208 (74) 220	4 136 (83) 144
Overall mean of medians	259	201	191	150
Change from baseline median/mean (%)	-	–22	–26	–42
Short-Term Use (Blood Phe Levels, µmol/L[b])				
Venipuncture method				
N Mean (SD) Median	2 490 (437) 490	4 421 (186) 366	2 179 (110) 179	-
Filter-paper method				
N Mean (SD) Median	4 297 (138) 248	7 332 (128) 279	3 234 (133) 188	1 224 224
Overall mean of medians	369	323	184	224
Change from baseline median/mean (%)	-	–13	–50	–39

Phe = phenylalanine; SD = standard deviation.
[a] Baseline: pre-sapropterin exposure.
[b] If more than one value available per time period, then the median was used.
[c] Number of patients (some had multiple samples collected).

Dietary Phe Intake

For children in the uninterrupted use group, the actual dietary Phe intake increased by 97% after one year, 141% after two years, and 26% after three years from a baseline median pre-SAP amount of 292 mg/day. For children in the short-term use group, the actual dietary Phe intake increased by 110% after one year, 99% after two years, and 262% after three years from a median pre-SAP baseline amount of 183 mg/day. Details regarding the prescribed, actual, and median change from baseline for these populations are provided in Table 29.

TABLE 29: DIETARY PHE IN PATIENTS AGED ZERO TO FOUR YEARS: PRESCRIBED, ACTUAL, AND CHANGE IN ACTUAL DIETARY PHE INTAKE

	Time Interval					
	Baseline[a]	0 to 3 Months	0 to ≤ 1 Year	> 1 to ≤ 2 Years	> 2 to ≤ 3 Years	> 3 to ≤ 4 Years
Uninterrupted Use						
Prescribed dietary Phe						
N Mean (SD) Median	8 371 (103) 359	9 524 (384) 525	22 580 (366) 525	9 625 (397) 400	3 400 (48) 375	-
Actual dietary Phe intake						
N Mean (SD) Median	6 343 (118) 292	9 450 (218) 472	21 545 (265) 575	8 850 (502) 703	3 328 (86) 367	-
Actual dietary Phe — % change from baseline						
Median Mean	-	62 (31)	97 (59)	141 (148)	26 (-4)	-
Short-Term Use						
Prescribed dietary Phe						
N Mean (SD) Median	3 182 (43) 180	1 250 (0) 250	3 345 (135) 285	2 413 (124) 413	-	-
Actual dietary Phe intake						
N Mean (SD) Median	2 183 (60) 183	1 250 (0) 250	4 380 (126) 385	2 364 (116) 364	1 663 (0) 663	-
Actual dietary Phe — % change from baseline						
Median Mean	-	37 (37)	110 (108)	99 (99)	262 (262)	-

Phe = phenylalanine.
[a] Baseline: pre-sapropterin exposure.

Blood Tyrosine

The data for this outcome were sparse, as fewer than 20 patients were assessed per time period; therefore, the results are not provided.

Safety

AEs considered unrelated to SAP use occurred in 31% of the total cohort, whereas AEs that were considered SAP-related occurred in 6% of the total cohort. The most common AEs that occurred in 5% or more of patients involved gastrointestinal, general disorders and administration-site conditions, infections and infestations, and respiratory and nervous system disorders. Actual AEs of interest in these groups were diarrhea and unspecified gastrointestinal disorder (gastrointestinal), dysphonia and rhinorrhea (respiratory), and amnesia, dizziness, memory impairment, headache, migraines, poor sleep quality, and psychomotor hyperactivity (nervous system). AEs that were considered to be SAP-related were observed in 6%, 4%, 5%, and 8% of patients younger than four years, four to 13 years, 13 to younger than 25 years, and 25 years or older, respectively. Of the AEs related to SAP use, most were considered mild or moderate, with only 4% being severe.

SAEs considered unrelated to SAP use occurred in 6% of the total cohort, whereas AEs that were considered SAP-related occurred in 1% of the total cohort. SAEs that were considered to be SAP-related were observed in 0%, 0.2%, 0.6%, and 1.6% of patients younger than four years, four to 13 years, 13 to younger than 25 years, and 25 years or older, respectively. Detailed harms information is provided in terms of System Organ Class in Table 30.

TABLE 30: HARMS

System Organ Class — Preferred Term	Total Unrelated AEs N = 1,189	Total SAP-Related AEs N = 1,189
All AEs, n (%)	**369 (31)**	**74 (6)**
AEs Occurring in ≥ 5% of Patients		
Gastrointestinal disorders	71 (6)	33 (3)
General disorders and administration-site conditions	62 (5)	0
Infections and infestations	173 (15)	0
Respiratory, thoracic, and mediastinal disorders	64 (5)	4 (0)
Nervous system disorders	61 (5)	22 (2)
All Other AEs		
Ear and labyrinth disorders	10 (1)	0
Eye disorders	11 (1)	0
Immune system disorders	9 (1)	0
Injury, poisoning, and procedural complications	47 (4)	0
Investigations	19 (2)	0
Metabolism and nutrition disorders	21 (2)	0
Musculoskeletal and connective tissue disorders	40 (3)	0
Renal and urinary disorders	17 (1)	0
Pregnancy, puerperium, and perinatal conditions	17 (1)	0
Psychiatric disorders	43 (4)	0
Skin and subcutaneous tissue disorders	27 (2)	0
Surgical and medical procedures	18 (2)	0
Uncoded	16 (1)	0
All SAEs, n (%)	**65 (6)**	**10 (1)**
General disorders and administration-site conditions	5 (0.4)	0
Infections and infestations	17 (1)	0
Investigations	4 (0.3)	0
Surgical and medical procedures	10 (0.8)	0

AE = adverse event; SAE = serious adverse event; SAP = sapropterin.

Limitations

As with most registries, the main limitation associated with assessing any results of any analyses is the potential for data to be incomplete, as well as not verified (at the source). In addition, the PKUDOS registry is a drug registry, as opposed to a disease registry, and it therefore lacks a control group of patients who have no experience with SAP. Variability in the blood Phe data is evident and is most likely due to differences in analytical measurements between centres and inconsistency in the number of blood Phe measurements acquired per patient. In addition, there appear to be asymmetrical differences between the mean and median blood and dietary Phe results (although the trends are similar). Another important limitation is the lack of baseline blood Phe levels pre-SAP use, which was not required and

thus not necessarily obtained for each patient. Other limitations included the lack of reporting of SAP compliance, the potential for under-reporting of AEs, and the lack of information regarding the intake of medical food.

Conclusions

It appears that SAP has an acceptable safety and tolerability profile, is associated with significant decreases in blood Phe levels, and improves dietary Phe tolerance in SAP responders (i.e., especially in those patients with uninterrupted SAP use over six years). Clinical effects were not as prominent in the SAP short-term use group as compared with the uninterrupted use group.

b) The Kuvan Adult Maternal Paediatric European Registry (KAMPER)[18]

The KAMPER study is an ongoing, European (where SAP is marketed), multicentre, observational drug registry for patients with hyperphenylalaninemia (HPA). Patients with HPA due to PAH deficiency who meet the inclusion criteria, are four years of age or older, and have mild PKU and mild HPA (phenylalanine levels greater than 360 µmol/L) are being enrolled. In addition, patients with HPA and tetrahydrobiopterin (BH_4) deficiency of any age who meet the inclusion criteria have been and will continue to be enrolled. Study eligibility includes patients currently on SAP treatment at one of the participating centres and who are SAP-responsive (i.e., defined as ≥ 30% reduction in blood Phe levels or are able to lower their levels to a physician-defined therapeutic blood Phe level) or BH_4-responsive. Exclusion criteria included breast-feeding and hypersensitivity to SAP. The maternal subregistry is available to eligible patients who become pregnant while in the study (and continue SAP treatment) and who are pregnant and unable to lower their blood Phe levels with a Phe-restricted diet.

Recruitment into the registry began in December 2009 and will continue until December 2019. Participating countries from which patients were recruited at the time of this interim analysis include Austria, France, Germany, Italy, the Netherlands, Slovakia, and Spain.

The primary outcome is the incidence of AEs and SAEs, in the total population and in specific subgroups of interest (e.g., patients ≥ 65 years of age, pediatric patients [younger than 18 years of age], and patients suffering from renal or hepatic insufficiency). Secondary outcomes include treatment adherence, population characteristics, Phe tolerance, metabolic control, auxological and nutritional outcomes, neurological and neuropsychiatric outcomes, and outcomes related to the maternal subregistry.

The registry plans to enrol 625 patients in order to obtain a population of 500 patients or more for evaluation at the end of the study period.

Findings

This descriptive interim analysis includes 325 patients (91.1% [n = 296] with PAH deficiency and 8.9% [n = 29] with BH_4 deficiency) with available baseline data who enrolled between December 2009 and November 2012. Patients who were aged 18 years or younger comprised 18.2% (n = 59) of the population and one-year follow-up analysis data were available for 180 patients (87.8% [n = 164] with PAH deficiency and 8.9% [n = 16] with BH_4 deficiency). No patients with renal or hepatic insufficiency were enrolled, nor were any elderly patients. The overall median age at baseline was 10.3 years of age, the age range with the largest population was that of four to less than eight years of age (n = 107), and males comprised 51.1% of the overall population. One or more medical conditions were observed in 85 of patients at baseline with 10 patients (n = 7 with PAH and n = 3 with BH_4 deficiency) diagnosed with intellectual disability. Most patients with BH_4 deficiency (95.7%) were concomitantly treated with

levodopa or carbidopa/levodopa. Detailed baseline characteristics and demographics (including gene mutation information) are provided in Table 31 and patient disposition is provided in Table 32.

