Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

Meta-analyses of cognitive functions in early-treated adults with phenylketonuria

Cristina Romani [a,*], Andrew Olson [b], Lynne Aitkenhead [c], Lucy Baker [a], Dhanesha Patel [a],
Francjan Van Spronsen [d], Anita MacDonald [e], Annemiek van Wegberg [d], Stephan Huijbregts [f]

[a] *Psychology Department, Aston University, UK*
[b] *Psychology Department, University of Birmingham, UK*
[c] *University College London Hospital NHS Foundation Trust, UK*
[d] *Beatrix Children's Hospital, the Netherlands*
[e] *Birmingham Women's and Children's NHS Foundation Trust, UK*
[f] *Leiden Institute for Brain and Cognition, Leiden University, UK*

A R T I C L E   I N F O

A B S T R A C T

Our study estimated size of impairment for different cognitive functions in early-treated adults with PKU (AwPKU) by combining literature results in a meta-analytic way. We analysed a large set of functions (N = 19), each probed by different measures (average = 12). Data were extracted from 26 PKU groups and matched controls, with 757 AwPKU contributing 220 measures. Effect sizes (ESs) were computed using Glass' Δ where differences in performance between clinical/PKU and control groups are standardized using the mean and standard deviation of the control groups. Significance was assessed using measures nested within independent PKU groups as a random factor. The weighted Glass' Δ was -0.44 for all functions taken together, and -0.60 for IQ, both highly significant. Separate, significant impairments were found for most functions, but with great variability (ESs from -1.02 to -0.18). The most severe impairments were in reasoning, visual-spatial attention speed, sustained attention, visuo-motor control, and flexibility. Effect sizes were larger with speed than accuracy measures, and with visuo-spatial than verbal stimuli. Results show a specific PKU profile that needs consideration when monitoring the disease.

1. Introduction

Phenylketonuria (PKU) is an autosomal recessive genetic disease with an incidence of about six cases per 100,000 live births, but with strong regional variations (e.g., Shoraka et al., 2020). It is characterized by variants in the gene encoding the enzyme phenylalanine hydroxylase (PHA) that metabolizes the essential amino acid phenylalanine (Phe) into another amino acid: tyrosine. PAH is mainly present in the liver. When it is not functioning, Phe accumulates in the blood and in the brain since Phe can cross the blood-brain barrier. How PKU impairs brain functions is not completely clear, but several different causes are likely (van Spronsen et al., 2021). High Phe levels may have a direct toxic effect on myelin, the sheet of cells which wraps around axons and increases speed of transmission (de Groot et al., 2010). In addition, no or

limited PAH activity will reduce the availability of important neurotransmitters such as dopamine and serotonin. Dopamine is synthesized from tyrosine which will be less available, both because less tyrosine is synthesized from Phe and because Phe will compete with the tyrosine which is supplied by food to cross the blood-brain barrier. Phe will also compete with other essential amino acids, such as tryptophan, from which serotonin is synthesized. Finally, elevated levels of Phe are thought to inhibit the activity of the enzymes tyrosine hydroxylase and tryptophan hydroxylase, which synthetize dopamine and serotonin respectively (e.g., Donlon et al., 2015; Gonzalez et al., 2016; see Boot et al., 2017 for a review). Through all these mechanisms, high levels of Phe seriously damage brain health and, if left untreated, PKU leads to severe disability.

Fortunately, Phe levels can be controlled by following a strict diet

---

that severely limits the ingestion of natural foods containing Phe (most protein foods), replacing them with foods where Phe is naturally low or has been artificially removed, and adding protein substitutes which guarantee adequate protein intake (see MacDonald et al., 2020 for general guidelines). If a strict PKU diet is followed from birth, the most serious consequences of high Phe levels can be avoided. Since the mid-sixties, programs to screen new-borns for PKU have been adopted by an increasing number of countries, allowing recent generations of people with PKU to reach good educational standards and lead fulfilling lives, although possibly not reaching their full potential. The treatment of PKU is one of the great achievements of modern medicine. However, outcomes remain suboptimal and cognitive impairments are well documented both in early-treated children and adults (van Spronsen, 2021).

Impairments are especially documented in IQ, speed of processing, and executive functions. Group performance is generally below average, but there is strong individual variability (for reviews see Burlina et al., 2019; Hofman, 2018; Trefz et al., 2011; van Spronsen et al., 2011; for adult results spanning a comprehensive set of functions see Brumm et al., 2004; Palermo et al., 2017). Most of these impairments and variability are attributable to poor metabolic control with many studies documenting strong correlations between blood Phe levels and performance in cognitive tasks (Jahja et al., 2017; Moyle et al., 2007a; Nardecchia et al., 2015; Ris et al., 1994; Romani et al., 2017, 2019; also for differences between groups with high vs low Phe see Bartus et al., 2018; Channon et al., 2007; Palermo et al., 2017). Treatment practices implemented in the last 30–40 years have dramatically reduced Phe levels in people with PKU. However, in most individuals current and historical Phe levels remain well above current therapeutic recommendations. European guidelines recommend maintaining Phe within 120–360 μmol/L until 12 years of age and within 120–600 μmol/L above 12 years (van Spronsen et al., 2017; van Wegberg et al., 2017). American guidelines recommend 120–360 μmol/L throughout life (American College of Medical Genetics and Genomics, ACMG, Vockley et al., 2014). Dietary recommendations have changed over the years. Keeping a strict diet was originally recommended only until six/eight years of age (see Koch et al., 2002 for examples of early discontinuation), then till adolescence (e.g., Report of Medical Research Council Working Party on Pheylketonuria, 1993) and more recently for life (e.g., van Spronsen et al., 2017). These changes, however, are only partially responsible for failures to maintain therapeutic targets. Instead, many individuals with PKU opt to abandon the diet in adolescence or even in late childhood because the PKU diet is expensive, time-consuming, socially difficult, and, despite continuous improvements, unpalatable. Thus, even when on diet, individuals with PKU show metabolic control which is generally far from perfect (e.g., see Romani et al., 2019).

