Title: Emotional health in early-treated adults with phenylketonuria (PKU): Relationship with cognitive abilities and blood Phenylalanine

Liana Palermo[1,2,3], Anita MacDonald[4], Ellie Limback[1], Louise Robertson[2], Sarah Howe[2], Tarekegn Geberhiwot[2], and Cristina Romani[1]

1. School of Life and Health Sciences, Aston University, Birmingham, UK

2. Queen Elizabeth Hospital, Birmingham, UK

3. Department of Medical and Surgical Sciences, Magna Graecia University of Catanzaro, Catanzaro, Italy

4. Birmingham Children's Hospital, Birmingham, UK

Running head: Emotional health in phenylketonuria

*Address correspondence to:

Liana Palermo, Ph.D.

Department of Medical and Surgical Sciences

Magna Graecia University of Catanzaro

Viale Europa - Loc. Germaneto,

88100 Catanzaro, Italy

Email: liana.palermo@unicz.it

Tel: 0039-0961-3694362

Abstract

Introduction: Mental health, physical health and cognitive skills have been scarcely investigated in the same sample of adults with PKU (AwPKU). This is striking since emotional difficulties may potentially contribute to cognitive impairments and vice-versa. Here we aim to fill this gap.

Method: Thirty-six early-treated AwPKU and 40 controls were given an extensive battery of cognitive tasks assessing complex executive functions, inhibitory control, short –term-memory, sustained attention, visuospatial attention, language production (reading and naming), visuo-motor coordination, spoken language and orthographic processing. In addition, participants were given tasks tapping emotion recognition and completed questionnaires to assess depression (BDI-II), empathy (IRI) and mental/physical health-related quality of life (SF-36).

Results: As a group, AwPKU performed significantly worse than controls especially in tasks tapping complex executive functions and across tasks when speed was measured but did not differ for emotional-health and physical health. In the PKU group, cognitive measures and measures of physical health- related quality of life were inter-correlated (differently than in the control group), and both measures were associated with metabolic control: better metabolic control, better cognition and better physical health. Instead, cognitive measures and measures of emotional-health/ mental-health related quality of life did not correlate with one another and better metabolic control was not associated with better emotional health. Instead, some negative correlations were found. Better metabolic control was associated with worse perspective taking and more distress in socially stressful situations. Furthermore, difficulties in keeping the diet were associated with less emotional well-being.

Conclusions: Taken together, these results indicate the advantages, but also possible emotional difficulties related to maintain a PKU diet, suggesting the importance of developing alternative therapy options.

Introduction

Phenylketonuria (PKU; ORPHA716) is a rare inherited metabolic disease caused by mutations on chromosome 12 in the gene coding for the phenylalanine hydroxylase enzyme (PAH). Normally, PAH converts the amino acid phenylalanine (Phe) into tyrosine, but in PKU a PAH deficiency disrupts this process. This causes accumulation of Phe in the blood, leading to toxic concentrations in the brain as well as a lack of tyrosine and possibly a lack of dopamine, of which tyrosine is a precursor (Boot et al., 2017; Burlina et al., 2000; McKean, 1972, Puglisi-Allegra et al., 2000). Consequently, untreated PKU results in severe intellectual disabilities, microcephaly and seizures (Blau, van Spronsen, & Levy, 2010). From the 1970s, detection through neonatal screening programs and early treatment with a low-protein diet has prevented severe neurological damage and mental disabilities.

Despite early treatment, however, individuals with PKU show some mild cognitive impairments (for a review see Christ, Huijbregts, de Sonneville, & White, 2010; Hofman, Champ, Lawton, Henderson, & Dye, 2018; Janzen & Nguyen, 2010; Moyle, Fox, Arthur, Bynevelt, & Burnett, 2007; Palermo et al., 2017; Romani et al., 2017). In both children and adults with PKU, the most common *cognitive deficits* include a reduction in speed of processing, especially in visuo-attentional tasks, and impairments in higher-order executive functions (e.g., Albrecht, Garbade, & Burgard, 2009; De Roche et al., 2008; Romani, MacDonald, De Felice, & Palermo, 2018; Romani et al., 2019). Regarding the influence of blood phenylalanine control on cognitive performance, processing speed is strongly associated with Phe levels in childhood, but this association appears to lessen in adulthood (see Channon, Mockler, & Lee, 2005; Moyle, Fox, Bynevelt, Arthur, & Burnett, 2007; Romani et al., 2017; for a meta-analysis, see Albrecht et al., 2009). In contrast, current Phe levels in adulthood seem to be particularly important for sustained attention (Romani et al., 2017; Schmidt et al., 1994; Ten Hoedt et al., 2011) and learning (Romani et al., 2017; Smith, Klim, Mallozzi, & Hanley, 1996), but results so far are limited.

Emotional and behavioural difficulties clearly co-occur with cognitive impairments in untreated individuals with PKU (Brumm, Bilder, & Waisbren, 2010; Wood, Friedmann, & Steisel, 1967). Difficulties in early-treated children and adolescents have also been reported (for a review see Brumm et al., 2010; see also Jahja et al., 2013; Weglage, Fünders, Ullrich, Rupp, & Schmidt, 1996; Weglage et al., 2000), but not consistently (e.g., see Sullivan, 2001, who did not report differences in psychiatric disorders and emotional symptoms among: a). adolescents with PKU, b). adolescents with a chronic illness and c). controls). Fewer studies have investigated the presence of mental health problems in AwPKU and possible relationships with Phe levels with conflicting results. A number of these studies have reported no significant difficulties (e.g., Channon, German, Cassina, & Lee, 2004; Channon et al., 2005; Channon, Goodman, Zlotowitz, Mockler, & Lee 2007; Ris et al., 1997). Other studies, however, have suggested possible mental health difficulties, and, especially, a higher incidence of internalizing symptoms (i.e. more depression, social isolation, and anxiety), but a normal incidence of externalizing symptoms (negative behaviours towards the external environment, such as aggression; see Bilder et al., 2013; Jahja et al., 2017a,b; Manti et al., 2016; Pietz et al., 1997). Even in studies reporting a higher incidence of mental difficulties in AwPKU, however, correlations with Phe levels often either have not been present and/or have been limited to associations with early childhood levels. For example, Jahja et al. (2017a) found that adults with childhood Phe ≥ 360 µmol/L had poorer mental health in terms of depressive and somatic symptomatology that those with childhood Phe < 360 µmol/L. However, Pietz et al. (1997) found no correlation between rate of internalizing symptoms and Phe levels across ages (childhood, 12 years to adulthood, recent levels, and concurrent levels). Instead, a restrictive controlling style of parenting was a risk factor. Similarly, Manti et al. (2016) found that mental health was not associated with life-long and concurrent quality of metabolic control, nor with Phe fluctuations. Instead, paradoxically, individuals with better metabolic control (≤500 µM) in the first 11 years of life showed higher frequency of psychiatric diagnosis than patients with worse control.

It is possible that mental health problems in children and adults with PKU, if at all present, are more the consequence of the stress of growing-up and living with a chronic disease than of having high Phe levels (see also Crone et al., 2005 for an association between parenting styles and dietary control and the advocacy to develop efficient strategies to make children adhere to the diet without being too rigid). Moreover, rather than being directly related to Phe, levels emotional difficulties may be related to cognitive difficulties. Deficits in speed or processing and/or executive functions (planning, problem solving, working memory, sustained attention) may reduce someone's ability to cope with daily life demands and this, in turn, may create mental health difficulties. The opposite may also be true. We know that cognitive performance could be affected by depression, anxiety and poor self-esteem (Marazziti, Consoli, Picchetti, Carlini, & Faravelli, 2010). These relationships are worth exploring, but, so far, only few studies have examined them, with especially few studies targeting adults with PKU (Jahja et al., 2017a; Manti et al., 2016; Pietz et al., 1997; Ris et al., 1997). Moreover, generally, relationship with mental health have been examined only in terms of IQ and few other functions, with mixed results. Ris et al. (1997) found significant correlations between cognitive measures (IQ, Wisconsin Card Sorting Test, and attention) and mental health measured, but these results were not replicated by further studies. Pietz et al. (1997) did not find any relationship between IQ and mental health variables (in terms of severity or frequency of psychiatric symptoms) assessed with four scales tapping antisocial behaviour, functional and emotional behaviour (e.g. depressed mood, anxiety), emotional instability and specific symptoms (e.g. tics, stereotypes). Jahja et al. (2017a) found that AwPKU had poorer mental health than controls, even when IQ was included as a covariate. Finally, Manti et al. (2016) did not find significant correlations between IQ and psychiatric assessment scores, although patients with an IQ ≥ 85 reported less mental health problems than patients with an IQ < 85.

Finally, it is important to assess how cognitive ability and mental health relate to physical health and possible mediating associations with Phe. Most of studies have not shown a higher incidence of medical issues in PKU using the physical health related quality of life questionnaire (pHRQoL). Landolt, Nuoffer, Steinmann, and Superti-Furga (2002) and Cazzorla et al. (2014) did not find a higher incidence of pHRQoL problems in children and adolescents with PKU. Similarly, studies with AwPKU generally failed to detect pHRQoL issues (Bosch et al., 2007; Cazzorla et al., 2014; Channon et al., 2007, Simon et al., 2008). Only Bosch et al. (2015), using the PKU-specific health related quality of life questionnaire (PKU-QOL), showed that children, adolescents and adults with PKU reported tiredness more frequently than expected. Only limited and mixed information is available on relationship with Phe levels. Landlot et al. (2002) did not find any association with Phe level in the first year of life or in the 12 months before testing, but Cazzorla et al. (2014) found an overall association between health-related quality of life and Phe level at time of testing and in the 12 months beforehand. Channon et al. (2007) did not find any difference between on diet and off diet adults with PKU on pHRQoL, but Koch et al. (2002) found more medical issues, such as headaches, eczema and neurological signs. No studies that we know of have examined an interrelation between cognitive performance and physical health difficulties.

In the present study, we aim to evaluate, in AwPKU, possible inter-relationships among a) cognitive abilities, b) mental health and c) physical health and relationships between each of these variables and Phe concentrations during the life span. To reduce error and increased the reliability of our results, we have used a range of measures for each of these domains. There is consistency in the PKU literature about which measures are correlated. This depends at least in part different measures being used to tap the same function in different studies (e.g., see Hofman et al., 2018). Tasks and questions in a questionnaire are no uniquely linked to single a cognitive/emotional/physical function and different researchers have their preference. Another important factor, however, is that individual correlations are notoriously unreliable especially in the range of sample sizes used in the PKU literature (typically N < 50). To address these difficulties, our analysis will not rely on the significance of individual correlations, but instead will consider *pools* of correlations and consider whether they provide evidence for a give association when considered together (see section on data analysis). This both reduces any personal bias in the choice of tasks/measures and, more importantly, reduces the error linked to variability in individual correlations.

Cognitive abilities were evaluated in terms of an comprehensive set of indices tapping: 1. Complex executive functions (from now on EF); 2. Inhibitory control; 3. Short-term/working memory; 4. Sustained attention, 5. Visuo-spatial attention speed; 6. Language speed; 7. Visuo-motor coordination; 8. Visuo-spatial attention accuracy; 9.