TABLE 31: BASELINE CHARACTERISTICS AND DEMOGRAPHICS

	PAH-Deficient Patients (n = 296)	BH_4 Deficient Patients (n = 29)	Total (N = 325)
Age, Years			
Median	10.3	12.8	10.3
Q1, Q3	7.2, 15.0	6.6, 18.9	7.1, 15.5
Age Group (Years), n			
< 4	0	5	5
4 to < 8	100	7	107
8 to < 12	75	1	76
12 to < 18	70	8	78
18 to < 65	51	8	59
Gender, n (%)			
Male	150 (50.7)	16 (55.2)	166 (51.1)
PAH Gene Mutation Information, n (%)			
Patients with data	212 (71.6)	-	-
Patients with a classified mutation	210 (70.9)	-	-
Mutation-Type[a,b]		-	-
p.L48S	44 (14.9)	-	-
p.Y414C	38 (12.8)	-	-
p.R261Q	39 (13.2)	-	-
IVS10-11G>A	30 (10.1)	-	-
Genotypes[c]		-	-
p.R408W/p.Y414C	7 (2.4)	-	-
p.L48S/p.L48S	6 (2.0)	-	-
p.R261Q/p.R261Q	6 (2.0)	-	-
Participating Countries, n			
Austria	[redacted]	[redacted]	[redacted]
France	[redacted]	[redacted]	[redacted]
Germany	[redacted]	[redacted]	[redacted]
Italy	[redacted]	[redacted]	[redacted]
The Netherlands	[redacted]	[redacted]	[redacted]
Slovakia	[redacted]	[redacted]	[redacted]
Spain	[redacted]	[redacted]	[redacted]

BH_4 = tetrahydrobiopterin; PAH = phenylalanine hydroxylase; Q = quartile.
Note: Demographics shown are for patients included in this interim analysis.
[a] Mutations on either allele 1 or 2.
[b] In ≥ 10% patients.
[c] In ≥ 2% patients; includes allele1-allele 2 or allele 2-allele 1 combinations.

TABLE 32: PATIENT DISPOSITION

	PAH-Deficient Patients	BH_4-Deficient Patients
Enrolled	329	
Withdrew[a]	4	
Included in third interim analysis[b]	325	
	296	29

BH_4 = tetrahydrobiopterin; PAH = phenylalanine hydroxylase.
[a] Patients remained in database but did not meet study eligibility criteria.
[b] Patients enrolled from December 2009 and November 2012 with baseline data.

Safety

Sixty-one patients of the total population reported AEs or SAEs, with 18.6% (n = 55) and 20.7% (n = 6) reporting them in the PAH and BH_4 deficient populations, respectively. The most frequent AE was headache (2.7% and 3.4% in PAH and BH_4 deficiency, respectively) followed by abdominal pain and cough (both at 1.4% in patients with PAH deficiency only). Most AEs were mild or moderate in intensity in patients with either PAH and BH_4 deficiency. In patients with PAH deficiency, AEs per patient-year were 226.3 and 199.5 in year 1 and year 2, respectively, while AEs per patient-year were 22.5 in year 1 for patients with BH_4 deficiency. Seven SAEs occurred in the 61 total patients, of which one was classified as severe (headache leading to hospitalization; possibly related to SAP treatment). Notable AEs occurring in BH_4-deficient HPA on concomitant carbidopa/levodopa treatment were chorea and tic in one patient and hypertonia in another patient. No deaths had occurred at the time of the interim analysis. Detailed harms data are provided in Table 33.

TABLE 33: HARMS

	PAH-Deficient Patients (n = 296)	BH_4-Deficient Patients (n = 29)
Any AE, n (%)	**55 (18.6)**	**6 (20.7)**
Specific AEs		
Headache	8 (2.7)	1 (3.4)
Abdominal pain	4 (1.4)	-
Cough	4 (1.4)	-
Acute tonsillitis	3 (1.0)	-
Nasopharyngitis	3 (1.0)	-
Tonsillitis	3 (1.0)	-
Decreased weight	2 (0.7)	-
Gastroenteritis	2 (0.7)	-
Overweight	2 (0.7)	-
Pharyngitis	2 (0.7)	-
Rhinorrhea	2 (0.7)	-
Acne	1 (0.3)	1 (3.4)
Vomiting	1 (0.3)	1 (3.4)
SAEs	7	
Deaths	0	

AE = adverse event; BH_4 = tetrahydrobiopterin; PAH = phenylalanine hydroxylase.

Clinical Effectiveness

In patients with PAH deficiency (n = 245), the median SAP dose was 12.7 mg/kg/day (quartile [Q] 1, Q3 of 10.0 and 18.9 mg/kg/day), while in patients with the BH_4 deficiency (n = 25), the median dose was 5.0 mg/kg/day (Q1, Q3 of 3.0 and 7.5 mg/kg/day). In pregnant patients, one patient received 3 mg/kg/day, another received 10 mg/kg/day, another patient originally receiving 8 mg/kg/day reduced their dose to 4 mg/kg/day, and the final pregnant patient received SAP doses between 9 and 17 mg/kg/day.

Blood Phe Levels

The median (Q1, Q3) concentration of blood Phe in PAH-deficient patients was 414 (289, 561) µmol/L, 349 (258, 503) µmol/L, and 340 (248, 486) µmol/L at baseline (n = 215), six months (n = 133), and 12 months (n = 121), respectively. The median (Q1, Q3) concentration of blood Phe in BH_4-deficient patients was 91 (67, 313) µmol/L, 103 (81, 254) µmol/L, and 89 (76, 117) µmol/L at baseline (n = 20), six months (n = 11), and 12 months (n = 6), respectively. Detailed results are provided for PAH-deficient patients only in Table 34; no detailed results were provided for BH_4-deficient patients.

Dietary Phe and Natural Protein Intake

Results were available for patients who had completed their dietary records. In patients with PAH deficiency, the median (Q1, Q3) dietary Phe intake appeared be higher at 12 months compared with baseline for all age groups but varied greatly over time; from 718 (385, 1,514) mg/day, 1,525 (750, 2,298) mg/day, and 1,205 (600, 2,549) mg/day at baseline (n = 136), six months (n = 79), and 12 months (n = 70), respectively. Natural protein intake appeared to follow a similar pattern to that for dietary Phe intake, as the median intake was higher when compared with the baseline values. Detailed results stratified by age group are provided in Table 34.

TABLE 34: BLOOD PHE LEVELS, DIETARY PHE INTAKE, AND NATURAL PROTEIN INTAKE IN PAH-DEFICIENT PATIENTS

	Age Group			
	4 to < 8 Years	8 to < 12 Years	12 to < 18 Years	18 to < 65 Years
Blood Phe Levels (µmol/L), Median (Q1, Q3)				
Baseline[a] [n]	371 (244, 494) [69]	367 (256, 508) [61]	414 (329, 555) [51]	609 (450, 869) [34]
Month 6 [n]	274 (229, 363) [45]	357 (288, 474) [41]	400 (309, 665) [31]	442 (327, 575) [16]
Month 12 [n]	321 (236, 376) [46]	340 (254, 486) [35]	446 (247, 564) [28]	482 (270, 678) [12]
Dietary Phe Intake (mg/day),[b] Median (Q1, Q3)				
Baseline[a] [n]	585 (370, 1,137) [42]	583 (300, 1,100) [39]	1,051 (600, 1,.880) [35]	1,057 (555, 1,763) [20]
Month 6 [n]	1,135 (600, 1,750) [25]	900 (600, 2,157) [24]	1,800 (1,407, 2,300) [15]	2,225 (1,860, 3487) [15]
Month 12 [n]	800 (444, 2,010) [31]	950 (600, 2,000) [18]	2,312 (1,315, 2,644) [16]	2,600 (2,400, 3,448) [5]
Natural protein Intake (g/day),[a] Median (Q1, Q3)				
Baseline[a] [n]	17.8 (13.0, 27.5) [8]	16.2 (9.0, 22.0) [10]	24.5 (17.0, 48.0) [15]	22.2 (15.8, 38.3) [10]

	Age Group			
	4 to < 8 Years	8 to < 12 Years	12 to < 18 Years	18 to < 65 Years
Month 6 [n]	36.0 (10.0, 47.0) [14]	30.0 (23.0, 46.0) [15]	38.0 (31.0, 47.0) [15]	44.8 (43.0, 50.5) [8]
Month 12 [n]	45.0 (26.0, 53.0) [19]	40.0 (14.0, 58.0) [9]	54.3 (50.0, 70.0) [11]	72.0 (62.0, 77.5) [4]

PAH = phenylalanine hydroxylase; Phe = phenylalanine; Q = quartile.
[a] Prior to sapropterin treatment.
[b] Patients who had data available.