Although there is accumulating evidence that cognitive outcomes in people with PKU are below what would be expected based on control groups, it is difficult to have accurate estimates of the size of the impairments. Understandably, there is variability in the outcomes reported by different studies. This is due to the limited number of participants per study, and to differences in level of metabolic control, both across studies and across participants within studies (see later demographic characteristics of the studies included in our review). Additionally, most studies have assessed only a few functions and with few tasks because of limitations in the resources of research teams and in the engagement that can be requested from participants. Meta-analyses combine results from different studies by standardizing differences between PKU and control groups (using effect sizes; ESs) and then averaging them across studies. This allows better estimates of the size of impairments, thus providing a more comprehensive and accurate picture of which functions are spared or impaired. How tasks are attributed to different functions is necessarily a matter of judgement. Tasks (especially more sensitive, complex tasks) involve multiple functions, and the decision about which 'main' function is tapped by a task is, to some degree, a matter of opinion. However, combining results across studies allows

each function to be assessed with multiple tasks overlapping in the skill of interest. This will produce a better overall estimate of the function, reducing the measurement error inherent in each task.

There are several meta-analyses of cognitive performance in children with PKU (DeRoche and Welsh, 2008; Waisbren et al., 2007). DeRoche and Welsh (2008) used meta-analyses of 33 studies to assess (Hedges' g) effect sizes for IQ and for executive functions such as flexibility, planning, inhibition and working memory. They found significant differences for all functions, with the largest difference for flexibility (ES = 1.15) and the smallest for IQ (ES = 0.42). Waisbren et al. (2007) estimated the effect of Phe levels on IQ by carrying out a meta-analysis of within-study correlations. In early-treated children, IQ correlated with life-time Phe ($r = -0.35$), Phe at 0–10 years of age ($r = -0.34$), and Phe at the time of testing ($r = -0.31$; N of studies = 12, 14, 29, respectively). An increase in Phe of 100 μmol/l predicted an average reduction in IQ of 1.3, 1.9, and 0.5 points, respectively, for the different Phe measures. Finally, Fonnesbeck et al. (2013) carried out meta-analyses to predict the probability of low IQ from a range of blood Phe levels. They pooled results from 17 unique studies which assessed individuals with PKU of different ages (children, adults or mixed). They found significant associations between the probability of IQ being < 85 and Phe level averaged in critical and non-critical periods (with a coefficient = −0.036), but no relationship with concurrent Phe.

These meta-analyses provide important results on the relationship between IQ and Phe. A recent review by Hofman et al. (2018) addressed, instead, the issue of the profile of cognitive outcomes in adults with PKU (from now on AwPKU). It summarized the results of 16 studies, investigating cognition in early-treated AwPKU and associations with Phe levels. The most consistent deficits involved *vigilance/sustained attention* (assessed by tasks such as the Continuous Performance Test; Rapid Visual Processing and Dot tasks; and impaired in 6/7 studies), *working memory* (assessed through both verbal and visuo-spatial tasks; and impaired in 12/15 studies) and *motor skills* (assessed by tasks such as Digit-symbol-coding, Grooved pegboard, Motor screening test, and pursuit tasks; and impaired in 5/6 studies). Other functions were impaired less consistently. For example, complex executive functions (assessed by tasks such as the Brixton test, Eithorn Perceptual Maze Test, Six Element Test, Stockings of Cambridge etc.) were impaired in only in 9/17 studies. This review provides important insights into the pattern of impairment present in PKU, but it reports study results in terms of presence/absence of significant impairment, without estimating size of impairment across studies. Only three meta-analyses of cognitive performance in AwPKU have been published so far. However, in a motivated way, these reviews focused on what are common impairments in PKU (mainly regarding executive functions) and did not address the breadth/spectrum of cognitive functions considered by Hofman et al. (2018).