Orthographic processing accuracy; 10. Spoken language accuracy; 11. Verbal learning and memory; 12 Visuo-spatial memory and learning (similar indexes were used in Palermo et al., 2017).

Emotional health was evaluated using questionnaires on depression, empathy, and quality of life related to mental health. Depression is important to examine because of reports of pathological increases in PKU and because of a potential relationship with cognitive performance. Empathy is important to examine because of its central role in social/personal relationships, which in turn are crucial for mental health. Moreover, both depression and pro-social-empathic abilities have been linked to the functioning of the dopaminergic system which may be affected in PKU. The mesolimbic dopamine pathway has been implicated in the control of mood (hypoactivity is related to altered reward processing and anhedonia, see Kulkarni & Dhir, 2009; Leggio et al., 2013) and dopamine has been argued to be important for the functioning of the "social brain" (Skuse & Gallagher, 2009) and, particularly, for empathy-motivated prosocial behaviour (Preston, 2013; Uzefovsky et al., 2014).

As a possible intermediate variable between cognitive performance and mental health, we assessed facial emotion recognition. There is a report of people with PKU showing difficulties in this domain (Jahja et al., 2016). Difficulties recognizing facial emotions may reduce empathy and, more generally, create difficulties with social interactions. This, in turn, may impact on mental health. (for a review see Kohler, Turner, Gur, & Gur, 2004). Facial emotional recognition was evaluated with three subtests from the Comprehensive Affect Test System (CATS; Froming, Levy, Schaffer, & Ekman, 2006).

Physical health was evaluated with a commonly used questionnaire assessing quality of life related to physical health. Phe concentrations were assessed both in terms of mean levels and Phe fluctuations since fluctuations have been found to be linked to cognitive performance both in our group of AwPKU (Romani et al., in press) and in children groups (Burgard et al., 1996; Arnold et al., 1998; Vilaseca et al., 2010; Hood, Grange, Christ, Steiner, & White, 2014; see for a review Cleary et al., 2013). Finally, we assessed attitudes towards diet in terms of deemed importance and ease of following it.

Materials and Methods

Participants

Early-treated patients with classical PKU were recruited from a pool of 88 individuals followed by the Department of Inherited Metabolic Disorders at the University Hospitals Birmingham who had been treated with a low-Phe diet since diagnosis in infancy and were still contactable. A total of 42 AwPKU responded to the invitation and consented to be tested. For six of them none of the mental health/emotion recognition measures were available, leaving a final sample of 36 participants. Diagnosis with newborn screening was at 5-7 days after birth. At the time of testing six AwPKU were on an unrestricted diet and 30 were still on a low-protein diet. The AwPKU were compared to a group of 40 healthy controls matched for age, gender, and education. Healthy volunteers were recruited through the Aston University volunteering website. Data on historical Phe concentrations were obtained from the PKU database at The Clinical Chemistry Department at the Birmingham Children's Hospital. Our experimental groups have a strong overlap with those reported in our previous studies: 31/36 AwPKU and 30/40 controls are the same participants assessed in Palermo et al. (2017) and Romani et al. (2017), which focused solely on cognitive functions and did not report on emotional processing and on questionnaire-derived measures of mental and physical health.

Table 1 reports demographic and metabolic information. Phe measures are reported as average Phe levels and Phe fluctuations at different ages. The average Phe level in each age band was calculated by averaging the mean Phe levels of each year and then averaging across the years in the age band. Similarly, Phe fluctuation in each age band is the average of the standard deviations of each year in the age band. Results show that Phe levels were well controlled in childhood, but diet was progressively relaxed after early childhood with increasingly high blood Phe levels, but not increasing fluctuations. Numbers are small, but AwPKU and controls did not differ in the frequency of unemployment or, for the people employed, in yearly salary. Similarly, AwPKU and controls did not differ for number of individuals in a stable relationship (64.7% vs. 46.7%; ($\chi^2(1) = 2.1$; p = .14) or being a parent (17.6% vs. 23.3% with one or more children; $\chi^2(1) = .2$; p = .61). Overall, these data indicate that AwPKU as a group are highly functional.

Insert Table 1 about here

Participants were tested in a quiet room in three separate testing sessions, two lasting approximately three hours and one lasting one hour. Blood Phe concentrations were measured on the morning of each testing session to determine current Phe levels. All participants gave voluntary informed consent to take part in the research that was approved by the NHS and Aston University Ethics committees. All efforts were made to administer all tasks to all participants, but some data points are missing because not everybody completed the three sessions (a mixture of failure to schedule another appointment for personal reasons and lack of resources on the part of the research team).

Assessment

Cognitive Abilities

IQ was measured using the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999). In addition, participants were given an extensive neuropsychological battery providing 28 different cognitive measures (described in more detail in Palermo et al., 2017). To facilitate report, *z*-scores from similar measures and/or from tasks in similar domains were aggregated into single scores (see also Arnold et al., 1998; Griffiths, Paterson, & Harvie, 1995; Nardecchia et al., 2015, Romani et al., 2017 for similar methodologies). The different domains considered by our study and relative tasks are described below[1].

Complex executive functions were calculated by averaging the z scores for the following four measures/tasks:

1. *The Tower of Hanoi puzzle* (Shallice, 1982): move rings of different sizes from an initial peg across three pegs to form a tower on the last peg; specific constraints have to be followed. Our score is based on the percentage of solved trials of different complexity (with three, four, and five rings).
2. *The Wisconsin Card Sorting Test* (Kongs, Thompson, Iverson, & Heaton, 2000): discover which rule to follow to match cards from a deck with four reference cards (e.g., according to shape, number or colour) using feedback. Our score aggregates different measures form the 64 Cards Version.

[1] Measures were aggregated following classical neuropsychological categories (e.g., visuo-spatial memory, verbal memory). EF measures were subdivided considering a commonly accepted taxonomy (see e.g., Latzman & Markon, 2010; Miyake et al., 2000).

3. *Semantic Fluency* (Rosen, 1980): say as many animals as possible in one minute of time.
4. *Trail Making Test B-A* (Reitan, 1958). This measures the difference between the time it takes participants to connect alternating number and letter circles (Trail Making B) and the time it takes to connect circles containing only numbers (Trail Making A).

Inhibitory Control was calculated by averaging the z scores for the following two measures/tasks:

1. *The Stroop colour–word test: Interference* (Stroop, 1935). This measured the difference in both time and errors between reporting the ink colour of coloured words where the colour of the ink was either incongruent with the meaning of the word (e.g. ‘red’ written with yellow ink), or congruent with the meaning of the word (e.g. ‘red’ written with red ink).
2. *Picture naming: Semantic interference.* This measured the differences in both RT and error rates between naming the first and last exemplar of a series of semantically related nouns.

Short-term memory memory was calculated by averaging the z scores for the following three tasks

1. *Digit Span*; repeat a sequence of digits spoken by the examiner, immediately after presentation).
2. *Nonword Repetition* (Romani, Tsouknida, & Olson, 2015): repeat a sequence of nonwords spoken by the examiner, immediately after presentation).
3. *Corsi Block Span* (Corsi, 1972): the examiner taps a sequence of blocks and the participant has to reproduce the sequence in the same order.

Sustained attention was measured using the *Rapid Visual Information Processing* (RVP) task (adapted from Sahakian, Jones, Levy, Gray, & Warburton, 1989) in which the participant must detect three target sequences of three digits in a continuously presented sequence of digits. The digits are presented on the computer screen and the participant has to press the response key when the last number of the sequence appears on the screen. The score is the percentage of correct responses.

Visuo-spatial attention-speed was calculated by averaging the z scores for mean correct RTs in the following six tasks:

1. *Simple Reaction Time*: press a button as soon as a ladybird appears on the screen.
2. *Detection with distractors*: press a button whenever a ladybird appears on the screen alone or with a green bug.
3. *Opposite Detection with distractors*: press a button whenever a green bug appears on the screen alone or with a ladybird.
4. *Choice Reaction Time* (from Cambridge Neuropsychological Test Automated Battery, CANTAB): press either a left or right response key, consistent with the direction of an arrow presented centrally.
5. *Feature Search:* detect a target among distractors not sharing features by pressing a ʻyes' or ʻno' button (e.g. a red ladybird among green bugs).
6. *Conjunction Search:* detect a target among distractors sharing features (e.g. red ladybird among red bugs and green ladybirds).

Visuo-spatial attention-accuracy was calculated by averaging accuracy z scores in *Detection with distractors, Opposite Detection with distractors, Choice Reaction Time, Feature Search and Conjunction Search* tasks described above (simple Detection excluded because of lack of errors

Language production -speed was calculated by averaging the z scores for mean correct RTs in the following four tasks:

1. *Word Reading* (adapted from Romani et al., 2015): read an English word as fast as possible.
2. *Nonword Reading* (adapted from Romani et al., 2015): read a nonword as fast as possible.

3. *Picture Naming:* name a picture as fast as possible.
4. *Colour Naming*: name, as fast as possible, the ink colour of a sequence of ‘X’ letters (i.e. XXX) or coloured words where the colour of the ink matched the meaning of the word (e.g. ‘red’ written with red ink).

Spoken language accuracy. This was calculated by averaging the accuracy z scores for the following four tasks:

Picture naming: name a picture as fast as possible;

1. *Color naming*: name as fast as possible the ink color of three X’s or colored words; only the congruent condition is considered here where the color of the ink matched the meaning of the word, for example, red written with red ink;
2. *Similarities* from the WASI (Wechsler, 1999): describe how similar in meaning two words are;
3. *Vocabulary* from the WASI (Wechsler, 1999): define a word.

Orthographic processing-accuracy. This was calculated by averaging the accuracy z scores for the following four tasks (phoneme deletion and spoonerisms are included in this index since they are strongly related to orthographic abilities):

1. *Word reading and Nonword reading* (adapted from Romani et al., 2015): read an English word/non-word as fast as possible (error rates was taken);
2. *Word Spelling and Nonword spelling* (Romani et al., 2015): spell words/nonwords to dictation;
3. *Phoneme deletion*: delete a sound from a word (e.g., powder; /d/ power);
4. *Spoonerisms (*adapted from Romani et al., 2015): exchange the initial sounds of two words to produce two newwords (e.g., bad-sin: sad-bin).

Visuo-spatial memory and learning was calculated by averaging the z scores of the following measures/tasks:

1. *Delayed Matching to Sample* (DMS; adapted from Sahakian et al., 1988): recognize a previously seen pattern among distracters.
2. *Paired Associates Visual Learning* (PAL; adapted from Sahakian et al., 1988): learn the association between objects and locations on the computer screen.

Verbal memory and learning was calculated by averaging the z scores of the following measures/tasks:

1. *The Rey Auditory Verbal Learning Test* (Rey, 1964; Schmidt, 1996) entailed learning, immediate recall and delayed recall of a list of 15 words.
2. *Paired Associates Verbal Learning* (Romani et al., 2015) involved learning and delayed recall of a nonword associated with the picture of an object or animal.