Other Secondary Outcomes

Growth measurement (z scores for height, weight, and body mass index [BMI]) medians were considered to be within the normal range for patients who were both PAH- and BH_4-deficient. Detailed results are provided in Table 35. Osteopenia and osteoporosis were detected in five and two patients with PAH deficiency, respectively, out of the 59 patients with available baseline bone density data. Of the 34 patients with available 12-month data who were PAH deficient, there were two and one cases of osteopenia and osteoporosis, respectively; these had not been recorded in these patients at baseline. Only one patient with BH_4 deficiency had baseline bone density data available and their data were normal; however, they did not have any follow-up 12-month data.

TABLE 35: OTHER SECONDARY OUTCOMES

	PAH-Deficient Patients		BH_4-Deficient Patients	
	Baseline	12 Months	Baseline	12 Months[a]
Growth Measurements (z Scores), Median (Q1, Q3)				
Height [n]	–0.1 (–0.8, 0.6) [244]	–0.2 (–0.9, 0.6) [115]	–0.1 (–1.4, 0.2) [23]	-
Weight [n]	0.3 (–0.2, 1.2) [118]	0.4 (–0.4, 1.2) [58]	–0.1 (–1.2, 1.0) [12]	-
BMI [n]	0.4 (–0.4, 1.2) [244]	0.4 (–0.3, 1.3) [115]	0.3 (–1.3, 1.2) [23]	-

BH_4 = tetrahydrobiopterin; BMI = body mass index; PAH = phenylalanine hydroxylase; Q = quartile.
[a] 12-month data not reported due to small number of patients (< 10).

Limitations

As with the PKUDOS registry, the main limitation of the KAMPER registry is that it is an observational drug registry with limited follow-up data (e.g., assessment of the blood), especially at the point of this interim analysis. There is also the potential for data to be incomplete, as well as not verified. In addition, due to the fact that KAMPER is a drug registry, it lacks a control group of patients who have no experience with SAP and thus any subsequent analyses of a statistical nature will need to be interpreted with caution. Another important consideration is that the information up to this point is limited as this study is ongoing and registration will occur up until 2019; therefore, these results can be considered preliminary in the scope of the entire registry.

Conclusions

It appears that SAP has an acceptable safety and tolerability profile, based on the KAMPER registry study as well. The most frequent AE was headache, which occurred in a very small percentage of patients. For the most part, there was a trend for blood Phe levels to be lower at 12 months when compared with

baseline in PAH- and BH_4-deficient patients. It also appears that the amount of dietary Phe and natural Phe intakes are higher at 12 months when compared with baseline in PAH-deficient patients. No statistical comparisons were conducted.

c) PKU-015 Trial[14]

The PKU-015 trial is an open-label, ongoing, multi-centre trial (United States and Canada) that aims to evaluate the effectiveness (i.e., neurocognitive function maintenance, growth parameters, and blood Phe levels) and safety of long-term SAP use in children diagnosed with PKU/HPA who initiated SAP treatment between zero and six years of age. This study will continue in order to obtain seven years of effectiveness and safety evidence in approximately 45 patients. At the time of screening, these children were required to have an official diagnosis of PKU/HPA along with two blood Phe concentrations of ≥ 360 µmol/L at least three days apart. In addition, patients were required to be compliant with local treatment standards prior to enrolling in the study and consent was required for adherence to the Phe-restricted diet and other study protocols. Primary exclusion criteria are children with a diagnosis of primary BH_4 deficiency, organ transplant, serious neuropsychiatric illness, concurrent disease that could potentially interfere with treatment, and hypersensitivity to SAP. The primary outcome of PKU-015 was to determine the long-term efficacy (preserving neurocognitive function) of SAP. Secondary outcomes included long-term safety, six-month efficacy of SAP in controlling blood Phe levels, and the effect of SAP on growth parameters.

Findings

This interim analysis includes clinical evidence in a subset of SAP-responsive patients (required to maintain blood Phe levels in the recommended rage of < 360 µmol/L) who, as of June 2012, had completed at least two years of SAP treatment. In addition, information was also obtained from three patients who terminated the study early prior to their two-year visit. Responsiveness to 20 mg/kg/day SAP was evaluated for the first four weeks of the study, with only SAP-responsive patients (i.e., defined as ≥ 30% average decrease from baseline in blood Phe levels) being allowed to continue. Within six weeks of confirmation of SAP-responsiveness, patients were administered a baseline neurocognitive assessment. Only patients who attained a ≥ 80 on either an infant development test or IQ were allowed to continue in the study. Dose reductions were permitted and occurred in five patients. Efficacy and safety evaluations occurred or will occur every six months for seven years. Detailed baseline characteristics of the included patients are provided in Table 36.

TABLE 36: BASELINE CHARACTERISTICS

	Age Group			
	< 1 Year n = 10	1 to 2 Years n = 19	3 to 4 Years n = 13	5 to 6 Years n = 13
Enrolment Age (Years)				
Mean (SD)	0.47 (0.27)	1.84 (0.66)	4.05 (0.62)	6.17 (0.63)
Gender, n (%)				
Male	4 (40.0)	9 (47.4)	4 (30.8)	3 (23.1)
Race, n (%)				
White	8 (80.0)	16 (84.2)	12 (92.3)	12 (92.3)
Asian	1 (10.0)	1 (5.3)	0	1 (7.7)
Other	1 (10.0)	2 (10.5)	1 (7.7)	0
Growth Measurements (z Score), Mean (SD)				
Head circumference	0.21 (0.60)	0.42 (1.18)	-	-

	Age Group			
	< 1 Year n = 10	1 to 2 Years n = 19	3 to 4 Years n = 13	5 to 6 Years n = 13
Height	0.35 (1.23)	0.38 (0.70)	0.78 (1.03)	0.41 (0.84)
Weight	0.41 (0.99)	0.25 (0.75)	0.58 (0.71)	0.64 (0.91)
Blood Phe Levels (µmol/L)				
Mean (SD)	284.4 (116.2)	337.9 (124.0)	352.8 (128.8)	335.7 (183.9)

Phe = phenylalanine; SD = standard deviation.

Ninety-five children enrolled in the study, of whom 71 were considered to be SAP responders. Sixty-five patients continued on to the first six-month efficacy and safety evaluation and 63 patients continued beyond this first six-month time point and were enrolled in the long-term neurocognitive evaluation. The interim analysis of the two-year outcome included data from 55 patients, of whom 48 were per-protocol SAP responders and seven were clinical SAP responders. In addition, safety and baseline data were included in the analysis for three patients who discontinued the study prior to the two-year follow-up. Details on patient responsiveness are provided in Table 37.

TABLE 37: SAPROPTERIN RESPONSIVENESS

	Age Group				
	< 1 Year	1 to 2 Years	3 to 4 Years	5 to 6 Years	Total
Responsiveness					
Tested for SAP response, N	-	-	-	-	95
SAP responders, n/N (%)	11/13 (85)	24/37 (65)	22/27 (81)	14/18 (78)	71 (75)
PP SAP responders,[a] n (%)	-	-	-	-	63 (89)
Clinical SAP responders,[b] n (%)	-	-	-	-	8 (11)

PP = per-protocol; SAP = sapropterin.
[a] ≥ 30% reduction in mean blood Phe levels.
[b] < 30% reduction in mean blood Phe levels; however, blood Phe levels maintained in advised range of 120 to 360 µmol/L, despite increased dietary Phe intake.

Efficacy Outcome Measures

The mean blood Phe concentrations in SAP-responsive children declined from baseline to week 4 in all children and then increased to levels still below baseline by the third month, with the exception of children aged three to four years in whom levels increased to baseline levels. By 24 months, 64% of the children had blood Phe levels of ≤ 240 µmol/L, while 84% of the children had blood Phe levels ≤ 360 µmol/L. In addition, from baseline to the 24-month time point, prescribed dietary Phe increased in all age groups (i.e., in all patients the prescribed dietary Phe increased from 30.3 to 37.1 mg/kg/day or from 377.5 to 667.0 mg/day. Details regarding blood Phe concentrations are provided in Table 38.

With regard to neurocognitive function and development, baseline testing indicated that Full Scale Intelligence Quotients (FSIQs) were not significantly different from the population norm of 100 and no significantly different changes were apparent at the 24-months follow-up. No significant changes from baseline to the two-year follow-up were apparent with the mean Bayley Scales of Infant and Toddler Development, Third Edition (Bayley-III) cognitive composite index; these scores were maintained within the 100 ± 15 normative range. In addition, no infant scored less than 85 on the Bayley-III cognitive composite index and were therefore not at additional risk during the study for developmental delay. Detailed FSIQ results are provided in Table 38.

Growth assessments (in terms of mean z scores for height, weight, and head circumference) demonstrated that all the children were slightly above the 50th percentile and that there were no statistically significant changes in values from baseline, indicating that these values were maintained. Details for the growth assessments are provided in Table 38.