Moyle et al. (2007a) assessed effect sizes for inhibition, attention, motor control, processing speed and working memory, combining results from five studies and 220 AwPKU. Hedges' g ESs for these functions ranged from .30 to .90 (number of measures 3–10). Working memory was not significant, and the largest effect size was obtained for tasks measuring processing speed. Bilder et al. (2016) carried out a meta-analysis of 13 studies focusing on executive functions. There were significant effects for attention (11 studies; 252 participants; ES= −0.74); inhibition (6 studies; 119 participants; ES = −0.41) and flexibility (7 studies; 157 participants; ES = −0.43). Working memory was again not significant (5 studies; 112 participants; ES = −0.08). Albrecht et al. (2009) carried out meta-analyses of studies which used timed tasks in children (N = 229), adolescents (N = 106) and adult with PKU (N = 174). Tasks assessed Trail making; Simple RT; Choice RT; Interference; N-back; Visual search; and rule identification tasks such as the Wisconsin card sorting test. Effect sizes were highly significant in all age groups (for children =−0.31; 95 % CI: −0.43, −0.19; adolescents = −0.13; 95 % CI: −0.29, −0.03; adults =−0.54; 95 % CI: −0.64, −0.44), but they were significantly moderated by the average Phe level only in
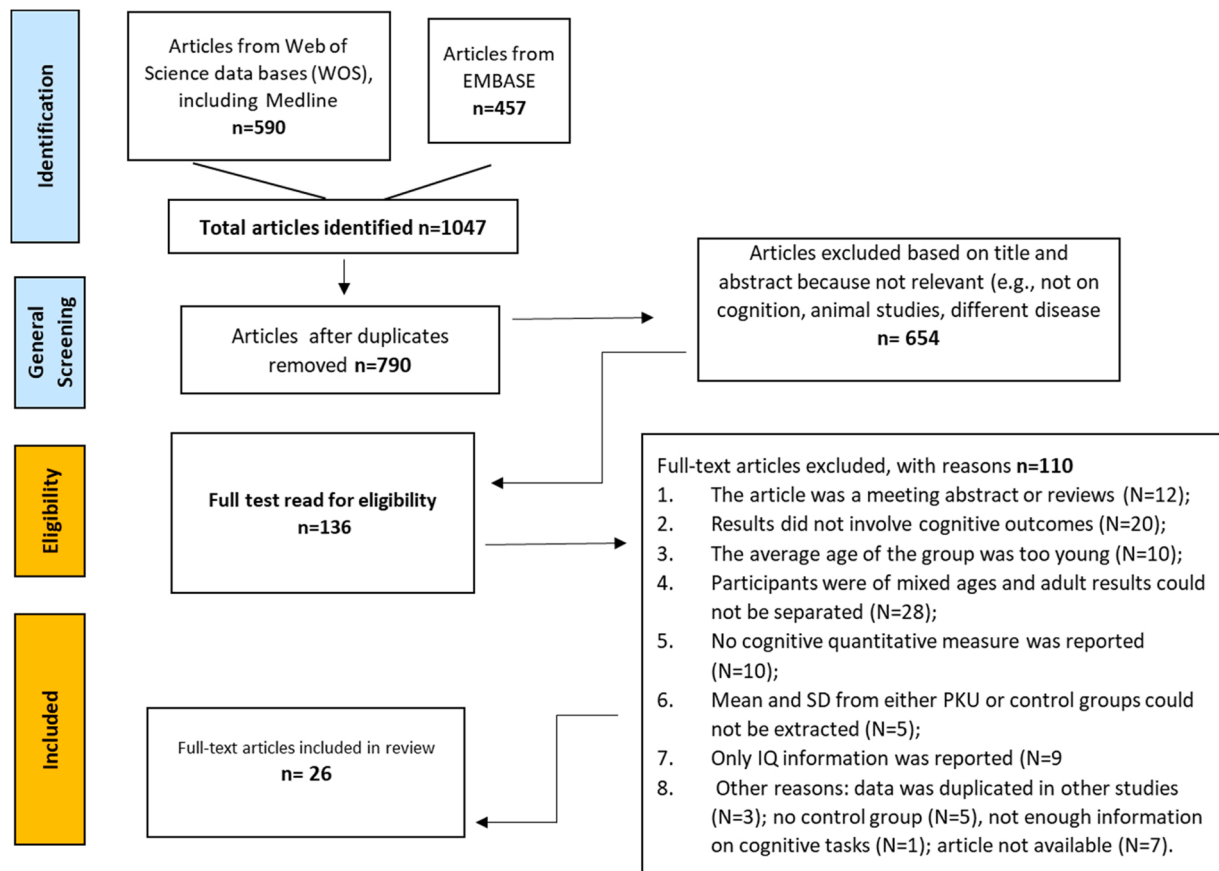
Fig. 1. Flow chart of process of article identification.

the child and adolescent groups.

The aim of the current study was to provide a comprehensive picture of the cognitive profile of early-treated AwPKU by considering a large set of functions and tasks as done by Hofman et al. (2018), but also combining the results from different studies through meta-analyses rather than merely counting numbers of significant results. In addition, and in contrast with previous meta-analyses which included mixed-age groups, we limited comparisons to cohorts of people > 16 years old, thus focusing more strictly on outcomes in adulthood. Finally, we used a slightly different measure of ES than previous meta-analyses. We used Glass' Δ as a measure of effect size which standardizes the difference between the means using only the standard deviation (SD) of what can be considered the reference group (the control group in our case). Previous meta-analyses have used Hedges' *g* ESs, where a difference between means is standardized using a combined SD from both the control and the PKU group. The Glass' Δ ES is more appropriate for clinical groups because one wants to measure the performance of a clinical group against a control group in a way conceptually analogous to a z-score. The fact that Glass' Δ is not affected by the SD of the clinical group is an added advantage, because, in the PKU groups, the SD is generally larger than in controls, since individual variation in performance is increased by differences in disease factors such as differences in metabolic control (e.g., Romani et al., 2019).