Visuo-motor coordination was calculated by averaging the z scores of the following tasks:

1. *Digit Symbol* (Wechsler, 1997): the participant is given 90 seconds to fill as many boxes as possible with symbols corresponding to number according to a key.
2. *Grooved Pegboard Test* (Trites, 1977): put pegs into the holes of a board as quickly as possible using only one hand.

Emotional Health and mental health-related quality of life

Emotional health and mental health-related quality of life were assessed with three types of standardized questionnaires.

The Beck Depression Inventory II (BDI-II; Beck, Steer, & Brown, 1996) evaluates the presence of depression. It includes 21 items, with each item assessing a specific symptom of depression (e.g. guilt, self-dislike, crying). Each item consists of four statements, and the participant must select which statement best applies to how he has been feeling in the last two weeks. Statements are graded for severity on a 0-3 scale for a maximum total score of 63. According to the standardized cut-offs, scoring 0–13 indicates minimal depression, 14–19 mild depression, 20–28 moderate depression, and 29–63 severe depression.

The **Interpersonal Reactivity Index** (IRI; Davis, 1980) evaluates empathy. It consists of four seven-item subscales. 1. The Perspective-Taking subscale assesses the tendency to adopt the point of view of other people; 2. the Fantasy subscale measures the tendency to identify with fictional characters; 3. the Empathic Concern subscale measures feelings of warmth, compassion, and concern for others, and 4. the Personal Distress subscale measures personal feelings of anxiety and discomfort in tense interpersonal situations. The Empathic Concern and the Personal Distress subscales assess affective aspects of empathy while the Perspective-Taking and the Fantasy subscales assess cognitive aspects. Participants had to rate how well each item of the scale described them using a 5-point Likert scale, ranging from 0 (doesn't describe me at all) to 4 (describes me very well). High scores on the Perspective-Taking, Fantasy and Empathic Concern subscales are associated with a pro-social tendency (a greater ability to take someone else's perspective being associated with more concern for others). Scores on these scales increase from childhood to adulthood. High scores on Personal Distress, instead, are associated with emotional difficulties (e.g., social anxiety and neuroticism) and scores consistently decrease with age (Batson, Early, & Salvanari, 1997; De Corte et al., 2007; Davis & Franzoi, 1991; Eisenberg & Eggum, 2009; for the Fantasy subscale see Hawk et al., 2013).

The Mental Health Quality of Life scale from SF-36 (Ware, Snow, Kosinski, & Gandek, 1993) assesses mental health-related quality of life (Farivar, Cunningham, & Hays, 2007). It includes four subscales assessing: 1. Vitality, 2. Social functioning, 3. Role limitations due to emotional problems and 4. Emotional well-being. Overall, the scale includes ten questions assessed on a 5/6-point Likert scale and three dichotomous (yes-no) questions. For each subscale, scores range from 0 to 100 with higher scores reflecting a better health state.

Emotion Recognition

This was investigated with three subtests from the Comprehensive Affect Test System (CATS; Froming, Levy, Schaffer, & Ekman, 2006), which uses stimuli from Ekman's internationally-normed faces collection. The Identity Discrimination subtest asks participants to decide if two faces represent the same or different individuals (22 trials with emotionally neutral faces) in order to exclude problems in face processing. The Three Faces Test asks participants to select the two portraits that express the same emotion (36 trials). The Select Affect Test asks participants to match the written name of an emotion with one out of five faces expressing different emotions (20 trials). For all subtests, we recorded number of correct answers and response latencies.

Physical health-related quality of life

This was examined with The Physical Health Quality of Life scale from SF-36 (Ware et al., 1993) which comprises the following four sub-scales: 1. Physical function, 2. Role limitations due to physical health problems, 3. Bodily pain, 4. General health. This scale includes 18 questions assessed on a 3/6-point Likert scales and four dichotomous (yes-no) questions. For each subscale, scores range from 0 to 100 with a higher score reflecting a better health state. The questionnaire was scored as suggested in Hays, Sherbourne & Mazel (1993). The scoring instructions are free available at www.rand.org/health-care/surveys_tools/mos/36-item-short-form/scoring.html.

Data Analysis

We will assess differences between the PKU and the control group with t-tests.

We will measure associations between different measures through bivariate Pearson r correlations. To facilitate interpretation, in all our analyses and for all our measures, higher scores will reflect better performance. We will refer to correlations as ‘positive’ if they meet expectations and better scores for one measure are associated with better scores in the other measure. We will refer to correlations as ‘negative’ if their value is in the opposite than expected direction.

Crucially, in order to assess associations of interest in a more reliable way, we will consider patterns of correlations in a given domain (1. cognitive-health, 2. emotional-health or 3. physical health) as well as compare patterns across domains. For this purpose, we will carry out three types of analyses.

1. <u>Within each domain</u>, we will:
   a. With χ^2s, compare the rate of positive vs negative correlations relative to a chance 50/50 rate; if there is an association, the rate of positive correlations will be significantly higher than chance, independently of whether any individual correlation reaches a significant difference from 0;
   b. As above, compare the rate of positive vs negative *significant* correlations; if there is a real association, the rate of significant positive correlations will be higher. Instead, if there is no real association --and the significant correlations are obtained by chance--, significant positive correlations will be counter-balanced by an equal number of significant negative correlations;
   c. With a one-sample t test, assess whether, the mean of the pool of correlations is significantly different from 0.

2. <u>Across domains</u>, we will:

   a-b) With χ^2s, compare rate of positive vs negative correlations and the rate of significant positive and negative correlations;

   c) With ANOVAS assess differences in the mean size of the correlation among domains (Bonferroni correction will be used for post-hoc comparisons).

Before running t-test/ANOVAS, the correlation coefficients will be Fisher Z transformed, as it is recommended (see Press, Teukolsky, Vetterling, & Flannery, 1992). Note that analyses which compare number/size of pools of correlations are standard in some field of investigation, although not used so far in the PKU literature (see neuroimaging studies; e.g. Boccia et al., 2019; Evans & Davis, 2015; Kriegeskorte, Mur, & Bandettin, 2008; Long et al., 2008; Smyser et al., 2016).

Results

General Performance

Cognitive Abilities. Results are presented in Table 2. AwPKU had a full-scale IQ in the average range (only one patient showed signs of impairment), but significantly lower than matched controls.

Insert Table 2 about here

Considering the whole neuropsychological assessment together, our PKU sample showed a performance only slightly lower than controls (0.5 *z*-scores below mean). Results, however, were very different for different functions. The areas showing most substantial difficulties were: 1. speed of processing which showed a systematic group difficulty across cognitive tasks (language and visuo-spatial attention); 2. complex executive functions and 3. visuo-motor coordination. The PKU sample showed also some difficulties in short-term memory. Performance in sustained attention was not statistically different from that of controls, but this may be due to the smallness of the sample size. No group difficulties were found in inhibitory control, verbal learning and memory, visuo-spatial learning and memory, and accuracy measures across verbal and visuo-spatial tasks, with the exception of spoken language accuracy

Results for Emotion Recognition, Emotional Health, and /Physical Health-related Quality of Life are reported in Table 3.

Insert Table 3 about here

Emotional Health and mental health-related quality of life. Scores on the BDI-II did not differ from controls. Indeed, only one of the 34 AwPKU (=2.9% of the sample) scored in the 'moderate depression' range and one in the 'severe depression' range (=2.9% of the sample). Scores on all the subscales of the Interpersonal Reactivity Index (IRI) and of Mental health-related quality of life scale also did not differ from controls.

***Emotion Recognition*.** AwPKU were as accurate as controls in recognizing faces (Identity Discrimination) or discriminating emotions (Three Faces Test and Select Affect Test), but they were slower.

Physical health-related quality of life. There were no significant differences from controls in terms of Physical Health-related Quality of Life.

Associations between Cognitive Abilities and Emotional Health/ Physical health-related quality of life

Correlations in the PKU group between our cognitive indices on the one hand and emotion recognition, emotional health and physical health-related quality of life, on the other are shown in Table 4.

Insert Table 4 about here

Cognitive Abilities and Emotional Health/ mental health-related quality of life. In the PKU group, results showed no association between cognitive measures and measures of emotional health/ mental-health related quality of life. Out of 108 correlations, only 52% were positive (better cognition better emotional health) which is not statistically different from chance (56/108; $\chi^2 = .02$; $p = .89$). There were only a few significant correlations (n = 13) and a similar number was positive or negative (8 vs. 5 $\chi^2 = 2.45$; $p =. 12$). Finally, the overall average correlation was close to 0 (r = .01; one-sample t-test; $t(107) = .55$; $p = .58$). With the IRI Empathy scale, there were eight significant, positive correlations (a higher desirable trait with better cognition), and two significant negative correlations. With the Quality of life scale, there was a single significant positive correlation and two significant negative correlations.

Correlations between cognitive measures and measures of emotional health/ mental health-related quality of life were similarly limited in the control group. Out of 108 total correlations, 59 were negative and 49 were positive ($\chi^2 = .32$; $p = .57$). Only very few correlations were significant: two positive (between Verbal Memory and Fantasy; r = 0.42; $p = .03$ and between Verbal Memory and Social Functioning; r = 0.36; $p = .05$) and two negative (between Distress and Verbal Learning and Memory; r= -0.39; $p = .04$; and between Sustained attention and Social Functioning; r = -0.39; $p < .03$). The overall average correlation was close to 0 (r = -.01; one-sample t-test: $t(107)= .63$; $p = .53$) and correlation coefficients between cognitive functions and emotional health were no different in the two groups ($t(214)= .03$; $p = .97$).

Cognitive Abilities and Physical ***health-related quality of life***. In the PKU group, there were widespread correlations between cognitive measures and measures of physical health, indicating that AwPKU who performed better in cognitive tasks also enjoyed a better quality of life related to physical health. Out of 48 correlations, 98% were positive (47/48; $\chi^2 = 57.4$; $p < .001$) and out of the significant correlations, 100% were positive (23/23; $p < .001$ Fisher's exact test). The overall average correlation was r=.34 which is significantly different from 0 (one-sample t-test: $t(47)= 11.6$; $p < .001$). Significant correlations were scattered across all the subscales, but were more prevalent in the subscales measuring Physical Functioning (which asks about limitations in vigorous and moderate activities, such as lifting objects, participating in strenuous sports, climbing stairs and walking different distances), Physical Health, which asks about effects on working life and Bodily Pain. There were only two significant correlations in the General Health subscale which asks for feeling of being healthy compared to others and tendency to get ill.

In contrast, in the control group, there was only one significant, negative correlation between STM and Physical Function (r = -0.58; p = .001) and the average of the correlation coefficients was close to 0 (r = .01; one-sample t-test: $t(47) = .15$; p = .88). Therefore, although both the PKU and the control group were generally in good health, correlation coefficients between cognitive functions and physical health were significantly higher in the PKU group than in the control groups ($t(94)= -8.88$; $p < .001$) and only in AwPKU cognition and physical health were associated with one another.

Associations with Phe Measures

Table 5 shows correlations between Phe measures at different ages on the one hand, and cognitive measures on the other (FSIQ and our composite cognitive scores; emotion recognition is not included given that it is not impaired and not associated with measures of cognitive/emotional or physical health). Table 6 shows correlations between Phe concentrations at different ages on the one hand, and measures of emotional and physical health on the other.