TABLE 38: BLOOD PHE CONCENTRATIONS, NEUROCOGNITIVE FUNCTION, AND GROWTH ASSESSMENTS IN CHILDREN ON SAPROPTERIN TREATMENT

	Time Period					
	Baseline	1 Week	2 Months	6 Months	12 Months	24 Months
Blood Phe Concentrations, n/N (%)						
≤ 240 µmol/L	18/55 (33)	42/53 (79)	33/48 (69)	35/52 (67)	33/52 (63)	32/50 (64)
≤ 360 µmol/L	36/55 (65)	50/53 (94)	40/48 (83)	43/52 (83)	44/52 (85)	42/50 (84)
Development and Neurocognitive Function, Score (SD)						
Mean FSIQ[a] *P* value	103 (12)	-	-	-	-	104 (10) 0.50
Growth Assessments, Mean z Scores (SD)						
Height	0.4 (0.9)	-	-	-	-	Maintained[b]
Weight	0.4 (0.8)	-	-	-	-	Maintained[b]
Head circumference	0.3 (1.0)	-	-	-	-	Maintained[b]

FSIQ = Full Scale Intelligence Quotient; Phe = phenylalanine; SD = standard deviation; WISC-IV = Wechsler Intelligence Scale for Children, Fourth Addition; WPPSI-III = Wechsler Preschool and Primary Scale of Intelligence.
[a] Includes data for whom baseline and two-year WPPSI-III and WISC-IV scores were available, n = 25.
[b] Values were maintained throughout the 2 years of follow-up with no statistically significant changes from baseline (data not provided). These values indicate that all three growth parameters were slightly above the 50th percentile.

Safety

The most common AEs that occurred in ≥ 5% of patients included abdominal pain, diarrhea, headache, ear and upper respiratory tract infections, nasal congestion, and nasal congestion. In addition, these were either possibly or probably related to SAP treatment. While some SAEs were reported (9%), none were determined to be related to SAP treatment. Details regarding harms are provided in Table 39.

TABLE 39: HARMS

	Age Group			
	< 1 Year n = 10	1 to 2 Years n = 19	3 to 4 Years n = 13	5 to 6 Years n = 13
Drug-related[a] AEs occurring in ≥ 5% patients, n (%)				
Abdominal pain	1 (10.0)	2 (10.5)	1 (7.7)	1 (7.7)
Diarrhea	0	4 (21.1)	2 (15.4)	0
Ear infection	2 (20.0)	1 (5.3)	0	0
Headache	0	0	2 (15.4)	1 (7.7)
Nasal congestion	1 (10.0)	1 (5.3)	1 (7.7)	2 (15.4)
Upper abdominal pain	0	1 (5.3)	1 (7.7)	2 (15.4)
URTI	1 (10.0)	1 (5.3)	1 (7.7)	3 (23.1)
Vomiting	2 (20.0)	3 (15.8)	1 (7.7)	1 (7.7)
SAEs,[b] n (%)				
Airway complication of anesthesia[c]	0	1 (5.3)	0	0
Constipation	1 (10.0)	0	0	0
Croup (infectious)	0	0	1 (7.7)	0
Injury[c]	0	1 (5.3)	0	0
Pneumonia	1 (10.0)	0	0	0
Seizure	0	1 (5.3)	0	0

AE = adverse event; SAE = serious adverse event; URTI = upper respiratory tract infection.
[a] Possibly or probably related to sapropterin treatment.
[b] All SAEs seemed unrelated to sapropterin treatment.
[c] Occurred in same patient.

Limitations

The main limitation of the PKU-015 study was that the performed interim analysis was conducted using a subset of patients with available data at the two-year follow-up. In addition, this is an open-label, single-arm study with no official comparator; therefore, this limits the interpretability of the results of this analysis. Specific to this study, there are additional challenges with the analysis associated with neurocognitive assessments, as different scales were used as the child increased in age.

Conclusions

SAP appears to have an acceptable safety and tolerability profile in young children and infants with PKU. In addition, it appears to reduce blood Phe levels and allows for dietary increases in Phe intake, although the clinical significance of the magnitude of the increase in dietary Phe is unknown. Neurocognitive function and growth assessment mean z scores were maintained over the two-year follow-up and none of the infants experienced developmental delays.

APPENDIX 7: DIETARY TREATMENT FOR PHENYLKETONURIA

This section was prepared by a metabolic dietician specializing in phenylketonuria (PKU) and updated by a CADTH Common Drug Review (CDR) reviewer. It has not been systematically reviewed.

1. Current Dietary Treatment

Dietary restriction of phenylalanine (Phe) has been the mainstay of treatment for PKU in Canada for more than 40 years, since the introduction of neonatal screening for the disease. Two recent new treatment strategies are now in limited use;[31] large neutral amino acids (LNAA) and tetrahydrobiopterin (BH_4). However, they are not targeted to the whole patient population, and are often used as an adjunct to dietary treatment.

2. Goals of Dietary Treatment

a) To maintain plasma Phe concentrations within the recommended treatment ranges[27] by:

- Controlled restriction of natural protein intake:
  - Exclusion of high-protein foods (e.g., meat, fish, poultry, eggs, milk, most dairy products, nuts, soy, and legumes) from the diet.
  - Daily allowance of dietary Phe from measured quantities of lower protein foods (e.g., cereals, vegetables, and fruit), allowed within Phe tolerance.
    - Phe tolerance varies widely from individual to individual, depending on the variant of PKU,[32] and in the same individual, depending on age,[33] adequacy of energy, protein intakes, and state of health.
- Supplementation of the diet with a medical food (i.e., synthetic protein consisting of Phe-free L-amino acids, and additional nutrients).
  - The optimal dose of medical food has not been determined, but it usually provides 52% to 80% of total dietary protein.[34]
  - Consumption of protein substitute should be evenly spread throughout the day to optimize protein utilization for anabolism, and to control plasma Phe concentrations.[35]
  - The distribution and dosage of the protein supplement are probably the most important determinants of plasma Phe concentrations in individuals with PKU on diet.[36,37]
  - Historically, patient compliance with protein substitutes is poor,[38] but palatability and presentation have improved in recent years. Many new novel protein substitutes have been introduced and short-term studies have demonstrated improved compliance.[39,40]
  - Glycomacropeptide (GMP) can also be used as an alternative to synthetic amino acids as it is dietary protein that contains no Phe. It occurs naturally in the whey fraction of bovine milk; does not contain aromatic amino acids, Phe, tryptophan, and tyrosine; and contains concentrations that are higher in isoleucine and threonine. In general, GMP medical foods are acceptable as an alternative to the synthetic amino acid diets, and increase variety, satiety, and palatability.[41]
  - There is the potential for LNAA to help lower blood and brain Phe levels as they competitively inhibit (using the four L-system transporters) the transportation of Phe across the blood–brain barrier and intestinal mucosa. LNAA has been prescribed (either alone or in combination with the low-Phe diet) for adults who either cannot adhere to or have a hard time adhering to the Phe-restricted diets. However, evidence of its efficacy is lacking; therefore, additional research is required before its implementation in everyday practice.[41]

b) To ensure that the diet meets requirements for Phe (an essential amino acid), total protein (natural plus synthetic), energy, fat, carbohydrate, and micronutrients (vitamins, minerals, trace elements, essential fatty acids):

- Provision of special low-protein foods (e.g., breads, pastas, baked goods, milk substitutes, cheese, vegetable burgers, scrambled egg mix, mashed potato mix), in order to meet energy requirement and/or relieve the monotony of the diet.
- Supplementation of the diet with micronutrients (e.g., vitamins, minerals, and trace elements) if the prescribed intake of medical food is not anticipated to meet estimated requirements.

c) To support normal growth and development in children and optimal health (physical and mental) in adults

d) To achieve micronutrient biochemical status that is ideal

e) To ensure that the diet is palatable, flexible, and compatible with a modern-day lifestyle:
Despite continued improvements, the diet remains highly restrictive, socially and financially burdensome, and time-consuming to manage;[42] hence, leading to poor adherence.

3. Dietary Adherence

Dietary adherence is determined by plasma Phe concentrations: patients are expected to provide routine blood samples for this purpose. However, in children older than 10 years, plasma Phe concentrations above the age-related treatment goals, and frequency of blood testing below recommendations, are the norm.[19,21]

Patients often have to wait days to weeks to obtain the results of a blood Phe test. However, a portable Phe monitoring device (currently in development) has the potential to improve dietary adherence, by allowing patients to monitor blood Phe levels in real time.[31]

4. Harms of the Phe-Restricted Diet

a) Deficiency of micronutrients (vitamins, minerals, trace elements):

- Vitamin B12 deficiency has been reported on a number of occasions and is most probably secondary to noncompliance with the medical food,[43] the only source of vitamin B12 in the diet (apart from a multivitamin and mineral supplement), if intake of high-protein foods of animal origin is avoided.
- Individuals with PKU are at risk for suboptimal plasma levels of iron, zinc, and vitamin A, despite adequate intake.[44]

b) Deficiency of essential fatty acids:
The diet, often low in fat and essential fatty acids (especially alpha-linolenic acid), puts children at risk of deficiencies in long-chain polyunsaturated fatty acids (LCPUFA),[45] especially in docosahexaenoic acid (DHA).[46]

c) Impaired growth in children:
There have been a number of reports of impaired growth, but this has not been a universal observation.[47]

d) Reduced bone mass:
Reduction in peak bone mass has been described in adults with PKU[48] but the etiology (diet and/or disease) has yet to be elucidated.

e) Obesity:

- Children with PKU have been reported to be at risk of obesity, despite energy intakes that appeared to be less than recommended.[49]

- Investigators have suggested that PKU patients with poor metabolic control may be subject to dysregulation of their neuroendocrine system, secondary to high plasma Phe levels.[50]

5. Current Practices in Managing Phenylketonuria in Canada

There are approximately 21 treatment centres in Canada: three devoted to adults and the remainder to individuals of all ages. A survey of metabolic dietitians across Canada was conducted in 2009. The response rate of the survey was approximately 50%. Of the 10 centres that responded, two followed fewer than 20 patients, five between 20 and 50 patients, two between 51 and 100 patients, and one between 101 and 150 patients. Nine of the 10 centres reported that they each had about 20 patients who were inactive, and the remaining centre between 51 and 100 patients. Five of the nine centres reported that they attempt to maintain contact with inactive patients.