Another benefit of our review is that it will provide a snapshot of the characteristics of studies carried out since the early 1990s when the first cohorts of early-treated AwPKU reached adulthood by considering number of participants, age, current Phe levels and cognitive functions assessed. We will assess how ESs are moderated by variables such as average Phe level, age when participants started treatment, or year the study was conducted. More recent studies may include participants with better metabolic control given the adoption of more stringent targets for

upper Phe levels. Note, however, that these analyses do not consider variability *within* studies and thus will underestimate the effects of Phe on cognition. Better estimates should be based on a meta-analysis of within-study correlations (see Waisbren et al., 2007; Fonnesbeck et al., 2013), which is outside the scope of the present study. Our primary aim is to provide estimates of impairment for different cognitive functions. These estimates are important to assess the success of current treatment practices and to have a baseline measure against which to evaluate the success of future policies. Additionally, better knowledge of the typical cognitive profile in PKU will allow affected people, their families, and their clinicians to have more accurate expectations about outcomes and to take strengths and weaknesses of the disease profile into consideration when evaluating educational achievements, job performance, and remediation strategies.

2. Method

2.1. Literature search and data extraction

2.1.1. Literature search

We comprehensively searched the literature for published studies which reported cognitive performance in early treated AwPKU and control groups. We used the Web of Science database (WOS) which includes the Medline database, among others, and Embase. The Pubmed database largely overlaps with Medline and was not separately searched (see Bramer et al., 2017). Databases were searched without any limitation for year of study up to May, 2022. The first relevant study was from 1994 (around this time, results from early-treated AwPKU first became available). The following combination of terms were used:

PKU OR Phenylketonuria AND Adult* AND cognitive OR IQ OR RTs OR reaction times OR speed OR executive OR inhibition OR flexibility

OR attention OR visuo-motor OR spatial OR learning OR language OR memory OR Stroop OR Go no-go.

A flowchart detailing our identification process is shown in Fig. 1. All abstracts and all full articles were read by at least two authors and screened for the following inclusion criteria:

- Participants with PKU were adults aged > 16 years;
- Participants with PKU were early-treated, with treatment started within three months of birth, or were defined as early-treated;
- Outcomes were measured in a quantitative way through a cognitive task (not through a questionnaire);
- An age-matched control group was included;
- Both mean and SD of the PKU group and the matched control group were reported or could be recovered.

Disagreements regarding whether a study could be included in the meta-analysis only occurred in a handful of cases and were addressed and resolved by the first author. Our procedure left 26 articles/publications with potential results to include in the meta-analysis. Given our age criterion, the following studies included in previous adult reviews were excluded because they included some younger participants: from Moyle et al. (2007a): Brunner and Berry (1987); Clarke et al. (1987); Luciana et al. (2001); 2004; Stemerdink et al. (1999); Griffiths et al. (1995). from Bilder et al. (2016): Antenor-Dorsey et al., 2013; Griffiths et al. (1995); Luciana et al. (2001). A list of adult studies was not provided by Albrecht et al. (2009) to allow a check on the age of participants.

2.1.2. Data extraction

To make sure that only unique sets of data were considered, we used the following criteria:

1. When results included different subgroups, only results for the whole PKU group were included. If only results for the subgroups were available these were separately reported. Exceptionally, Burgard et al. (1997) assessed a French and a German PKU group; only the German group was included since diet was relaxed very early in the French group (after five years old) and a French control group was not included.
2. When results were reported for the same participants at different times, only one set of results was selected according to the following criteria:

    a) Nardecchia et al. (2015) measured cognitive functions over a 14-year gap. Only the measures taken at time 2 were included, since participants were adults at only this time;

    b) Weglage et al. (2013) assessed AwPKU at baseline and after five years. Only results at baseline were included, since all participants were already adults at this time and there were no differences between the two testing times.

    c) Feldmann et al. (2019) assessed a subgroup of the AwPKU assessed by Weglage et al. (2013) after a further 5 years (10 from baseline). Only the Weglage results at baseline were included.

    d) Schmidt et al. (1994) assessed AwPKU in three conditions (no diet at time of testing, diet, no diet again), only Time 1, baseline measures, were included.

    e) Sundermann et al. (2011) assessed AwPKU after receiving either a Phe loading capsule or a placebo (within-participant design). Only results for the placebo condition were included whether it was assessed after (group 1) or before Phe loading (group 2).
3. In some studies, overlapping PKU groups were assessed using different tasks. This was the case for Palermo et al., (2017, 2020) and De Felice et al. (2018), and for Jahja et al. (2016) and (2017). Results were separately included. However, for overlapping data only results from the most recent study were included.

After applying these criteria, we were able to extract results for 26

separate pairs of PKU and control groups (see Table 5 in the Appendix). For every group, averages and SDs of cognitive measures were entered in a spreadsheet. When several measures were reported for the same task, we limited input to those likely to be more independent. For example, for the peg-board task, we averaged measures from the dominant and non-dominant hand; for the Rey verbal learning test, we reported only two measures: recall over five trials and delayed recall; for the Wisconsin Card Sorting test, we only reported number of categories correct and number of perseverations. However, we always considered both speed and accuracy measures separately since they could be affected differently in people with PKU. We included measures of IQ only from studies where participants in PKU and control groups have been 'roughly' matched for IQ by excluding PKU participants with pathologically low IQ. This is a conservative choice. We will see that in spite of this PKU and control groups differed for IQ.

All raw results were extracted from the studies by two of the authors working together (LB and DP). They were then thoroughly and comprehensively checked (and rechecked) by two other authors (CR and AO), with any disagreement resolved by discussion.