Insert Table 5 and 6 about here

Cognitive measures. Out of all 91 correlations, 89% were positive (81/91; $\chi^2 = 34$; $p < .001$); better cognition was associated with better metabolic control. There was a high number of significant correlations and out of them, 100% were positive (30/30; $p < .001$ Fisher's exact test) indicating better cognition with reduced Phe levels. Out of 30 significant correlations 6 were with historical average Phe level, 3 with current Phe and 21 with historical variation of Phe. The average overall r correlation coefficient was 0.25 which is significantly different from 0 (one-sample t-test: $t(83) = 11.22$; $p < .001$)

Emotional health/ mental healt-related quality of life. Out of 63 correlations, only 41% were positive which is not statistically different from chance (26/63; $\chi^2 = 1.3$; $p = .25$). There was a small number of significant correlations and most were negative (4/5), that is, better metabolic control was associated with more emotional difficulties. The average overall r correlation coefficient was -.06 which is significantly different from 0, but in the direction opposite to that expected (better mental health, worse metabolic control; one-sample t-test: $t(62) = -2.8$; $p = .04$). With the IRI, lower Phe level in childhood and adolescence were associated with 1. a lower tendency to adopt someone else's point of view (Perspective scale[2]); and 2. more anxiety and discomfort from observing another person's negative experience (Personal Distress scale). There was only one significant, positive correlation: lower Phe fluctuations in adulthood was associated with more Vitality.

Physical health-related quality of life. Out of 28 correlations, 86% were positive: better physical health was associated with lower Phe levels and reduced fluctuations (24/28; $\chi^2(1) = 28.1$; $p = .001$). There was a high proportion of significant correlations (13/28 = 46%) and, out of them, 100% were positive (13/13; $p < .001$ Fisher's exact test). The average overall r correlation coefficient was .30 which is significantly different from 0 (one-sample t-test: $t(27)= 6.23$; $p < .001$).

Comparisons across domains. Chi-square analyses showed that more significant, positive correlations were obtained between metabolic variables and either cognitive or physical health measures than between metabolic variables and measures of emotional health (33% and

[2] Note this has a negative connotation because it is from an emotional/empathic perspective not a reasoning perspective.

46% vs 1.6%; $\chi^2 = 31.2$ and 50.7, respectively; $p < .001$ for both). A three-way ANOVA with domain as the independent variable (Cognition, Emotional-health/ mental health-related quality of life, Physical health-related quality of life) and correlation coefficient with Phe measures as the dependent variable showed a highly significant difference ($F(2,12) = 26.50$; $p < .001$; $\eta p^2 = .82$). Post-hoc comparisons, after Bonferroni correction, showed that the strength of the correlation with Phe measures was significantly higher for the physical health and cognitive measures than for the emotional health measures ($p < .01$). Instead, the strength of the correlation between Phe measures and either physical health or cognitive measures was not significantly different ($p = .91$).

Attitudes Towards Diet

We evaluated attitudes towards the PKU diet with two questions to be answered using a seven 7-point Likert scale. We asked: 1. "How *important* is it for you to follow the diet?" with responses ranging from 1 (not important at all) to 7 (very important); and 2. "How *easy* is it for you to follow the diet?" with responses ranging from 1 (not easy at all) to 7 (very easy). In response to question 1, the average score was 5.9 (SD =1.2; range: 3-7), indicating that most participants thought it important to follow the diet. In response to question 2, the average score was 4.6 (SD =1.6; range: 1-7), indicating that on average participants found relatively easy to follow the diet, but variability was high. Table 7 shows correlations between attitudes towards diet and measures of emotional health and physical health. There was a lack of significant correlations between how important it is deemed to be on diet. However, there were two significant correlations with ease of keeping to the diet. The easier our participants found being on diet, the lower the depressive symptomatology and the higher the emotional well-being.

Insert Table 7 about here

Discussion

We evaluated emotional health, physical health and cognitive abilities in a sample of adults with PKU, as well as their inter-relations and relations with Phe measures. We had described cognitive performance in an overlapping sample of AwPKU in previous publications (Palermo et al., 2017; Romani et al., 2017). Here we report relationships with emotional health and physical health measures. AwPKU had a full-scale IQ in the average range, but significantly lower than that of controls matched for education[3]. This result is consistent with previous studies and suggest that AwPKU do not reach their full potential (see also DeRoche & Welsh, 2008; Moyle, Fox, Arthur et al., 2007 for a meta-analysis).

Considering the whole neuropsychological assessment together, as a group, AwPKU performed well in a variety of cognitive domains, but also showed specific deficits, mainly involving speed of processing (in visuo-spatial tasks, language tasks, tasks involving emotional recognition) and executive functions (EFs) such as planning, reasoning and monitoring (poor performance on Tower of Hanoi, WCST, semantic fluency, Trail-making test, Similarities of the WASI, STM tasks). Performance was also poor in tasks involving visuo-motor coordination such as the peg-board task and the digit symbol tasks (which also taps EFs). Performance in sustained attention was not statistically different from that of controls, but this may be due to the smallness of the simple size since significant results were found in Palermo et al. (2017) with a

[3] . The control group had an average IQ above 100, but this is expected given that participants engaging in research often have an educational level higher than that of the average population.

larger sample that included six more participants and with a mixed English and Italian sample (see Romani et al., in press). No group difficulties were found in inhibitory control, verbal learning and memory, visuo-spatial learning and memory, and accuracy measures across verbal and visuo-spatial tasks, with exception of spoken language accuracy where the index includes tasks with a strong abstract reasoning component and/or strongly related to general intelligence (Similarities and Vocabulary and form the WASI; Jensen, 2001; McCrimmon & Smith, 2013).

Variability, however, was high, with about 1/3 of participants showing average or above average performance (see Palermo et al., 2017 for more details). Cognitive capacities were modulated by Phe levels and Phe fluctuations. Constancy in maintaining a diet was, if anything, more important than average levels since fluctuations had an even stronger negative effect on performance (see also Cleary et al., 2013; Romani et al., in press).

As a group, AwPKU showed no evidence of ill physical health, with self-reported scores of quality of life related to physical health being very similar to those of the controls. However, variations in physical health correlated with cognition in the PKU group, but not in the control group. This is likely to be due to that in AwPKU both physical health and cognitive abilities are mediated by an intermediate variable which is blood Phe. In fact, in our sample, AwPKU with lower Phe levels enjoyed both better cognition and better physical health, this result being consistent with reports of more medical issues, such as headaches, eczema and neurological signs, in participants who relax the diet compared with participants on diet (see Koch et al., 2002). It is unlikely that causation would be in the opposite direction (less subjective physical health, less ability to keep on the diet) because we found no relation between the physical health related quality of life questionnaire and our question concerning how easy it was to keep on diet.

Taken together, these results make a case for maintaining low Phe levels through life (see also Weglage et al., 2013 for an indication of the cognitive benefits of maintaining a diet in adulthood; in this study, older patients who relaxed the diet in adolescence showed worse cognitive performance than younger patients who maintained better control although differences could also be due to poorer treatment options/support available for the older patient which were treated in the 1970s and 80s).

As a group, our PKU participants showed no significant difficulties in emotional recognition accuracy. They were slower that controls in discriminating emotions, but they were also slower in the identity discrimination task which does not involve emotional processing, suggesting a general speed reduction in processing of visuo-perceptual stimuli rather than specific difficulties in processing emotions (see also Romani et al., 2017). They also showed no increased symptoms of depression, a good ability to show empathy and no problems with quality of life related to emotional health. Moreover, when compared to controls matched for gender and education, they showed the same incidence of individuals in a stable relationship and/or with children, and no difference in income for those employed. Our sample is too small to derive strong demographic conclusions. Nevertheless, taken together, our results suggest that, as a group, early treated AwPKU function very well (see also Bosch et al., 2007 and Simon et al., 2008 for similar results showing good quality of life, social capacity and sociodemographic outcomes in AwPKU).

Our results are consistent with those of previous studies that also reported no emotional difficulties in AwPKU (mainly on UK samples of patients: see Channon et al., 2004; 2005; 2007), but at odds with other studies which, instead, detected some emotional difficulties both in early treated children and adults with PKU (e.g., Jahja et al., 2017a,b; Manti et al. 2016; Pietz et al., 1997). There could be different possibilities of these discrepancies. First of all, we should note that the range of possible difficulties that we measured was limited. For example, we did not measure anxiety. Another possible reason is the good metabolic control in our group of

AwPKU compared to others described in the literature (average Phe level at testing time 670 µmol/L) which may have limited the scope for impairment. Alternatively, presence/absence of emotional difficulties may be due to differences in clinical management across centres with different levels of psychological support received by the patients and/or to socio-cultural differences in style of parenting and in the centrality of food for socialization. It is also possible that in PKU some genetic mutations, but not others are linked to emotional difficulties. This is very speculative, but the mutation R261Q is relatively prevalent in the Netherlands (for a review see Zschocke, 2003), and studies with Dutch patients seem to report more mental difficulties than studies on other populations.

We have found no positive association between measures of emotional-health and either cognitive abilities or measures of metabolic control. Relationship with cognitive abilities have not been assessed before. Importantly, there were no significant correlations between cognition and the depression questionnaire. This suggests that in our participants depression had no role in reducing speed of processing and/or in affecting executive functions. Overall, we found a small negative correlation between emotional health and metabolic control (-.06) which was significantly different from 0 and a small number of significant individual negative correlations (4/5), indicating more emotional difficulties with better metabolic control. There was only one significant, positive correlation between Phe fluctuations in adulthood and Vitality. Vitality, however, is strongly related to physical health (see Ware, Kosinski, & Keller, 1994) and thus, this correlation fits with the widespread correlations between Phe levels and physical health discussed above. These results are marginal, but they point to the emotional costs of keeping on a strict life-long diet.

Other studies have found no positive relation between emotional health and metabolic control (e.g., Manti et al., 2016; Pietz et al., 1997; Ris et al., 1997), but, here again, there have been exceptions (e.g., Bilder et al., 2013). As customary, we have assessed emotional health through self-report questionnaires and it is also possible that people with better metabolic control are more aware and better able to report difficulties than people with worse control, cancelling some differences out. However, our comparisons of patterns of correlations across domains clearly indicates that the impact of metabolic control on the emotional domains examined is much weaker than on cognitive abilities and physical health. We can speculate that the dopaminergic pathways involved in emotional processing (the *mesolimbic pathway*, which connects the ventro-tegmental area to the limbic system) is less sensitive to dopamine depletion that the pathways involved in cognitive processing (the *mesocortical pathway*, which connects the ventro-tegmental area to the prefrontal cortex; see Wise, 2004; but see Björklund & Dunnett, 2007 for anatomical overlap of these networks). If so, a relation between Phe level and emotional health would emerge only in a sample of participants with higher Phe concentrations than our own.