There is currently an initiative underway by a physician to try to achieve more comprehensive information than that which was collected in the above-mentioned 2009 survey. There was general agreement among the respondents of the survey in terms of the frequency with which patients are expected to attend clinic visits and to provide blood samples (to monitor Phe concentrations). The majority of the respondents (six out of nine) indicated that their patients had to visit the hospital to provide a blood draw: patients in the remaining centres (three out of nine responses) were able to send in blood spots. The majority of respondents (seven out of nine) indicated that their treatment centre does a BH_4 loading test, before initiation of diet therapy in an infant with a positive newborn screen for PKU.

On the contentious issue of whether diet therapy is indicated for mild hyperphenylalaninemia,[51] six respondents (out of 10) indicated that their treatment centres initiate the Phe-restricted diet, if blood Phe concentration is more than 360 µmol/L, and another respondent if it is more than 240 µmol/L. The remaining three respondents (out of 10) indicated that their institutions do not initiate the Phe-restricted diet unless blood Phe concentration is more than 600 µmol/L; i.e., unless the newborn has PKU. In general, there was consensus among the 10 respondents regarding target blood Phe concentrations,[27] and they all indicated that their treatment centres recommend “diet for life.”

APPENDIX 8: PROPOSED REIMBURSEMENT CRITERIA FROM ONTARIO SPECIALISTS

A) Ontario Drug Program Exceptional Access Program Reimbursement Criteria

The current Ontario Drug Program Exceptional Access Program (EAP) reimbursement criteria for Kuvan are listed below. The standard approval duration is one year for ongoing funding. For initial funding in non-pregnant patients and patients planning a pregnancy, the approval duration is six months and is funded by the manufacturer. The approval duration in pregnant patients is six months or to the end of pregnancy and is also funded by the manufacturer.

<u>Ongoing funding</u> of sapropterin (Kuvan) will be considered through the EAP for non-pregnant patients and patients actively planning pregnancy who have a diagnosis of Phenylketonuria (PKU) and who have demonstrated a response to the initial six-month trial of sapropterin (reimbursed through the manufacturer [see details below]) and who meet ALL of the following criteria:

1. Compliance with low-protein diet, formulas, and treatment with sapropterin;

AND

2. Has achieved:
   a) normal sustained blood phenylalanine (Phe) levels [greater than 120 µmol/L and less than 360 µmol/L] (at least two levels measured at least one month apart); OR
   b) sustained blood Phe reduction of at least 30% (at least two levels measured at least one month apart) compared with baseline if the Phe baseline level is less than 1,200 µmol/L*;* OR
   c) sustained blood Phe reduction of at least 50% (at least two levels measured at least one month apart) compared with baseline if the Phe baseline level is greater than 1,200 µmol/L;

AND

3. Demonstrated increase of dietary protein tolerance based on targets set between the clinician and patient;

AND

4. Clinically meaningful age-appropriate improvement in:
   a) neurobehavioural or neurocognitive function or impairment for patients with such impairments as determined by peer-demonstrated improvement in Quality of Life using peer-reviewed validated scales; AND
5. Managed by a physician specialized in metabolic/biochemical diseases.

Please note that sapropterin is considered only through the EAP for responders to an initial six-month trial period funded through the manufacturer (described in the next two pages). The exclusion criteria for initial funding of sapropterin also apply to funding through the EAP. (See next page.)

THE INITIAL SIX MONTHS OF FUNDING OF SAPROPTERIN (KUVAN) IS PROVIDED BY THE MANUFACTURER, BIOMARIN PHARMACEUTICAL CANADA INC. THE CRITERIA FOR INITIAL COVERAGE ARE DESCRIBED BELOW.

<u>Initial funding</u> of sapropterin will be considered by the manufacturer for the management of non-pregnant patients and patients actively planning pregnancy who have a diagnosis of PKU and who meet ALL of the following criteria:

1. Compliance with a low-protein diet and formulas
2. Baseline blood Phe levels are greater than 360 µmol/L despite compliance with low-protein diet (require at least two levels during three- to six-month time frame)
3. Baseline protein intake assessment by a dietitian
4. Ability to comply with medication regimen
5. Managed by a physician specialized in metabolic/ biochemical diseases.

Initial funding of sapropterin will also be considered by the manufacturer for the management of pregnant patients who have a diagnosis of PKU and who meet ALL of the following criteria:

- Managed by a physician specialized in metabolic/biochemical diseases
- Baseline blood Phe levels > 360 µmol/L despite compliance with all recommendations for dietary intervention and monitoring.

Funding will not be considered for patients meeting any of the following exclusion criteria:

- Known hypersensitivity to sapropterin or its excipients
- Any other contraindications
- Baseline Phe Levels less than 360 µmol/L in a non-pregnant patient
- Baseline Phe Levels less than 360 µmol/L in a pregnant patient
- Women who are nursing/breastfeeding
- Patients who are not on the special diet or who are not compliant with their special diet.

Note that sapropterin should be used with caution when the patient is taking medication known to inhibit folate synthesis (e.g., methotrexate) and/or has any condition that requires treatment with levodopa or any phosphodiesterase type 5 (PDE5) inhibitor.

Additionally, consideration for initial funding of sapropterin requires that the patient completes an eligibility test called the "72-Hour Kuvan Challenge," described below.

Test for Eligibility: "72-Hour Kuvan Challenge"

- 72-hour challenge with sapropterin at 20 mg/kg/day
- Blood Phe concentrations are measured at 48 hours, 24 hours, and time "0" PRIOR TO the sapropterin dose and THEREAFTER at 4, 12, 24, 48, and 72 hours following the dose; OR as per clinic's protocol.

Note that the recommended dose of sapropterin to establish clinical benefit is 20 mg/kg/day.

Responders to the "72-Hour Kuvan Challenge"

For non-pregnant patients and patients actively planning pregnancy, responders to the Kuvan challenge are those who meet the following criteria:

- Reduction in Phe blood level of at least 30% compared with baseline; AND
- Patient must have a baseline assessment of neurobehavioural or neurocognitive impairment[1] and quality of life assessment due to PKU after the 72-hour Kuvan challenge but before start of Kuvan therapy (this assessment does not apply to pregnant women).

[1] For children younger than four years of age, clinically validated age-appropriate neurobehavioural, neurocognitive, or developmental tests may be selected at the clinician's discretion rather than PKU-specific tests.

Note: A baseline Phe tolerance level must be documented and Phe tolerance levels must be documented at months 1 to 2 and 4 to 6 during the initial six months of therapy.

For pregnant patients and patients actively planning pregnancy, responders to the Kuvan challenge are those who meet the following criteria:

- Reduction in Phe blood level of at least 30% compared with baseline after 72 hours.

Pregnant patients who meet the "responder" definition to the 72-hour Kuvan Challenge may be eligible for Kuvan funding if the following criteria are met:

- A decrease in Phe concentration to less than 360 µmol/L is to be maintained for the duration of pregnancy to be eligible for continued funding.

Renewals for sapropterin in pregnant patients will not be considered.

B) Alternate Exceptional Access Program Reimbursement Criteria Proposed by Ontario Specialist Physicians

A separate submission was made by Ontario PKU experts in March 2014 suggesting alternate reimbursement criteria for Kuvan to the Ontario Drug Program, which are as follows:

Re: Exceptional Access Program Reimbursement Criteria for Kuvan (Sapropterin) for the Treatment of Phenylketonuria (PKU)

Phenylketonuria experts representing all of the PKU clinics in Ontario would like to propose alternative EAP reimbursement criteria for Kuvan for the treatment of PKU.

The current EAP in Ontario were developed with little input from the physicians and other health care providers in Ontario who actually treat these individuals. Unfortunately, these criteria are difficult, if not impossible, to meet because of contradictory requirements and because of lack of resources in most centres to perform serial neurobehavioural/neurocognitive testing in these individuals as required by the criteria.

The proposed new guidelines (Appendix 8.1) were developed by consensus of all of the undersigned physicians. All of these doctors are experts in phenylketonuria and represent the PKU clinics in Toronto, Ottawa, London, Hamilton, and Kingston that follow the vast majority of individuals with PKU who are being actively treated in this province. An external expert, Dr. John Mitchell of Montreal, was also consulted on the proposed guidelines. A more detailed explanation that includes the rationale and evidence for the proposed guidelines and which highlights the differences from the current criteria is included as Appendix 8.2.