2.2. Computation of effect sizes

2.2.1. Types of effect sizes

Moyle et al. (2007a), Albrecht et al. (2009) and Bilder et al. (2016) all carried out their meta-analyses based on Hedges' *g* effect sizes, which are calculated by dividing the difference between the PKU and control mean by a weighted standard deviation (SD) that pools the SDs of the two groups. Hedges' *g* is almost identical to the best-known effect size measure, Cohen's *d,* but adds a correction for small samples. We carried out the bulk of our analyses, however, using Glass' Δ, reporting Hedges' *g* only for comparison in Supplementary materials. The Glass' Δ ES differs because the difference between groups is standardized using only the SD of the reference group (the control group in our case; see Ialongo, 2016). Arguably, therefore, it is a better measure of deviation from normality for clinical samples.

These are the formulas for the three types of ESs.

$$\text{Cohen's } d = \frac{\overline{X}_E - \overline{X}_C}{\sqrt{\frac{(n_C-1)S_C^2 + (n_E-1)S_E^2}{(n_C + n_E - 2)}}}$$

$$\text{Hedges' } g = d \times \left(1 - \frac{3}{4(n_C + n_E - 2) - 1}\right)$$

[where $d = $ Cohen's d]

$$\text{Adjusted } Glass' \Delta = \frac{(\overline{X}_E - \overline{X}_C)}{S_C} \times \left(1 - \frac{3}{4(n_C-1)-1}\right)$$

where $\overline{X}_C$ and $\overline{X}_E$ are the mean values of the control group and the PKU group; S_C and S_E are the respective standard deviations (SD); and n_C and n_E are the number of participants in the two groups.

Following Pustejovsky (2016), Glass' Δ ESs were adjusted for small samples using the same adjustment that is applied to Hedges' *g*. When only t-values (rather than mean and standard deviations) were available (e.g., Aitkenhead et al., 2021) we computed Glass Δ and variance of Glass Δ values according to the formulas given by Pustejovsky (2016).

We checked for unusual effect sizes by calculating the mean and standard deviation of effect sizes for a cognitive function. We eliminated studies when the effect size was more than 2.5 standard deviations from the average effect size for that function. Six measures were excluded on this basis (=2 % of effect sizes).

2.2.2. Combining ESs and statistical analyses for individual functions

To compute individual ESs, the control mean of each measure was subtracted from PKU mean. However, to combine results across tasks and measures we wanted all differences between PKU and control groups to be in the same direction. Therefore, we reversed the difference (multiplied by −1) when a higher mean was indicative of lower performance, as when performance was reported in terms of reaction time

or errors. Thus, negative ES always indicated worse performance for the PKU group.

When combining ESs it is also important to weight effects according to some study variables. More weight must be given to estimates obtained with more participants, using more tasks, and where more independent groups were involved, (i.e., adjusting when several measures came from the same group of participants). To combine ESs by function, we used a random effects model. A fixed effects model assumes that all ESs reflect a fixed, true effect shared across PKU groups and measures. Thus, results would not generalize to different studies with different characteristics. A random effects model, instead, assumes that ESs are sampled from a population of ESs which vary depending on the degree of impairment in different PKU groups and the sensitivity of different tasks. Therefore, estimates consider both sampling error and heterogeneity in the population of ESs. Heterogeneity is smaller when results from different measures are consistent. Moyle et al. (2007a) used either a fixed or a random effects model based on a heterogeneity test. If measures were not significantly heterogeneous, a fixed effects model was used. Albrecht et al. (2009) and Bilder et al. (2016) used a random effects model for all comparisons. We also always used a random effects model as the more conservative and theoretically sound model.

To assess significance, we used mixed-model regression analyses. We used a three-level random effects model because we wanted to account, not only for sampling error and heterogeneity in the population, but also for structure in the participant groups, since some measures were taken from the same group of participants. We used *'measure' nested within 'PKU-group'* as a random factor, which clustered together measures which came from the same people. Our model, therefore, accounted for potential differences in sensitivity among measures, for potential differences in ability among groups, and for clusters of measures taken from the same people. We carried out our analyses using the Metafor package from R (see: bookdown.org/MathiasHarrer/Doing_-Meta_Analysis_in_R/mlma.html) as follows: M1 < -rma.mv(g,Vg, random = ~ 1 | GroupNum / MeasureNum, test="z", data=MyData, method = "REML"). Significance of coefficients was assessed using z values. Complete code for the analyses is available at OSF https://osf.io/awr5e. The materials are also available through github at: https://github.com/olsonac/pku_cognitive_functions_meta_analysis.

Studies with a small sample size could unduly inflate ES estimates if they are published more often when they obtain results matching the expectation that the PKU groups are impaired. We examined this possibility using a funnel plot and a sensitivity analysis. With a funnel plot analysis ESs are plotted against a measure of study precision. If precision is plotted on the vertical axis, then one expects that the ESs derived by studies with more precision will cluster closer to the mean at the top (closer to the true value of the effect) and be more widely distributed at the bottom, thus producing a plot with a funnel shape. Moreover, if a publication bias exists, one would expect that some studies at the top of the distribution will be missing on the side inconsistent with predictions because small studies with inconsistent outcomes will be less likely to be published. We will assess a potential asymmetry in the distribution of a funnel plot using an Eggar's test (see Egger et al., 1997). In addition, we will carry out a sensitivity analysis where individual ESs are ordered by precision (based on the number of participants contributing to the estimate) and, then subdivided into bins which include progressively more ESs of low precision. If there is a publication bias and studies with low precision inflate the estimate, this should increase with the addition of low precision studies (for an outline of this logic see in Borenstein et al., 2009, Chapter 30).