Finally, one should consider, that emotional difficulties, whenever present, may be related more to the stress of living with a chronic disease or a restrictive/controlling style of parenting than to high levels of Phe (see also Pietz et al., 1997; Manti et al., 2016). Higher scores in perspective taking are indicative of adaptive and prosocial affective responses (Batson, Fultz, & Schoenrade, 1987; Batson et al., 1997). In our sample, a better ability to take other people's perspectives and less distress in emotionally tense interpersonal situations were associated with higher Phe levels and, therefore, less dietary control in childhood and adolescence. It is possible that those who find harder to keep on a diet (and have less dietary control) develop more empathy for others. It is also possible that those who have a more relaxed approach to diet have a better socialization and less social anxiety. Keeping a strict diet may be particularly difficult in adolescence when socialization though food is important. Consistent with this possibility, we found that those who reported harder to keep on a diet had worse emotional well-being (although this relation could also suggest that, vice-versa, the presence of emotional difficulties can make harder to adhere to the diet). The diverse emotional reactions caused by keeping a PKU diet are well exemplified by the different answers given by two of our PKU participants when we asked them to provide any comments about the diet as part of our Attitudes Toward Diet Questionnaire. Participant A, who decided to resume the diet after stopping it for a while said: "When I was off diet I felt like a monster. I was depressed and anxious". Participant B, who decided to stop the diet said; "I felt worse on diet, very awkward at work and I was very low". These considerations underline the difficulties for some AwPKU to stay on diet and highlight the importance of developing alternative therapies to offer to the people with PKU more options of treatment (e.g., enzyme substitution therapy; see Pascucci et al., 2018; Thomas et al., 2018).

In conclusion, we found that cognitive measures and measures of quality of life related to physical health were strongly associated to dietary control in our sample of early-treated adults with PKU. In contrast, we found no evidence that emotional difficulties are a core issue in this population and only limited (and negative) associations with metabolic control. These null results could have several causes, but they highlight that cognitive measures are more sensitive indicators of outcomes related to dietary control than emotional health measures. More speculatively, our results indicate that, in some adults with PKU, the stress of being on a diet counterbalances the possible positive effects on mood.

Funding

This study was funded by a Marie Curie Intra European Fellowship within the 7th European Community Framework Programme granted to Liana Palermo under the supervision of Cristina Romani (TREPAPHEN; 329423) and by a grant of the University Hospital Birmingham Charity to Tarekegn Gerberhiwot.

Disclosure of Interest

Anita MacDonald has received research funding and honoraria from Nutricia, Vitaflo International and Merck Serono. She is a member of the European Nutrition Expert Panel (Biomarin), member of Sapropterin Advisory Board (Biomarin), member of the Advisory Board entitled ELEMENT (Danone-Nutricia), and member of an Advisory Board for Arla and Applied Pharma Research. Liana Palermo has received honoraria as speaker from Biomarin and Nutricia. Louise Robertson has received honoraria as speaker/member of advisory boards from Nutricia and a sponsorship from Vitaflo to attend a conference.

Tarekegn Geberhiwot, Sarah Howe, Ellie Limback and Cristina Romani declare that they have no conflict of interest.

References

1. Albrecht, J., Garbade, S. F., & Burgard, P. (2009). Neuropsychological speed tests and blood phenylalanine levels in patients with phenylketonuria: a meta-analysis. *Neuroscience and Biobehavioral Reviews, 33,* 414–421.
2. Arnold, G.L., Kramer, B.M., Kirby, R.S., Plumeau, P.B., Blakely, E.M., Sanger Cregan, L.S., & Davidson, P.W. (1998). Factors affecting cognitive, motor, behavioral and executive functioning in children with phenylketonuria. *Acta Paediatrica, 87*, 565-570.
3. Batson, C. D., Early, S., & Salvanari, G. (1997). Perspective taking: Imagining how another feels versus imagining how you would feel. *Personality and Social Psychology Bulletin, 23*, 751–758.
4. Batson, C. D., Fultz, J., & Schoenrade, P. A. (1987). Distress and empathy: Two qualitatively distinct vicarious emotions with different motivational consequences. *Journal of Personality, 55,* 19–39.
5. Beck, A.T., Steer, R.A., & Brown, G.K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.
6. Björklund, A., & Dunnett, S.B. (2007). Dopamine neuron systems in the brain: an update. *Trends in Neuroscience, 30*, 194-202.
7. Bilder, D.A., Burton, B.K., Coon, H., Leviton, L., Ashworth, J., Lundy, B.D., Vespa, H., Bakian, A.V., Longo, N. (2013). Psychiatric symptoms in adults with phenylketonuria. *Molecular Genetics and Metabolism, 108*, 155-160
8. Blau, N., van Spronsen, F. J., & Levy, H. L. (2010). Phenylketonuria. *Lancet, 376*, 1417–27.

9. Boccia, M., Sulpizio, V., Teghil, A., Palermo, L., Piccardi, L., Galati, G., & Guariglia, C. (2019). The dynamic contribution of the high- level visual cortex to imagery and perception. *Human brain mapping, 40,* 2449-2463.
10. Boot, E., Hollak, C. E. M., Huijbregts, S. C. J., Jahja, R., van Vliet, D., Nederveen, A. J., ... & Abeling, N. G. G. M. (2017). Cerebral dopamine deficiency, plasma monoamine alterations and neurocognitive deficits in adults with phenylketonuria. *Psychological Medicine*, 1-12.
11. Bosch, A. M., Burlina, A., Cunningham, A., Bettiol, E., Moreau-Stucker, F., Koledova, E., ... & Regnault, A. (2015). Assessment of the impact of phenylketonuria and its treatment on quality of life of patients and parents from seven European countries. *Orphanet journal of rare diseases, 10*, 80.
12. Bosch, A.M., Tybout, W., van Spronsen, F.J., de Valk, H.W., Wijburg, F.A., & Grootenhuis, M.A. (2007). The course of life and quality of life of early and continuously treated Dutch patients with phenylketonuria. *Journal of inherited metabolic disease*, 30, 29–34.
13. Brumm, V., Azen, C., Moats, R., Stern, A.M., Broomand, C., Nelson, M.D., & Koch, R. (2004). Neuropsychological outcome of subjects participating in the PKU adult collaborative study: a preliminary review. *Journal of inherited metabolic disease, 27,* 549-566.
14. Brumm, V.L., Bilder, D., & Waisbren, S.E. (2010). Psychiatric symptoms and disorders in phenylketonuria. *Molecular Genetics and Metabolism, 99 Suppl 1*, S59-63.
15. Burgard, P., Rupp, A., Konecki, D.S., Trefz, F.K., Schmidt, H., & Lichter-Konecki, U. (1996). Phenylalanine hydroxylase genotypes, predicted residual enzyme activity and

phenotypic parameters of diagnosis and treatment of phenylketonuria. *European Journal of Pediatrics, 155,* S11-S15.

16. Burlina, A., Bonafé, L., Ferrari, V., Suppiej, A., Zacchello, F., & Burlina, A. (2000). Measurement of neurotransmitter metabolites in the cerebrospinal fluid of phenylketonuric patients under dietary treatment. *Journal of Inherited Metabolic Disease, 23*, 313–316.
17. Cazzorla, C., Cegolon, L., Burlina, A. P., Celato, A., Massa, P., Giordano, L., ... & Burlina, A. B. (2014). Quality of Life (QoL) assessment in a cohort of patients with phenylketonuria. *BMC Public Health, 14*, 1243.
18. Channon, S., German, E., Cassina, C., & Lee, P. (2004). Executive functioning, memory, and learning in phenylketonuria. *Neuropsychology, 18*, 613–620.
19. Channon, S., Goodman, G., Zlotowitz, S., Mockler, C., & Lee, P. J. (2007). Effects of dietary management of phenylketonuria on long-term cognitive outcome. *Archives of Disease in Childhood, 92*, 213–218.
20. Channon, S., Mockler, C., & Lee, P. (2005). Executive functioning and speed of processing in phenylketonuria. *Neuropsychology, 19*, 679–686.
21. Christ, S. E., Huijbregts, S. C. J., de Sonneville, L. M. J., & White, D. A. (2010). Executive function in early-treated phenylketonuria: profile and underlying mechanisms. *Molecular Genetics and Metabolism, 99 Suppl 1,* S22–32.
22. Cleary, M., Trefz, F., Muntau, A. C., Feillet, F., van Spronsen, F. J., Burlina, A. … MacDonald, A. (2013). Fluctuations in phenylalanine concentrations in phenylketonuria: a review of possible relationships with outcomes. *Molecular Genetics and Metabolism, 110*, 418–423.

23. Corsi, P. M. (1972). Human memory and the medial temporal region of the brain. McGill University, Montreal: Unpublished doctoral dissertation.

24. Crone, M.R., van Spronsen, F.J., Oudshoorn, K., Bekhof, J., van Rijn, G., & Verkerk PH. (2005). Behavioural factors related to metabolic control in patients with phenylketonuria. *Journal of inherited metabolic disease*, *28,* 627-637.

25. Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10,* 85.

26. Davis, M. H., & Franzoi, S. (1991). Stability and change in adolescent selfconsciousness and empathy. *Journal of Research in Personality, 25,* 70–87.

27. De Corte, K., Buysse, A., Verhofstadt, L. L., Roeyers, H., Ponnet, K., & Davis, M. H. (2007). Measuring empathic tendencies: Reliability and validity of the Dutch version of the Interpersonal Reactivity Index. *Psychologica Belgica, 47*, 235–260.

28. DeRoche, K., & Welsh, M. (2008). Twenty-five years of research on neurocognitive outcomes in early-treated phenylketonuria: intelligence and executive function. *Developmental Neuropsychology, 33,* 474–504.

29. Eisenberg, N., & Eggum, N. D. (2009). Empathic responding: Sympathy and personal distress. In J. Decety & W. Ickes (Eds.), *The social neuroscience of empathy* (pp. 71–83). Cambridge, MA: MIT Press.

30. Evans, S., & Davis, M. H. (2015). Hierarchical organization of auditory and motor representations in speech perception: evidence from searchlight similarity analysis. *Cerebral cortex*, *25*, 4772-4788.

31. Farivar, S.S., Cunningham, W.E., & Hays, R.D. (2007). Correlated physical and mental health summary scores for the SF-36 and SF-12 Health Survey, V.I. *Health Qual Life Outcomes*, *5*, 54.

32. Froming, K. B., Levy, C. M., Schaffer, S. G., & Ekman, P. (2006). *The comprehensive affect testing system.* Gainesville, FL: Psychology Software, Inc.

33. Griffiths, P., Paterson, L., & Harvie, A. (1995). Neuropsychological effects of subsequent exposure to phenylalanine in adolescents and young adults with early-treated phenylketonuria. *Journal of Intellectual Disability Research, 39*, 365–372.

34. Hays, R. D., Sherbourne, C. D., & Mazel, R. M. (1993). The rand 36- item health survey 1.0. *Health economics*, 2, 217-227.

35. Hawk, S.T., Keijsers, L., Branje, S.J., Graaff, J.V., Wied, M.d., & Meeus, W. (2013). Examining the Interpersonal Reactivity Index (IRI) among early and late adolescents and their mothers. *Journal of Personality Assessment, 95*, 96-106.