We look forward to your response to this proposal.

Sincerely,
Dr. Andreas Schulze Hospital for Sick Children, Toronto
Dr. Komudi Siriwardena Hospital for Sick Children, Toronto
Dr. Pranesh Chakraborty Children's Hospital of Eastern Ontario, Ottawa
Dr. Michael Geraghty Children's Hospital of Eastern Ontario, Ottawa
Dr. Chitra Prasad London Health Sciences Centre, London
Dr. Murray Potter McMaster Children's Hospital, Hamilton
Dr. Jennifer MacKenzie Kingston General Hospital, Kingston
Dr. John Mitchell Montreal Children's Hospital, Montreal

(Signature pages and disclosures of any conflicts of interest by the physicians were attached to the proposal)

Appendix 8.1: Suggested Guidelines for Kuvan Reimbursement

Individuals will be considered eligible for funding for Kuvan if they:

1. Have hyperphenylalaninemia requiring treatment, defined as having at least one documented plasma phenylalanine level of 600 µmol/l or greater, or greater than 360 µmol/l if pregnant or planning pregnancy.

It is recognized that non-pregnant individuals with peak phenylalanine levels between 360 and 600 µmol/l may also benefit from Kuvan, but the largest benefit of Kuvan will likely be seen in individuals who meet the above definition. At this time, individuals not meeting the above definition are not eligible for Kuvan reimbursement through the Exceptional Access Program.

AND

2. Respond to Kuvan, defined as a decrease of blood phenylalanine levels of at least 30% from baseline during a 4 week Kuvan challenge or during a neonatal Kuvan challenge.

Both plasma and whole-blood dried blood spot samples (DBS) are acceptable, but the baseline and challenge sample type (plasma versus DBS) and collection conditions (overnight fast, 4 hour preprandial, etc.) should be consistent for the entire determination of Kuvan responsiveness.

For the 4 week challenge, baseline levels must be determined by at least 2 plasma phenylalanine levels within a 2 week period before starting Kuvan. Baseline measurements must be stable and average >360 µmol/L for the study to be considered valid, with no measurement more than 20% above or below the last level taken. Individuals who have previously taken Kuvan should have a wash-out period of at least 4 weeks prior to determining Kuvan responsiveness. Individuals with baseline levels <=360 µmol/L should have their dietary phenylalanine adjusted to raise the plasma phenylalanine >360 µmol/L before beginning the Kuvan challenge. For the neonatal challenge, at least one pre-Kuvan baseline sample must be collected within 4 hours before the Kuvan loading dose.

Kuvan should be administered at a dose of 20 mg/kg/day, taken once daily for the duration of the challenge, during which time the dietary management is kept stable and similar to the baseline period.

For the 4 week challenge, blood phenylalanine levels should ideally be determined according to the following schedule (± 1 day for any time point): T-14d, T-7d, T0 (just prior to starting Kuvan), T+2d, T+7d, T+14d, T+21d, T+28d, with a minimum of 3 levels obtained after starting Kuvan. At least 3 post-Kuvan levels should be averaged and used to determine the response to Kuvan. If an intercurrent illness prevents collection of at least 3 suitable samples then the challenge may be extended one additional week. For a single dose neonatal challenge, at least one post-Kuvan load sample must be collected within 28 hours of the Kuvan load. The lowest phenylalanine level post-load is used to determine responsiveness. For a multi-day neonatal challenge, the lowest phenylalanine level from each day should be averaged to determine responsiveness.

AND

3. Demonstrate therapeutic blood phenylalanine levels while on Kuvan, defined as blood phenylalanine levels consistently in the control range.

For individuals < 16 years of age this is defined as <= 360 µmol/L in more than 50% of the monitoring samples and <600 µmol/L in more than 80% of the monitoring samples.

For individuals >=16 years of age this is defined as <600 µmol/L in more than 80% of the monitoring samples.

Phenylalanine control will be audited at 6 months after completing the Kuvan responsiveness trial and thereafter every 12 months. Monitoring samples should be collected at least monthly while on Kuvan. An individual will be considered to have failed to demonstrate phenylalanine control if there are no monitoring samples for any period exceeding 3 months duration.

Appendix 8.2: Background and Rationale

PKU is an inherited genetic disorder leading to an accumulation of phenylalanine in the blood and body tissues.[52] The degree to which phenylalanine increases in any one individual on a normal diet is determined by the specific genetic changes in the gene that causes PKU. There is a high degree of consensus that sustained levels of blood phenylalanine less than 360 µmol/L do not cause damage to the brain but that levels 600 µmol/L and higher can cause significant damage.[1,26,27] Untreated, this damage starts soon after birth and can progress to a severe, permanent disability. There is also a high degree of consensus that consistently maintaining blood phenylalanine levels in the range of 120 – 360 µmol/L can minimize or eliminate this brain damage.[1,27] For most people with PKU, a highly specialized diet is required to achieve this goal. This dietary therapy has been the standard of care for approximately 50 years, and, as you are aware, is funded in Ontario by the Inherited Metabolic Diseases Program. Kuvan is an alternative or adjunctive therapy to help reach this goal by lowering the phenylalanine level in some "Kuvan responsive" individuals.[53]

"Maternal PKU" is a distinct entity where a woman with PKU has high phenylalanine levels while pregnant. The developing fetus (which does not usually have PKU itself) is subjected to high phenylalanine from the mother's circulation, causing birth defects. The risk of these defects is reduced dramatically by the woman maintaining her blood phenylalanine levels between 120-360 while pregnant.[54,55]

With this brief background, we would like to explain the rationale and evidence for the proposed criteria as well as highlight the main differences between the proposal and the current criteria:

1. The new proposal has a clear definition of what type of hyperphenylalaninemia is eligible for reimbursement (**criterion #1**). As outlined in the background material above, there is consensus that individuals with hyperphenylalaninemia >=600 µmol/L while on an unrestricted diet require treatment, while those with phenylalanine < 360 µmol/L do not. There is not strong consensus on what treatment, if any, is needed for those with levels between 360-600 µmol/L on an unrestricted diet. There is a growing body of evidence that individuals with levels in this intermediate range may suffer some harm if they are not treated, but it is not yet clear.[51,56] The apparent paradox is that individuals with PKU who are on treatment to lower their phenylalanine levels to a "safe" level (<360 µmol/L) but who do not achieve this goal and instead have levels (while treated) in the 360-600 µmol/L range do suffer harm. The recent ACMG Practice Guidelines[1] recommends treating these individuals with hyperphenylalaninemia where the phenylalanine levels are between 360-600 µmol/L and that "any combination of therapies", including Kuvan, should be used to achieve this goal. We have chosen a more conservative approach that limits reimbursement for non-pregnant individuals to those with blood phenylalanine levels on an unrestricted diet of 600 µmol/L or greater. This, by definition, excludes non-pregnant individuals with milder forms of (non-PKU)

hyperphenylalaninemia. There is ample evidence that these excluded individuals actually have a very good chance of being responsive to Kuvan, but our expert consensus opinion is that the evidence at this time does not support that the clinical benefit of treating these individuals with Kuvan outweighs the cost of the therapy.

In contrast, our expert consensus is that the potential risk to a developing fetus in women with hyperphenylalaninemia in the intermediate range (360-600 µmol/L on an unrestricted diet) is significant enough that any and all measures, including Kuvan, should be available to control the phenylalanine levels. Therefore, we recommend a lower threshold of >360 µmol/L for individuals who are pregnant / planning pregnancy.

The current reimbursement criteria do not address the issue of the different types of hyperphenylalaninemia. There is an exclusion for patients with baseline Phe levels less than 360 µmol/L in both pregnant and non-pregnant patients, but this exclusion is speaking to the blood Phe attained on therapy, not the type of hyperphenylalaninemia (which, when defined by blood Phe levels as outline above, is defined on an unrestricted diet; i.e., off therapy).

2. The definition of a response to Kuvan in the proposal is outlined in **criterion #2**. The current reimbursement criteria uses a 72 hour Kuvan Challenge as a test for eligibility. While this has the advantage over other protocols in that it can be completed in 6 days (including the baseline measurements), it may not detect all patients who are responsive to Kuvan. We propose the widely accepted 4-week challenge be used to determine responsiveness, as this challenge is frequently used in the definition of Kuvan responsiveness in clinical trials, is very familiar to the clinics in Ontario, and has been shown to detect more Kuvan responders. We also propose that the single dose neonatal challenge, which has been used for decades as a test in the diagnostic evaluation of infants with positive newborn screens for PKU, is an acceptable and economical alternative demonstration of Kuvan responsiveness, recognizing that those who are nonresponsive to a neonatal challenge should also be offered the 4 week challenge later in life to more definitively determine responsiveness.