We considered the effects of the Phe level of the PKU samples at the time of assessment in three different ways. When we compared groups of measures (e.g., visuo-spatial vs. language tasks), we factored out Phe level since this could influence the size of the effect for a group of measures. Secondly, we ran bivariate Pearson correlations between individual effect sizes, Phe levels, age of diet initiation, and year of study to assess the relationship among these variables. Finally, we assessed the contribution of the average current Phe (as well as the contribution of age of treatment initiation and year of study) to our ES estimates by considering whether these variables were significant moderators.

2.3. Cognitive outcome measures

We computed effect sizes for 19 different functions. They can be grouped into different descriptive domains, but our focus is on functions which are theoretically coherent and which can be potentially independently impaired. We considered the following functions, each assessed through different tasks (see also Hofman et al., 2018; Lezak et al., 2012).

Executive functions:

- a) *Flexibility and planning:* Digit-symbol coding; Letter and semantic fluency; Object alternation; Trail making test B-A; WCST categories; WCST perseverations.
- b) *Reasoning:* Tower of London*;* Elithorns puzzle; Brixton rule finding test; Six Elements ordering task.
- c) *Inhibitory control:* Stroop incongruous-congruous condition; Hayling sentence completion B; Flanker task incompatible-compatible condition; Block cyclic naming semantic interference; Continuous naming semantic interference.
- d) *Sustained attention*: Dot pattern exercise average performance; Dot pattern exercise stability; D2 test; Telephone search test; Rapid visual processing of number sequences.

Language:

- a) *Higher-order language skills*: Verbal narrative: speech rate; mean utterance length; number of main concepts; number of correct information units; Right Hemisphere Battery: Prosody; Conflicting prosody; Metaphors; Humour; Inferred meaning; WASI Vocabulary; WASI Similarities.
- b) *Naming speed*: Continuous Picture Naming-RT; Blocked cycling naming-RT; word reading-RT; Stroop colour naming neutral condition-RT.
- c) *Naming accuracy*: Continuous Picture Naming-accuracy; Blocked cycling naming-accuracy; Word reading-accuracy; Nonword reading accuracy; Stroop colour naming neutral condition-accuracy; Word accuracy in narrative.
- d) *Spelling*: Spelling to dictation words; Spelling to dictation nonwords.
- e) *Phonological tasks*: Phoneme deletions; Spoonerisms.

Memory:

- a) *Verbal short-term memory (STM)/working memory (WM):* 0,.1, 2-back tasks; WMS-III Auditory immediate memory; Digit span; Nonword repetition.
- b) *Verbal learning/long-term memory (LTM)*: Rey AVLT; WMS-III - Auditory delayed recall & recognition; Paired-associates verbal learning; Story memory recall from WMS.
- c) *Visual STM/WM:* Self ordered pointing test; Feature integration pattern selection task; Visuospatial sequencing; WMS-III Visual immediate memory; Corsi span.
- d) *Visual learning/LTM*: Rey Figure recall; Complex figure recognition - immediate & delayed; WMS-III Visual delayed; Paired-associates visual learning - immediate & delayed.

Visuo-spatial cognition:

- a) *Visuo-spatial skills:* Rey Figure copy; Complex figure copy; Perceptual judgment task-shape and function.
- b) *Visual detection speed*: Finger motor speed; Simple detection-RT; Saccade latency.

- *c) Visuo-spatial attention speed* [1]*:* Memory search-high load-RT; Choice-RT; Detection w/distractors-RT; Feature search-RT; Conjunction search-RT.
- *d) Visuo-spatial attention accuracy:* Memory search-high load-accuracy; Choice-accuracy; Detection w/distractors-accuracy; Feature search-accuracy; Conjunction search-accuracy.
- *e) Visuo-motor control.* Long pins/ short pins pegboard; Grooved pegboard; MLS (Motor performance series battery): steadiness; line following and rotor cycle; - Motor screening test.

Social cognition:

- a) Reading the mind in the eyes; ANT Facial Recognition - speed and accuracy; ANT identification of facial emotions - speed and accuracy; Faux-pas Recognition Test – score; Identity Test - speed and accuracy; Affect Selection - speed and accuracy; 3 Face test -speed and accuracy.

We also evaluated IQ and an overall effect size averaged across all functions, except IQ, to have an indication of overall performance.

It should be recognized that how tasks are attributed to functions is not always consistent in the literature. Disagreements, however, do not regard whether a function is involved in a task at all, but rather under which function heading the author chooses to categorize a task. Thus, categorization difficulties are mitigated when, as in our case, several tasks are used in the assessment and heterogeneity is considered. Our attributions of tasks to functions followed common assessment practices and were agreed by the three authors with experience in clinical and experimental neuropsychology (SR, LA, SH).[2]

---

[1] Note visuo-spatial skills refers to the ability to process and analyse visuo-spatial information; visuo-spatial attention refers to the ability to engage and disengage attention when processing visual displays.