36. Hood, A., Grange, D.K., Christ, S.E., Steiner, R., & White, D.A. (2014). Variability in phenylalanine control predicts IQ and executive abilities in children with phenylketonuria. *Molecular Genetics and Metabolism, 111*, 445-451.

37. Hofman D.L., Champ C.L., Lawton C.L., Henderson M., & Dye L. (2018). A systematic review of cognitive functioning in early treated adults with phenylketonuria. Orphanet Journal of Rare Disease, 13, 150. Doi: 10.1186/s13023-018-0893-4

38. Jahja, R., Huijbregts, S.C., de Sonneville, L.M., van der Meere, J.J., Bosch, A.M., Hollak, C.E. … van Spronsen, F.J. (2013). Mental health and social functioning in early treated Phenylketonuria: the PKU-COBESO study. *Molecular Genetics and Metabolism, 110 Suppl,* S57-61.

39. Jahja, R., Huijbregts, S. C., de Sonneville, L. M., van der Meere, J. J., Legemaat, A. M., Bosch, A. M., ... & de Vries, M. C. (2017). Cognitive profile and mental health in adult phenylketonuria: A PKU-COBESO study. *Neuropsychology, 31*(4), 437.

40. Jahja, R., van Spronsen, F. J., de Sonneville, L. M., van der Meere, J. J., Bosch, A. M., Hollak, C. E., ... & Janssen, M. C. (2016). Social-cognitive functioning and social skills in patients with early treated phenylketonuria: a PKU-COBESO study. *Journal of inherited metabolic disease*, *39,* 355-362.

41. Jahja, R., van Spronsen, F.J., De Sonneville, L.M., van der Meere, J.J., Bosch, A.M., Hollak, C.E., Rubio-Gozalbo, M.E., Brouwers, M.C., Hofstede, F.C., de Vries, M.C., & Janssen, M.C. (2017b) Long-term follow-up of cognition and mental health in adult phenylketonuria: a PKU-COBESO study. *Behavior Genetics*, 47, 486-497.

42. Janzen, D., & Nguyen, M. (2010). Beyond executive function: non-executive cognitive abilities in individuals with PKU. *Molecular Genetics and Metabolism, 99 Suppl* , S47–51.

43. Jensen, A. R. (2001). Vocabulary and general intelligence. *Behavioral and Brain Science,* 24, 1109–1110.

44. Koch, R., Burton, B., Hoganson, G., Peterson, R., Rhead, W., Rouse, B., Scott, R., … Azen, C. (2002). Phenylketonuria in adulthood: a collaborative study. *Journal of inherited metabolic disease*, *25*, 333–346.

45. Kohler, C.G., Turner, T.H., Gur, R.E., & Gur, R.C. (2004). Recognition of facial emotions in neuropsychiatric disorders. *CNS Spectrums, 9*, 267-274.

46. Kongs, S. K., Thompson, L. L., Iverson, G. L., & Heaton, R. K. (2000). *WCST-64: Wisconsin Card Sorting Test–64 card version, professional manual.* Odessa, FL: Psychological Assessment Resources.

47. Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in systems neuroscience, 2*, 4.

48. Kulkarni, S.K., & Dhir, A. (2009). Current investigational drugs for major depression. *Expert opinion on investigational drugs*, *18*, 767–788.
49. Landolt, M.A., Nuoffer, J.M., Steinmann, B., & Superti-Furga, A. (2002). Quality of life and psychologic adjustment in children and adolescents with early treated phenylketonuria can be normal. *Journal of Pediatrics*, *140*, 516–521.
50. Latzman, R. D., & Markon, K. E. (2010). The factor structure and agerelated factorial invariance of the Delis–Kaplan Executive Function System (D-KEFS). *Assessment*, *17,* 172–184.
51. Leggio, G.M., Salomone, S., Bucolo, C., Platania, C., Micale, V., Caraci, F., & Drago, F. (2013). Dopamine D(3) receptor as a new pharmacological target for the treatment of depression. *European Journal of Pharmacology, 719*, 25-33.
52. Long, X. Y., Zuo, X. N., Kiviniemi, V., Yang, Y., Zou, Q. H., Zhu, C. Z., ... & Li, K. C. (2008). Default mode network as revealed with multiple methods for resting-state functional MRI analysis. *Journal of neuroscience methods*, *171*, 349-355.
53. Manti, F., Nardecchia, F., Chiarotti, F., Carducci, C., Carducci, C., & Leuzzi, V. (2016). Psychiatric disorders in adolescent and young adult patients with phenylketonuria. *Molecular genetics and metabolism*, *117*, 12-18.
54. Marazziti, D., Consoli, G., Picchetti, M., Carlini, M., & Faravelli, L. (2010). Cognitive impairment in major depression. *European Journal of Pharmacology*, 626, 83-86.
55. McCrimmon, A. W., & Smith, A. D. (2013). Review of the Wechsler abbreviated scale of intelligence, (WASI-II). *Journal of Psychoeducational Assessment*, 31, 337-341.
56. McKean, C. (1972). The effects of high phenylalanine concentrations on serotonin and catecholamine metabolism in the human brain. *Brain Research*, *47*, 469–476.

57. Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49–100.

58. Moyle, J. J., Fox, A. M., Arthur, M., Bynevelt, M., & Burnett, J. R. (2007). Meta-analysis of neuropsychological symptoms of adolescents and adults with PKU. *Neuropsychology Review, 17,* 91–101.

59. Moyle, J. J., Fox, A. M., Bynevelt, M., Arthur, M., & Burnett, J. R. (2007). A neuropsychological profile of off-diet adults with phenylketonuria. *Journal of Clinical and Experimental Neuropsychology, 29,* 436–441.

60. Nardecchia, F., Manti, F., Chiarotti, F., Carducci, Cl., Carducci, Ca., & Leuzzi, V. (2015). Neurocognitive and neuroimaging outcome of early treated young adult PKU patients: A longitudinal study. *Molecular Genetics and Metabolism, 115,* 84–90.

61. Palermo, L., Geberhiwot, T., MacDonald, A., Limback, E., Hall, S. K., & Romani, C. (2017). Cognitive outcomes in early-treated adults with phenylketonuria (PKU): A comprehensive picture across domains. *Neuropsychology*, *31*, 255-267.

62. Pascucci, T., Rossi, L., Colamartino, M., Gabucci, C., Carducci, C., Valzania, A., Sasso, V., Bigini, N., Pierigè, F., Viscomi, M.T., & Ventura, R. (2018). A new therapy prevents intellectual disability in mouse with phenylketonuria. *Molecular Genetics and Metabolism*, 124, 39-49.

63. Pietz, J., Fätkenheuer, B., Burgard, P., Armbruster, M., Esser, G., & Schmidt, H. (1997). Psychiatric disorders in adult patients with early-treated phenylketonuria. *Pediatrics, 99,* 345-350.

64. Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P. (1992). Numerical recipes in fortran77: the art of scientific computing. 2nd ed. Cambridge: Cambridge University Press.

65. Preston, S. D. (2013). The origins of altruism in offspring care. *Psychological Bulletin, 139,* 1305–1341.

66. Puglisi-Allegra, S., Cabib, S., Pascucci, T., Ventura, R., Cali, F., & Romano V. (2000) Dramatic brain aminergic deficit in a genetic mouse model of phenylketonuria. *Neuroreport*, *11*, 1361-1364.

67. Reitan, R. M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor Skills*, *8*, 271-276.

68. Rey, A. (1964). *L'examen clinique en psychologie* [Clinical examination in psychology]. Paris: Presses Universitaires de France.

69. Ris, M.D., Weber, A.M., Hunt, M.M., Berry, H.K., Williams, S.E., & Leslie, N. (1997). Adult psychosocial outcome in early-treated phenylketonuria. *Journal of inherited metabolic disease, 20,* 499-508.

70. Romani, C., Manti, F., Nardecchia, F., Valentini, F., Fallarino, F., Carducci, C. , De Leo, S., MacDonald, A., Palermo, L., & Leuzzi, V. (in press). Adult cognitive outcomes in phenylketonuria: Explaining causes of variability beyond average Phe levels. *Orphanet*

71. Romani, C., MacDonald, A., De Felice, S., & Palermo, L. (2018). Speed of processing and executive functions in adults with phenylketonuria: Quick in finding the word, but not the ladybird. *Cognitive neuropsychology*. *35*, 222-226.

72. Romani, C., Palermo, L., MacDonald, A., Limback, E., Hall, S. K., & Geberhiwot, T. (2017). The impact of phenylalanine levels on cognitive outcomes in adults with

phenylketonuria: Effects across tasks and developmental stages. *Neuropsychology*, *31*, 242-254.

73. Romani, C., Tsouknida, E., & Olson, A. (2015). Encoding order and developmental dyslexia: A family of skills predicting different orthographic components. *Quarterly Journal of Experimental Psychology, 68*, 99–128.

74. Rosen, W. G. (1980). Verbal fluency in aging and dementia. *Journal of Clinical Neuropsychology, 2,* 135–146.

75. Sahakian, B. J., Jones, G. M. M., Levy, R., Gray, J. A., & Warburton, D. M. (1989). The effects of nicotine on attention, information processing, and short-term memory in patients with dementia of the Alzheimer type. *British Journal of Psychiatry, 154,* 797-800.

76. Sahakian, B. J., Morris, R.G., Evenden, J.L., Heald, A., Levy, R., Philpot, M., & Robbins, T.W. (1988). A Comparative Study of Visuospatial Memory and Learning in Alzheimer-Type Dementia and Parkinson's Disease. *Brain, 111*, 695–718.

77. Schmidt, E., Rupp, A., Burgard, P., Pietz, J., Weglage, J., & de Sonneville, L. (1994). Sustained attention in adult phenylketonuria: the influence of the concurrent phenylalanine-blood-level. *Journal of Clinical and Experimental Neuropsychology*, *16*, 681-688.

78. Schmidt, M. (1996). *Rey auditory verbal learning test: A handbook*. Los Angeles, CA: Western Psychological Services.

79. Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London*, *298*, 199–209.

80. Skuse, D. H., & Gallagher, L. (2009). Dopaminergic-neuropeptide interactions in the social brain. *Trends in Cognitive Sciences, 13*, 27–35.

81. Simon, E., Schwarz, M., Roos, J., Dragano, N., Geraedts, M., Siegrist, J., Kamp, G., & Wende, U. (2008). Evaluation of quality of life and description of the sociodemographic state in adolescent and young adult patients with phenylketonuria (PKU). *Health and Quality of Life Outcomes, 6*, 25.
82. Smyser, C. D., Snyder, A. Z., Shimony, J. S., Mitra, A., Inder, T. E., & Neil, J. J. (2014). Resting-state network complexity and magnitude are reduced in prematurely born infants. *Cerebral cortex, 26*, 322-333.
83. Smith, M.L., Klim, P., Mallozzi, E., & Hanley, W.B. (1996). A test of frontal-specificity hypothesis in the cognitive performance of adults with phenylketonuria. *Developmental Neuropsychology, 12*, 327-341.
84. Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662.
85. Sullivan, J.E. (2001). Emotional outcome of adolescents and young adults with early and continuously treated phenylketonuria. *Journal of Pediatric Psychology, 26,* 477–484.
86. Ten Hoedt, A. E., de Sonneville, L. M. J., Francois, B., ter Horst, N. M., Janssen, M. C. H., Rubio-Gozalbo, M. E., … Bosch, A. M. (2011). High phenylalanine levels directly affect mood and sustained attention in adults with phenylketonuria: a randomised, double-blind, placebo-controlled, crossover trial. *Journal of Inherited Metabolic Disease, 34*, 165–171.
87. Thomas, J., Levy, H., Amato, S., Vockley, J., Zori, R., Dimmock, D., ... & Merilainen, M. (2018). Pegvaliase for the treatment of phenylketonuria: Results of a long-term phase 3 clinical trial program (PRISM). *Molecular genetics and metabolism*.
88. Trites, R. (1977). *Grooved Pegboard Test*. Lafayette, IN: Lafayette Instrument.