3. The current inclusion criteria have conflicting requirements around “compliance” and blood Phe levels. Inclusion criteria are variously stated as having “compliance with a low protein diet and formula” and “baseline blood phenylalanine (Phe) levels are greater than 360 µmol/L despite compliance with low protein diet” for non-pregnant patients and “baseline blood phenylalanine (Phe) levels are >360 µmol/L despite compliance with all recommendations for dietary intervention and monitoring” for pregnant patients. The target of diet therapy is to reduce the blood Phe to 120-360 µmol/L, so compliance with diet is generally defined as maintaining Phe levels in this target range. Therefore it is impossible to meet both components of this requirement (being compliant *and* having Phe >360) in the current reimbursement criteria. The proposed criteria do not define baseline blood Phe requirements for funding beyond the definition of type of hyperphenylalaninemia eligible for funding as outlined in criterion #1, and do not require “compliance” with therapy as a prerequisite to starting Kuvan.

4. We feel compliance is not a reasonable inclusion criterion, but agree that o*ngoing* funding of Kuvan requires demonstration of compliance, as outlined in **criterion #3**. The targets and definition of compliance in the proposed criteria differ from the existing criteria for ongoing funding. The new criteria require meeting absolute age-based targets for blood Phe while the current criteria allow targets based on baseline Phe. Allowing different targets based on baseline Phe puts undue emphasis on the meaning of the baseline Phe (which, as defined in the existing criteria may represent only a very small sampling of the actual pre-Kuvan blood Phe levels in that patient) and could allow Phe decreases that are mathematically significant but not clinically significant to be deemed acceptable.

5. Increased dietary protein or Phe tolerance has been completely removed from the proposed criteria. In a Kuvan responsive patient, giving Kuvan can decrease the blood Phe when the dietary protein or Phe is kept constant. This decrease in blood Phe is generally regarded to be caused by the Kuvan increasing metabolism of the Phe that is consumed in the diet. The increased metabolism of Phe (and subsequent decrease in blood Phe) can be counteracted by increasing the dietary Phe up to a point where the blood Phe increases back up to the baseline level; i.e., the increased metabolism is now balanced by increased dietary Phe, or increased Phe "tolerance". The decrease in blood Phe and the increase in Phe tolerance caused by Kuvan are fundamentally linked and opposing processes. Kuvan causes increased Phe metabolism, which can be used to decrease the blood Phe (hold the dietary Phe the same) or increase the Phe tolerance (hold the blood Phe the same). How the increased Phe metabolism is used in an individual patient depends on a number of factors and clinical goals. The current criteria for ongoing funding that require demonstrating of both a maximal decrease in blood Phe AND an increase in protein tolerance are illogical.

6. Neurobehavioural / neurocognitive measurement or quality of life improvements are not included in the proposed criteria. The concept of quality of life improvements with Kuvan is mainly based on the supposition that increased dietary tolerance for protein will lead to measureable improvements in quality of life scales. Removal of dietary tolerance from the criteria makes quality of life measurements less applicable. Serial neurobehavioural /neurocognitive testing is not available to most metabolic clinics, which creates an equitable access issue across the province. As well, evidence for neurobehavioural / neurocognitive impairments in PKU is linked to increased Phe levels. Poor compliance (high Phe levels) is not an inclusion criterion in the proposed criteria, which makes testing for these impairments less relevant.

REFERENCES

1. Vockley J, Andersson HC, Antshel KM, Braverman NE, Burton BK, Frazier DM, et al. Phenylalanine hydroxylase deficiency: diagnosis and management guideline. Genet Med. 2014 Feb;16(2):188-200.
2. Leuders S, Wolfgart E, Ott T, du Moulin M, van Teeffelen-Heithoff A, Vogelpohl L, et al. Influence of PAH genotype on sapropterin response in PKU: Results of a single-center cohort study. JIMD Rep [Internet]. 2014 [cited 2016 Mar 7];13:101-9. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4110336/pdf/978-3-642-54149-0_Chapter_263.pdf
3. About PKU. 2015 [cited 2016 Mar 27]. In: Canadian PKU and Allied Disorders [Internet]. Toronto: Canadian PKU and Allied Disorders Inc.; 2015 - . Available from: http://www.canpku.org/about-pku.
4. Aldamiz-Echevarria L, Bueno MA, Couce ML, Lage S, Dalmau J, Vitoria I, et al. Tetrahydrobiopterin therapy vs phenylalanine-restricted diet: impact on growth in PKU. Mol Genet Metab. 2013 Aug;109(4):331-8.
5. Kuvan (sapropterin dihydrochloride): 100 mg tablets [product monograph]. Toronto (ON): BioMarin Pharmaceutical (Canada) Inc.; 2014 Dec 4.
6. CADTH Canadian Expert Drug Advisory Committee (CEDAC) final recommendation: Kuvan (Biomarin Pharmaceutical (Canada) Inc) [Internet]. Ottawa: CADTH; 2011 Jan 26. [cited 2016 Mar 21]. Available from: https://www.cadth.ca/sites/default/files/cdr/complete/cdr_complete_Kuvan_Jan-28-2011.pdf
7. CDR submission: Kuvan resubmission (Request for advice), 100 mg tablets. Company: BioMarin Pharmaceutical (Canada) Inc. [**CONFIDENTIAL** manufacturer's submission]. Toronto (ON): BioMarin Pharmaceutical (Canada) Inc.; 2011 Oct 26.
8. European Society for Phenylketonuria and Allied Disorders Treated as Phenylketonuria [Internet]. Kirchberg am Wagram [Austria]: E.S. PKU. European phenylketonuria - PKU guidelines [video recording]; 2015 [cited 2016 Apr 15]. Available from: http://www.espku.org/2016/01/25/european-phenylketonuria-pku-guidelines/
9. Burton B, Grant M, Feigenbaum A, Singh R, Hendren R, Siriwardena K, et al. A randomized, placebo-controlled, double-blind study of sapropterin to treat ADHD symptoms and executive function impairment in children and adults with sapropterin-responsive phenylketonuria. Mol Genet Metab. 2015 Mar;114(3):415-24.
10. Clinical Study Report: PKU-016. A double-blind, placebo-controlled, randomized study to evaluate the safety and therapeutic effects of sapropterin dihydrochloride on neuropsychiatric symptoms in subjects with phenylketonuria [**CONFIDENTIAL** internal manufacturer's report]. Novato (CA): BioMarin Pharmaceutical Inc.; 2014 May 5.
11. Clinical study report EMR700773-003: a phase IIIb, multicenter, open-label randomized, controlled study of the efficacy, safety, and population pharmacokinetics of sapropterin dihydrochloride (Kuvan®) in phenylketonuria (PKU) patients <4 years old. [**CONFIDENTIAL** internal manufacturer's report]. Geneva: Merck Serono S.A.; 2014 Jun 25.
12. CDR submission: Kuvan resubmission, 100 mg tablets. Company: BioMarin Pharmaceutical (Canada) Inc. [**CONFIDENTIAL** manufacturer's submission]. Toronto (ON): BioMarin Pharmaceutical (Canada) Inc.; 2016 Feb 18.
13. Muntau AC, Roschinger W, Habich M, Demmelmair H, Hoffmann B, Sommerhoff CP, et al. Tetrahydrobiopterin as an alternative treatment for mild phenylketonuria. N Engl J Med. 2002 Dec 26;347(26):2122-32.

14. Longo N, Siriwardena K, Feigenbaum A, Dimmock D, Burton BK, Stockler S, et al. Long-term developmental progression in infants and young children taking sapropterin for phenylketonuria: a two-year analysis of safety and efficacy. Genet Med. 2015 May;17(5):365-73.

15. Longo N, Arnold GL, Pridjian G, Enns GM, Ficicioglu C, Parker S, et al. Long-term safety and efficacy of sapropterin: the PKUDOS registry experience. Mol Genet Metab. 2015 Apr;114(4):557-63.

16. Center for Drug Evaluation and Research, U.S. Food and Drug Administration. Medical review(s). In: Kuvan (sapropterin dihydrochloride) . Company: BioMarin Pharmaceutical, Inc. Application no.:22181. Approval date: 12/13/2007 [Internet]. Rockville (MD): FDA; 2007 Dec 13 [cited 2016 Feb 22]. (FDA drug approval package). Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2007/022181s000_MedR_P1.pdf

17. Center for Drug Evaluation and Research, U.S. Food and Drug Administration. Statistical review(s). In: Kuvan (sapropterin dihydrochloride) . Company: BioMarin Pharmaceutical, Inc. Application no.:22181. Approval date: 12/13/2007 [Internet]. Rockville (MD): FDA; 2007 Dec 13 [cited 2016 Feb 22]. (FDA drug approval package). Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2007/022181s000_StatR.pdf

18. Trefz FK, Muntau AC, Lagler FB, Moreau F, Alm J, Burlina A, et al. The Kuvan((R)) Adult Maternal Paediatric European Registry (KAMPER) multinational observational study: Baseline and 1-year data in phenylketonuria patients responsive to sapropterin. JIMD Rep [Internet]. 2015 [cited 2016 Mar 15];23:35-43. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4484899

19. Walter JH, White FJ, Hall SK, MacDonald A, Rylance G, Boneh A, et al. How practical are recommendations for dietary control in phenylketonuria? Lancet. 2002 Jul 6;360(9326):55-7.