[2] For transparency, we report here differences between our categorizations and those of Hofman et al. (2018). With few exceptions, these differences involved us splitting functions into more homogenous subgroups, as shown below:

1. Among executive functions, we separated reasoning from flexibility and planning.
2. H&al had a category 'attentional capacity' encompassing a number of tasks that we attributed to different functions. We attributed choice RT, feature search, conjoined search, and detection with distractors to *visuo-spatial attention*; the Telephone search test to *sustained attention*; the California Verbal Learning test to *Verbal LTM*; the digit span and n-back to *STM/WM*; the video tracking task to *motor skills*; the control condition of the Stroop to *naming*.
3. We distinguished between visual and verbal memory tasks because of previous evidence of more severe impairments with visual stimuli (see Romani et al., 201).
4. In assessing LTM, we did not distinguish between recall and recognition tasks since these conditions are generally very correlated.
5. We attributed Hayling Sentence completion part B and Naming-semantic interference to inhibitory control. H&al attributed them to 'complex language skills'.
6. H&al attributed non-word reading and semantic interference effects to complex language tasks. We attributed them to reading and inhibitory control, respectively.
7. We attributed processing prosody and emotional context to complex language skills. H&al attributed them to basic language skills.
8. We attributed Digit-symbol coding to flexibility and planning. H&al attributed it to motor skills (both attributions are equally reasonable).
9. Finally, H&al had a category named 'Processing speed' which brought together tasks which tapped different functions, but where speed of processing was measured, for example, simple motor tasks (e.g., motor screening test, saccadic latencies, simple detection, simple RT, Trail making part A) as well as tasks that involved reasoning (Stockings of Cambridge initial thinking time). We have separated tasks according to function, as well as according to RTs or accuracy measures when available.

Table 1

Demographics of studies and participants included in the meta-analysis. SES=socio-economic status.

N of articles/publications = 26; 3 with multiple groups; 3 with the same participants;				
N of PKU samples 23 + 3–3 = 26				
	PKU		**CONTROLS**	
Participants overall N	N = 757		N = 676	
	Mean	SD across samples	Mean	SD across samples
Participants N per sample	31.2	30.4	29.9	17.7
mean age in years across samples	26.5	4.5	26.4	4.6
% females across sample	61	*21*	60	*19*
Current Phe-average μmol/L across samples	954	298		
Current Phe-average SD within sample	307	132		
When treated	N			
Close to Birth	8			
Within 1month	7			
Within 3 months	5			
Continuity of dietary treatment				
Continuous until adolescence	6			
Some ppts stopped before adolescence	3			
Continuous/or not better specified	14			
Matched to controls for				
IQ	6			
Education	8			
SES	3			

Table 2

Distribution of studies and participants according to nationality.

Nationality	N of studies	N of ppts with PKU
Australian	2	21
British	7	338
Canadian	2	22
Dutch	1	56
Swiss	1	20
German	9	196
Hungarian	1	46
Italian	2	33
USA	1	25
Total	26	757

A task that was particularly difficult to categorize was Digit-symbol coding. It involves writing the appropriate numbers below symbols according to a key. We attributed this task to flexibility and planning because it involves moving back and forth between the symbol list and the key showing the proper associations. However, this is a complex task which also involves visuo-motor control and working memory (it is facilitated by keeping in mind the association between digits and symbols). The fact that involves a set of different functions (as well as processing speed) could be one reason why it is often severely impaired (e.g. it is the most impaired task in Aitkenhead et al., 2021).

For a complete list of tasks and measures organized by functions, and a brief description of tasks see Supplementary Materials 1. For the number of measures contributing to each function see Table 3.

3. Results

3.1. Characteristics of reviewed studies

The characteristics of the 26 PKU samples included in our meta-analyses and the distribution of studies and participants according to

Table 3

Summary of weighted Glass' Δ effect sizes for different functions. Significance established with a random effects model and taking into consideration variability of tasks and groups of participants (with group nested in study). Highly significant results are shown in bold. Heterogeneity refers to whether effects are or are not homogeneous across measures.