89. Uzefovsky, F., Shalev, I., Israel, S., Edelman, S., Raz, Y., Perach-Barzilay, N., … & Ebstein, R.P. (2014). The dopamine D4 receptor gene shows a gender-sensitive association with cognitive empathy: evidence from two independent samples. *Emotion, 14*, 712-21.

90. Vilaseca, M.A., Lambruschini, N., Gómez-López, L., Gutiérrez, A., Fusté, E., Gassió, R., Artuch, R., & Campistol, J. (2010). Quality of dietary control in phenylketonuric patients and its relationship with general intelligence. *Nutricion Hospitalaria*, *25*, 60–66.

91. Ware, J.E., Kosinski, M., & Keller, S.K. (1994). *SF-36® Physical and Mental Health Summary Scales: A User's Manual*. Boston, MA: The Health Institute.

92. Ware, J.E., Snow, K.K., Kosinski, M., & Gandek, B. (1993). *SF-36® Health Survey Manual and Interpretation Guide.* Boston: New England Medical Center, The Health Institute.

93. Wechsler, D. (1997). *Wechsler Adult Intelligence Scale—Third Edition*. San Antonio, TX: The Psychological Corporation.

94. Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI).* San Antonio, TX: Harcourt Assessment.

95. Weglage, J., Fromm, J., van Teeffelen-Heithoff, a, Möller, H. E., Koletzko, B., Marquardt, T., … Feldmann, R. (2013). Neurocognitive functioning in adults with phenylketonuria: results of a long term study. *Molecular Genetics and Metabolism, 110,* S44–8.

96. Weglage, J., Fünders, B., Ullrich, K., Rupp, A., & Schmidt, E. (1996) Psychosocial aspects in phenylketonuria. *European Journal of Pediatrics, 155,* 101-104.

97. Weglage, J., Grenzebach, M., Pietsch, M., Feldmann, R., Linnenbank, R., Denecke, J., & Koch, H.G. (2000). Behavioural and emotional problems in early-treated adolescents

with phenylketonuria in comparison with diabetic patients and healthy controls, *Journal of inherited metabolic disease, 23,* 487–496.

98. Wise, R., A. (2004). Dopamine, learning and motivation. *Nature Reviews Neuroscience 5*, 483-494.

99. Wood, A.C., Friedmann, C.J., & Steisel, I.M. (1967). Psychological factors in phenylketonuria*, American Journal of Orthopsychiatry, 37,* 671–679.

100. Zschocke, J. (2003). Phenylketonuria mutations in Europe. *Human Mutation, 21,* 345-356.

Table 1. Demographic variables for control and PKU participants and metabolic variables for PKU participants

	Controls			PKU Group			PKU vs Controls
	Mean	SD	*n*	Mean	SD	*n*	*t*-test and *p*
Age	**28.0** (range: 18-46)	*8.5*	40	**27.4** (range: 18-47)	*8.3*	36	$t_{(1,74)}$= 0.3; *p*= .78
Education years	**14.7** (range: 9-19)	*2.0*	40	**14.3** (range: 11-18)	*2.1*	36	$t_{(1,74)}$= 0.9; *p*=.37
Gender (M/F)	**10/30**		40	**13/23**		36	$\chi^2(1)$= 1.1; *p*=.29
% Unemployed	**10.0**		30	**14.7**		34	$\chi^2(1)$= 0.3; *p*=.57
Salary	**£28822**	*11662*	17	**£23845**	*12025*	26	$t_{(1,41)}$= 1.3; *p*=.19
0-10 Years old							
N of blood samples				**189**	*67*		
Phe Level (per year)				**436**	*192*	29	
Phe Fluctuations (per year)				**189**	*102*	29	
11-16 Years old							
N of blood samples				**74**	*69*		
Phe Level (per year)				**670**	*315*	33	
Phe Fluctuations (per year)				**146**	*53*	33	
17 years old to present							
N of blood samples				**60**	*60*		
Phe Level (per year)				**747**	*304*	36	
Phe Fluctuations (per year)				**135**	*53*	36	
Phe Level at Time of Testing				**670**	*319*	34	

Note: Phe levels and Phe fluctuations across an age band were calculated by averaging year averages for each patient and then averaging values across the group. The SDs refer to differences between the participants' averages. Fluctuations were calculated in terms of SD of values within PKU participants. Blood Phe measured in μmol/L.

Table 2. Cognitive performance in control and PKU participants. Performance for different composite scores is reported in terms of *z*-scores (the higher *z*-score the worse the performance). * indicates significant differences also after Bonferroni correction (.05/14 = .004). g= Hedges' g for effect size.

	Control			**PKU Group**			**PKU vs. Controls**
	Mean	SD	*n*	Mean	SD	*n*	*t*-test and *p*
FSIQ	**113.4**	*10.1*	40	**104.3**	*14.2*	36	$t_{(1,74)}$= 3.2; ***p*=.002***; *g*=0.71
Complex EF	**0**	*0.6*	30	**0.8**	*1.2*	31	$t_{(1,59)}$= -3.4; ***p*=.001***; *g*=-0.83
Visuo-spatial attention -speed	**0**	*0.7*	30	**0.8**	*1.3*	31	$t_{(1,59)}$= -3.1; ***p*=.003***; *g*=-0.75
Language – speed	**0**	*0.8*	30	**1.1**	*1.5*	29	$t_{(1,57)}$= -3.4; ***p*=.001***; *g*=-0.91
Visuo-motor coordination	**0**	*0.8*	30	**1.0**	*1.5*	31	$t_{(1,59)}$= -3.2; ***p*=.002***; *g*=-0.82
STM	**0**	*0.7*	30	**0.5**	*0.8*	31	$t_{(1,59)}$= -2.7; ***p*=.009**; *g*=-0.66
Sustained attention	**0**	*1.0*	30	0.4	*1*	31	$t_{(1,59)}$= -1.4; *p*=*.17*; *g*=-0.39
Inhibitory control	**0**	*0.4*	30	0.3	*0.7*	30	$t_{(1,58)}$= -1.6; *p*=.11; *g*=-0.52
Visuo-spatial attention -accuracy	**0**	*0.6*	30	0.0	*0.6*	31	$t_{(1,59)}$= -0.1; *p*=.95; *g*= 0
Orthographic process. – accuracy	**0**	*0.7*	30	0.1	*1.1*	30	$t_{(1,58)}$= -0.5; *p*=.59; *g*= -0.11
Spoken language – accuracy	**0**	*0.7*	30	**0.5**	*1.1*	31	$t_{(1,59)}$= -2.2; ***p*=.03**; *g*= -0.53
Verbal Learning and memory	**0**	*0.8*	30	0.0	*0.8*	30	$t_{(1,58)}$= 0.1; *p*=.91; *g*= 0
Visuo-spatial learning and memory	**0**	0.7	30	0.5	*1.4*	31	$t_{(1,59)}$= -1.6; *p*= .11; *g*= -0.44
Overall Cognitive Score	**0**	*0.4*	30	**0.5**	*0.7*	31	$t_{(1,59)}$= -3.4; ***p*=.001*;** *g*= -0.86

Note. Complex EF: *The Tower of Hanoi puzzle; The Wisconsin Card Sorting Test; Semantic Fluency; Trail Making Test B-A*; Visuo-spatial attention - speed: *Simple Detection; Detection with distractors; Opposite Detection with distractor; Choice Reaction Time; Feature Search;* Language production- speed: *Word Reading; Nonword Reading; Picture Naming; Colour Naming.* Visuo-motor coordination: *Digit Symbol; Grooved Pegboard Test*; STM: *Digit Span; Nonword Repetition; Corsi Block Span*; Sustained attention: *Rapid Visual Information Processing* (RVP); Inhibitory control: *Stroop interference; Picture naming interference*; Visuo-spatial attention - accuracy: same task but in terms of accuracy (simple Detection excluded because of lack of errors; Orthographic processing – accuracy: *Word reading and Nonword reading; Word Spelling and Nonword spelling; Phoneme deletion; Spoonerisms;* Spoken language - accuracy*: Picture naming; Color naming; Similarities; Vocabulary*; Visuo-spatial learning and memory: *Delayed Matching to Sample*; *Paired Associates Visual Learning;* Verbal learning and memory: *The Rey Auditory Verbal Learning Test*

Table 3. Emotion recognition, emotional health/ mental health-related quality of life, and physical health-related quality of life in control and PKU participants.