20. Meli C, Bianca S. Dietary control of phenylketonuria. Lancet. 2002 Dec 21;360(9350):2075-6.

21. Mundy H, Lilburn M, Cousins A, Lee P. Dietary control of phenylketonuria. Lancet. 2002 Dec 21;360(9350):2076.

22. Cleary M, Trefz F, Muntau AC, Feillet F, van Spronsen FJ, Burlina A, et al. Fluctuations in phenylalanine concentrations in phenylketonuria: a review of possible relationships with outcomes. Mol Genet Metab. 2013 Dec;110(4):418-23.

23. Hood A, Grange DK, Christ SE, Steiner R, White DA. Variability in phenylalanine control predicts IQ and executive abilities in children with phenylketonuria. Mol Genet Metab [Internet]. 2014 Apr [cited 2016 Apr 4];111(4):445-51. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4144445/pdf/nihms569983.pdf

24. Hood A, Antenor-Dorsey JA, Rutlin J, Hershey T, Shimony JS, McKinstry RC, et al. Prolonged exposure to high and variable phenylalanine levels over the lifetime predicts brain white matter integrity in children with phenylketonuria. Mol Genet Metab [Internet]. 2015 Jan [cited 2016 Apr 4];114(1):19-24. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4277899/pdf/nihms642536.pdf

25. Wyrwich KW, Auguste P, Yu R, Zhang C, Dewees B, Winslow B, et al. Evaluation of neuropsychiatric function in phenylketonuria: Psychometric properties of the ADHD Rating Scale-IV and Adult ADHD Self-Report Scale Inattention Subscale in phenylketonuria. Value Health. 2015 Jun;18(4):404-12.

26. Weglage J, Ullrich K, Pietsch M, Funders B, Zass R, Koch HG. No fine motor deficits in patients with untreated non-phenylketonuria hyperphenylalaninaemia. Acta Paediatr. 1996 Mar;85(3):320-3.

27. National Institutes of Health Consensus Development Panel. National Institutes of Health Consensus Development Conference Statement: phenylketonuria: screening and management, October 16-18, 2000. Pediatrics. 2001 Oct;108(4):972-82.

28. Singh RH, Quirk ME. Using change in plasma phenylalanine concentrations and ability to liberalize diet to classify responsiveness to tetrahydrobiopterin therapy in patients with phenylketonuria. Mol Genet Metab [Internet]. 2011 Dec [cited 2016 Mar 15];104(4):485-91. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4029441

29. Trefz FK, Burton BK, Longo N, Casanova MM, Gruskin DJ, Dorenbaum A, et al. Efficacy of sapropterin dihydrochloride in increasing phenylalanine tolerance in children with phenylketonuria: a phase III, randomized, double-blind, placebo-controlled study. J Pediatr. 2009 May;154(5):700-7.

30. Viau KS, Wengreen HJ, Ernst SL, Cantor NL, Furtado LV, Longo N. Correlation of age-specific phenylalanine levels with intellectual outcome in patients with phenylketonuria. J Inherit Metab Dis. 2011 Aug;34(4):963-71.

31. van Spronsen FJ, Enns GM. Future treatment strategies in phenylketonuria. Mol Genet Metab. 2010;99 Suppl 1:S90-S95.

32. Guldberg P, Rey F, Zschocke J, Romano V, Francois B, Michiels L, et al. A 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 [Internet]. 1998 Jul [cited 2016 Mar 15];63(1):71-9. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1377241

33. van Spronsen FJ, van Rijn M, Dorgelo B, Hoeksma M, Bosch AM, Mulder MF, et al. Phenylalanine tolerance can already reliably be assessed at the age of 2 years in patients with PKU. J Inherit Metab Dis. 2009 Feb;32(1):27-31.

34. Yi SH, Singh RH. Protein substitute for children and adults with phenylketonuria. Cochrane Database Syst Rev. 2008;(4):CD004731.

35. MacDonald A. The dietary management of phenylketonuria [dissertation] [Internet]. Birmingham (United Kingdom): University of Birmingham, Faculty of Medicine and Dentistry; 1999. [cited 2016 Mar 15]. Available from: http://etheses.bham.ac.uk/29/3/MacDonald99PhD.pdf

36. MacDonald A, Rylance G, Hall SK, Asplin D, Booth IW. Factors affecting the variation in plasma phenylalanine in patients with phenylketonuria on diet. Arch Dis Child [Internet]. 1996 May [cited 2016 Mar 15];74(5):412-7. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1511531

37. MacDonald A, Chakrapani A, Hendriksz C, Daly A, Davies P, Asplin D, et al. Protein substitute dosage in PKU: how much do young patients need? Arch Dis Child [Internet]. 2006 Jul [cited 2016 Mar 15];91(7):588-93. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2082836

38. Schulz B, Bremer HJ. Nutrient intake and food consumption of adolescents and young adults with phenylketonuria. Acta Paediatr. 1995 Jul;84(7):743-8.

39. Rohr FJ, Munier AW, Levy HL. Acceptability of a new modular protein substitute for the dietary treatment of phenylketonuria. J Inherit Metab Dis. 2001 Nov;24(6):623-30.

40. MacDonald A, Lilburn M, Davies P, Evans S, Daly A, Hall SK, et al. 'Ready to drink' protein substitute is easier is for people with phenylketonuria. J Inherit Metab Dis. 2006 Aug;29(4):526-31.

41. Ney DM, Blank RD, Hansen KE. Advances in the nutritional and pharmacological management of phenylketonuria. Curr Opin Clin Nutr Metab Care [Internet]. 2014 Jan [cited 2016 Mar 24];17(1):61-8. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4004170/pdf/nihms567467.pdf

42. Bilginsoy C, Waitzman N, Leonard CO, Ernst SL. Living with phenylketonuria: perspectives of patients and their families. J Inherit Metab Dis. 2005;28(5):639-49.

43. Hvas AM, Nexo E, Nielsen JB. Vitamin B12 and vitamin B6 supplementation is needed among adults with phenylketonuria (PKU). J Inherit Metab Dis. 2006 Feb;29(1):47-53.

44. Acosta PB, Yannicelli S, Singh RH, Elsas LJ, Mofidi S, Steiner RD. Iron status of children with phenylketonuria undergoing nutrition therapy assessed by transferrin receptors. Genet Med. 2004 Mar;6(2):96-101.
45. Rose HJ, White F, MacDonald A, Rutherford PJ, Favre E. Fat intakes of children with PKU on low phenylalanine diets. J Hum Nutr Diet. 2005 Oct;18(5):395-400.
46. Beblo S, Reinhardt H, Demmelmair H, Muntau AC, Koletzko B. Effect of fish oil supplementation on fatty acid status, coordination, and fine motor skills in children with phenylketonuria. J Pediatr. 2007 May;150(5):479-84.
47. Huemer M, Huemer C, Moslinger D, Huter D, Stockler-Ipsiroglu S. Growth and body composition in children with classical phenylketonuria: results in 34 patients and review of the literature. J Inherit Metab Dis. 2007 Oct;30(5):694-9.
48. Modan-Moses D, Vered I, Schwartz G, Anikster Y, Abraham S, Segev R, et al. Peak bone mass in patients with phenylketonuria. J Inherit Metab Dis. 2007 Apr;30(2):202-8.
49. Acosta PB, Yannicelli S, Singh R, Mofidi S, Steiner R, DeVincentis E, et al. Nutrient intakes and physical growth of children with phenylketonuria undergoing nutrition therapy. J Am Diet Assoc. 2003 Sep;103(9):1167-73.
50. Schulpis KH, Papassotiriou I, Vounatsou M, Karikas GA, Tsakiris S, Chrousos GP. Morning preprandial plasma ghrelin and catecholamine concentrations in patients with phenylketonuria and normal controls: evidence for catecholamine-mediated ghrelin regulation. J Clin Endocrinol Metab. 2004 Aug;89(8):3983-7.
51. Weglage J, Pietsch M, Feldmann R, Koch HG, Zschocke J, Hoffmann G, et al. Normal clinical outcome in untreated subjects with mild hyperphenylalaninemia. Pediatr Res. 2001 Apr;49(4):532-6.
52. Mitchell JJ. Phenylalanine hydroxylase deficiency. In: Pagon RA, Adam MP, Bird TD, editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 2013 [cited 2016 Apr 15]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK1504/
53. Levy HL, Milanowski A, Chakrapani A, Cleary M, Lee P, Trefz FK, et al. Efficacy of sapropterin dihydrochloride (tetrahydrobiopterin, 6R-BH4) for reduction of phenylalanine concentration in patients with phenylketonuria: a phase III randomised placebo-controlled study. Lancet. 2007 Aug 11;370(9586):504-10.
54. Matalon KM, Acosta PB, Azen C. Role of nutrition in pregnancy with phenylketonuria and birth defects. Pediatrics. 2003 Dec;112(6 Pt 2):1534-6.
55. Ng TW, Rae A, Wright H, Gurry D, Wray J. Maternal phenylketonuria in Western Australia: pregnancy outcomes and developmental outcomes in offspring. J Paediatr Child Health. 2003 Jul;39(5):358-63.
56. Gassio R, Artuch R, Vilaseca MA, Fuste E, Boix C, Sans A, et al. Cognitive functions in classic phenylketonuria and mild hyperphenylalaninaemia: experience in a paediatric population. Dev Med Child Neurol. 2005 Jul;47(7):443-8.