	Control			PKU Group				PKU vs. Controls
	Mean	SD	*n*	Mean	SD	Average *z*-score	*n*	*t*-test, *p and Hedges's g*
EMOTIONAL HEALTH								
BDI-II – Depression	**7.5**	*8.7*	39	**7.5**	*7.2*	0.0	35	$t_{(1,72)}$= .04; *p*=. 97; *g*= 0
IRI – Empathy								
Perspective-Taking	**18.4**	*4.0*	37	**18.7**	*4.0*	- 0.1	30	$t_{(1,65)}$= -0.3; *p*=.78; *g*= -0.07
Fantasy	**15.1**	*6.3*	37	**13.9**	*5.7*	0.2	30	$t_{(1,65)}$= 0.8; *p*=.42; g= 0.19
Empathic Concern	**19.4**	*3.5*	37	**20.3**	*3.4*	-0.3	30	$t_{(1,65)}$= -1.1; *p*=.26; *g*=- 0.26
Personal Distress	**11.3**	*4.6*	37	**10.1**	*5.1*	-0.3	30	$t_{(1,65)}$= 1.0; *p*=.32; *g*= 0.24
SF-36 QoL- Mental Health								
Vitality	**53.8**	*16.5*	30	**60.0**	*20.6*	-0.4	31	$t_{(1,59)}$= 1.3; *p*=.19; *g*= -0.33
Social functioning	**90.1**	*18.9*	30	**88.7**	*14.9*	0.1	31	$t_{(1,59)}$= -.3; *p*=.75; *g*= -0.08
Limitations-emotional health	**80.0**	*35.7*	30	**79.6**	*35.1*	0.0	31	$t_{(1,59)}$= -.05; *p*=.96; *g*= 0.01
Emotional well-being	**71.3**	*15.8*	30	**75.2**	*15.5*	0.2	31	$t_{(1,59)}$= 1.0; *p*=.33; *g*= -0.25
EMOTION RECOGNITION								
Identity Test – Speed	**2.2**	*0.7*	25	**2.8**	*1.1*	**0.9**	20	**$t_{(1,43)}$= -2.3; *p*=.024;** *g*= -0.66
Identity Test – Accuracy	**21.4**	*1.0*	25	**21.5**	*0.7*	-0.1	20	$t_{(1,43)}$= -.39; *p*=.70; *g*= -0.11
3 Faces Test – Speed	**4.5**	*1.9*	25	**5.5**	*2.0*	0.5	20	$t_{(1,43)}$= -1.7; *p*=.09; *g*= -0.51
3 Faces Test – Accuracy	**22.6**	*4.8*	25	**23.1**	*2.9*	-0.1	20	$t_{(1,43)}$= -0.45; *p*=.66; *g*= -0.12
Affect Selection - Speed	**4.7**	*1.3*	25	**5.6**	*1.2*	**0.7**	20	**$t_{(1,43)}$= -2.5; *p*=.018**; *g*= -0.70
Affect Selection – Accuracy	**18.2**	*2.0*	25	**18.5**	*2.0*	-0.1	20	$t_{(1,43)}$= -0.44; *p*=.66; *g*= -0.15
• *z* emotion recognition Speed	**0.0**	*0.9*	25	*1.0*	**0.7**		20	**$t_{(1,43)}$= -2.5; *p*=.017;** *g*= -0.73
• *z* emotion recognition Accuracy	**0.0**	*0.7*	25	*0.6*	-0.1		20	$t_{(1,43)}$= 0.6; *p*= .55; *g*= 0.15
PHYSICAL HEALTH								
SF-36 QoL- Physical Health								
Physical function	**96.2**	*8.8*	30	**93.9**	*17.6*	0.3	31	$t_{(1,59)}$= 0.6; *p*= .53; *g*= 0.16
Limitations - physical health	**94.2**	*20.4*	30	**91.1**	*26.3*	0.1	31	$t_{(1,59)}$= 0.5; *p*=.62; *g*=0.13
Bodily pain	**85.3**	*17.8*	30	**87.2**	*16.3*	-0.1	31	$t_{(1,59)}$= 0.4; *p*=.66; *g*= -0.11
General health	**71.7**	*18.6*	30	**76.3**	*19.6*	-0.2	31	$t_{(1,59)}$= 0.9; *p*=.35; *g*= -0.24

Note. For emotion recognition, speed was measured in seconds; a composite *z*-score averages speed and accuracy *z*-scores in the 3 Faces Test and the Affect Selection Test. For consistency across tasks, a higher *z*-score always reflect worse performance. No differences are significant after Bonferroni correction (0.05/22= 0.002). g= Hedges' g for effect size. QoL=quality of life.

.

Table 4. Correlations between cognitive measures on the one hand (IQ, composite indexes and emotion recognition), and emotional health/ mental-health-related quality of life and physical health-related quality of life on the other.

	EMOTIONAL/MENTAL HEALT									**PHYSICAL HEALTH**			
	BDI-II Depression	**IRI Empathy**				**SF36 QoL Mental Health**				***SF36 QoL***			
		Perspect. Taking	**Fantasy**	**Empathic Concern**	**Distress**	**Vitality**	**Social Function**	**Em. Limits**	**Em. well-being**	***Physical Function***	***Physical Limits***	***Bodily Pain***	***General health***
FSIQ	.10	.03	.08	-.13	.17	.22	.02	-.15	-.08	**.60****	**.58****	**.50****	**.46****
N	*35*	*30*	*30*	*30*	*30*	*31*	*31*	*31*	*31*	*31*	*31*	*31*	*31*
Composite cognitive scores													
Complex EF	-.29	-.13	.11	-.04	-.11	.09	-.24	**-.44***	-.11	**.37***	.28	**.37***	.02
Visuo-spatial attention -speed	-.14	-.06	.25	-.02	-.16	.15	-.15	-.13	-.28	**.63****	**.69****	**.40***	.14
Language -speed	.20	**.51****	**.51****	.21	.24	.17	.05	.06	-.01	.15	**.43***	**.38***	.31
Visuo-motor coordination	-.10	.06	**.47***	.24	-.20	.18	-.20	-.21	-.16	**.64****	**.58****	.33	.13
STM	.20	.21	**.48***	.05	**.54****	.28	.02	-.06	-.05	**.53****	**.54****	.28	**.46***
Sustained attention	.01	.02	.08	.13	-.17	.16	.07	-.01	.15	.32	**.36***	**.37***	.25
Inhibitory control	-.09	-.10	.36	-.11	.10	-.14	**-.39***	-.23	-.29	.22	.26	.22	-.07
Visuo-spatial atten. -accuracy	.04	.04	-.21	-.02	-.20	.18	.02	.20	**.37***	.25	.21	.09	.13
Orthographic process. - accuracy	.14	**.40***	**.41***	.17	.32	.09	.02	.21	.08	.21	**.38***	.31	.32
Spoken language	-.07	-.12	-.10	-.07	.06	.16	-.32	-.29	-.20	**.74****	**.67****	**.40***	.19
Verbal learning and memory	-.08	.06	.09	.07	**-.44***	.28	-.03	-.14	.05	**.38***	.29	**.40***	.21
Visuo-spatial learn. and memory	-.20	**-.42***	.02	-.05	**-.47***	.16	-.25	-.24	-.10	**.57****	**.45***	**.54****	.08
$n^{\#}$	*31*	*26*	*26*	*26*	*26*	*30*	*30*	*30*	*30*	*30*	*30*	*30*	*30*
Emotion Recognition													
Speed	.23	.28	.48	-.09	.48	.09	.13	.39	-.13				
Accuracy	-.07	.26	.05	-.08	.11	.20	-.16	-.09	-.19				

n	*19*	*16*	*16*	*16*	*16*	*15*	*15*	*15*	*15*

Note: For consistency, across tasks, positive correlations always indicate that better cognitive skills were linked to better health. For example, negative correlations with verbal memory and learning indicate that those with better memory did worse on some heath measures; positive correlation between FSIQ and BDI-II indicates that better the IQ better the mental health on the depression dimension (i.e., lower the depression).

Significant correlations are in bold. ***p*<.01 **p*<.05.

Table 5. Correlations between cognitive measures (IQ, cognitive composite scores, emotion recognition) on the one hand and Phe measures (average levels and fluctuations) on the other.

	Average Phe levels				**Phe fluctuations-SD**		
	0-10 yr	**11-16 yr**	**17 yr-now**	**Con-current**	**0-10 yr**	**11-16 yr**	**17 yr-now**
FSIQ	.26	.13	.24	.24	**.41***	**.46****	**.43****
Cognitive composite scores							
Complex EF	.36	.32	.27	.28	**.51****	**.42***	.21
Visuo-spatial attention – speed	**.53****	.32	.33	.07	**.45***	.16	**.36***
Language – speed	.11	.04	.20	-.02	.05	-.01	.01
Visuo-motor coordination	.34	.23	.34	**.41***	**.42***	**.43***	**.39***
STM	.10	-.05	-.05	-.23	.10	.21	.23
Sustained attention	.17	.04	.15	.32	.08	.01	.26
Inhibitory control	.18	-.02	-.04	.08	.17	.34	.18
Vsuo-spatial attention - accuracy	.21	.16	.34	.34	**.42***	**.51****	**.36***
Orthographic processing - accuracy	-.01	-.18	.01	-.02	.03	.14	.16
Spoken language	.33	.08	.11	.16	**.44***	**.53****	**.56****
Verbal learning and memory	.23	**.38***	**.45***	**.63****	**.44***	**.40***	.19
Visuo-spatial learning and memory	**.51****	**.58****	**.40***	**.38***	**.62****	**.45***	**.54****
Emotion Recognition							
Speed	-.05	-.08	.32	-.03	.02	-.18	.23
Accuracy	-.14	-.30	-.11	-.15	.16	.28	.15

Note. To facilitate interpretation, positive correlations indicate that high Phe was related to worse performance. Thus, in the case of the IQ, as Phe increased, IQ decreased.
Note the number of AwPKU is slightly different across tests (see table 2), as is the number of Phe measures available (see table 1); this means that the same value of *r* may have different probabilities. Significant correlations are in bold. Significant correlations are in bold. **p<.01 *p<.05.

Table 6. Correlations between emotional health/ mental-health-related quality of life and physical health-related quality of life measures on the one hand, and Phe measures (average levels and fluctuations) on the other.

	Average Phe levels				**Phe fluctuations-SD**		
	0-10 yr	**11-16 yr**	**17 yr-now**	**Con-current**	**0-10 yr**	**11-16 yr**	**17 yr-now**
EMOTIONAL/MENTAL HEALTH							
BDI-II Depression	-.16	-.07	.12	.02	-.30	-.21	.28
IRI Empathy							
Perspective-Taking	**-.49***	**-.54****	-.10	-.07	-.24	.09	-.30
Fantasy	-.26	-.05	.13	-.04	-.15	.04	.32
Empathic Concern	-.24	-.16	0	.17	-.01	-.11	-.23
Personal Distress	-.32	**-.52****	-.29	**-.51****	-.34	-.06	.02
SF-36 QoL Emotional Health							
Vitality	.10	.24	.21	-.11	.27	.28	**.39***
Social functioning	-.18	-.04	.06	.01	-.26	-.17	.20
Emotional limitations	-.13	-.05	.11	0	-.24	-.16	.26
Emotional well-being	-.03	.14	.16	.13	-.13	-.07	.05
PHYSICAL HEALTH							
SF-36 QoL Physical Health							
Physical functioning	**.48***	.20	**.37***	.03	**.57****	**.51****	**.71****
Physical limitations	**.55****	.21	**.42***	-.04	**.55****	**.39***	**.62****
Bodily pain	.36	.36	**.40***	.14	.38	.17	**.50****
General health	-.23	-.06	.30	.14	-.17	.07	**.49****

Note. To facilitate interpretation, positive correlations indicate that high Phe was related to reduced mental/physical health. Thus, in the case of the BDI, perspective taking and personal distress, negative correlations indicate that as Phe increased depression and personal distress decreased, while perspective-taking increased. Significant correlations are in bold. ***p*<.01 **p*<.05.

Table 7. Correlations among metal health (BDI and IRI), health-related quality of life (SF-36) and diet.

	Importance of Diet	**Ease to Keep on Diet**
EMOTIONAL HEALTH		
BDI-score	-.13	**-.39***
n	*29*	*29*
IRI-score		
• Perspective-Taking	.10	.13
• Fantasy	-.01	.19
• Empathic Concern	.29	.12
• Personal Distress	.21	.04
n	*24*	*24*
SF-36-score – Mental health		
• Vitality	-.06	.21
• Social functioning	-.06	.34
• Limitations emotional health	.09	.28
• Emotional well-being	.27	**.51****
n	*28*	*28*
PHYSICAL HEALTH		
SF-36-score – Physical health		
• Physical functioning	.08	-.02
• Limitations physical health	.18	-.03
• Bodily pain	.25	.10
• General health	-.08	.27
n	*28*	*28*

Note. ***p*<.01 **p*<.05. Highlighted correlations are significant at *p*<.01.

.