	Samples characteristics					Effect size				
	Groups	**Measures**	**PKU Participants**		**Control Participants**	**Estimated adjusted Glass' Δ**			**Heterogeneity**	
EXECUTIVE FUNCTIONS	**N**	**N**	**N**	**Current Phe**	**N**	**mean**	**Conf Interval**	**P**	**Q**	**p**
Flexibility and planning	13	34	386	928	343	**-0.51**	(−0.85, −0.16)	**.004**	74.8	**<.001**
Reasoning	4	6	90	825	89	**-1.02**	(−1.7, −0.33)	**.004**	13.3	**.02**
Inhibition	10	18	166	1006	193	-0.13	(−0.39, 0.13)	0.33	27.2	**.05**
Sustained attention	8	11	225	943	208	**-0.65**	(−0.87, −0.43)	**< 0001**	13.2	.21
LANGUAGE										
Higher Language Skills	1	15	37	747	32	**-0.34**	(−0.59, −0.09)	**.008**	34.2	**.002**
Naming RT	3	7	54	1025	60	-0.45	(−0.98, 0.08)	.09	13.5	**.04**
Naming accuracy	2	8	56	810	57	-0.11	(−0.31, 0.09)	.30	1.2	.99
Spelling	1	2	37	720	30	0.01	(−0.33, 0.34)	.97	0.9	.33
Phonological Tasks	1	2	37	720	30	-0.45	(−0.94, 0.04)	.07	0.2	.65
MEMORY										
Verbal STM	7	11	291	903	247	**-0.31**	(−0.56, −0.06)	**.01**	12.5	.25
Verbal learning LTM	7	15	262	936	201	-0.28	(−0.58, 0.02)	.07	17.6	.23
Visual STM	7	9	167	889	182	-0.60	(−1.23, 0.03)	.06	28.2	**<.001**
Visual learning LTM	7	9	146	1067	135	**-0.49**	(−0.75, −0.23)	**<.001**	10.6	.22
VISUO-SPATIAL COGNITION										
Visuo-spatial skills	5	11	103	957	129	**-0.18**	(−0.34, −0.02)	**.03**	6.5	.77
Visual Detection - RT	4	4	133	966	126	-0.31	(−0.67, 0.06)	.11	6.4	.09
Visual spatial attention RT	4	10	138	1091	121	**-0.80**	(−1.02, −0.59)	**<.001**	8.4	.50
Visuo-spatial attention accuracy	3	9	113	806	106	-0.06	(−0.22, 0.1)	.48	3.9	.86
Visuo motor control	6	16	235	857	191	**-0.53**	(−0.79, −0.26)	**<.001**	50.7	**<.001**
SOCIAL COGNITION	4	16	98	860	122	**-0.42**	(−0.66, −0.19)	**<001**	16.9	.33
SUM	97	213	2774		2602	–				
MEAN	5	11	146	898	137	**-0.44**	(−0.56, −0.32)	**<.001**	417	**<.001**
range	–	–	–		–	-1.02–0.006 –		–		
IQ	16	19	407	1008	439	**-0.60**	(−0.78, −0.42)	**<.001**	26.8	.08

nationality are shown in Tables 1 and 2. A full list of studies and their characteristics is presented in Appendix 1.

Participants were all young adults aged up to their early 40 s. Within and across studies there were similar numbers of male and female participants. Average current Phe level was well above the current target of 120–600 μmol/L recommended by European guidelines (average of our PKU groups = 954 μmol/L), but there was high variability both across studies and among participants within studies. While treatment always started within three months of age (this was our selection criterion), age of treatment initiation varied within this period. In our analyses, we will consider potential effects of Phe level and age of treatment initiation on ESs.

All studies assessed AwPKU who were under the care of metabolic centres and had been continuously monitored since infancy. For most studies, it can be assumed that diet was continued at least until adolescence, although this was not always explicitly stated. In two samples, diet was interrupted by 10 years of age: Ris et al. (1994), N = 8 participants; and Ullrich et al., (1996), for an unspecified number of their 8 participants. These samples were only a small proportion of our participants (<16/756, ~2 % of all participants). In Channon et al. (2004), the 25 participants discontinued the diet after 10 years of age, but no more specific details were given. When age of diet discontinuation was reported this was specified in Appendix 1.

To best appraise the consequences of PKU, PKU and control groups should be matched for the education and SEC of the parents because these variables will affect outcomes independent of any effect of PKU. However, among our reviewed studies 10/26 = 38 % matched PKU and control groups for their own education/SES rather than their parents'. This is a conservative choice. It will control for possible environmental effects since the SES of adult participants will be related to that of their parents. It will also, however, reduce any potential differences from controls. Education/SES are related to cognition. Thus, matching groups for education/SES will necessarily reduce differences in cognition. Therefore, the profile shown by our review could be attenuated compared to the profile one would have obtained with unmatched samples.

The large number of AwPKU and controls contributing to our meta-analysis gave excellent power to assess overall differences in performance. Moreover, PKU groups were assessed with a variety of tasks tapping different functions (N = 220 different ESs excluding IQ), allowing generalizations across functions. We had less power for the investigation of individual functions since each has been assessed by only a subset of studies: a mean of 5.5 studies (SD = 3.2) and 12 measures (SD = 7.4) per function.

3.2. Meta-analyses of effect sizes

3.2.1. Effect sizes by functions

All individual Glass' Δ ESs for different tasks and measures, organized by function, are reported in the Supplementary materials 2; Hedges' *g* ESs are also included for comparison. Table 3 shows Glass' Δ estimates for each individual function weighted by variance, number of tasks administered, and number of independent PKU samples who carried out the tasks. Examples of forest plots are shown in Fig. 2. A complete listing of all the forest plots is presented in Supplementary materials 3.

With the exception of spelling, all functions showed negative ESs, reflecting worse performance of PKU groups compared to controls. The performance of AwPKU was − 0.44 of a SD below that of the controls (unweighted average = −0.40) and even more impaired for IQ (ES −0.60). Not all the pooled effect sizes reached significance given the stringent criteria we used for assessment, but overall, results paint a picture of clear and widespread impairment. Importantly, however, there was great variability among functions, with some showing no or very small differences and others showing a strong impairment. The largest impairments were, in order, in reasoning, visuo-spatial attention RT, sustained attention, visuo-motor control, flexibility and planning, visual learning/LTM, social cognition, higher language skills, verbal STM/WM and visuo-spatial skills (all p < .05, ES range = −1.02 to -0.18). Performance, instead, was normal or close to normal in spelling, visuo-spatial attention accuracy, language accuracy, inhibition, simple RT and naming RT (p > .09). Among language tasks, phonological tasks,
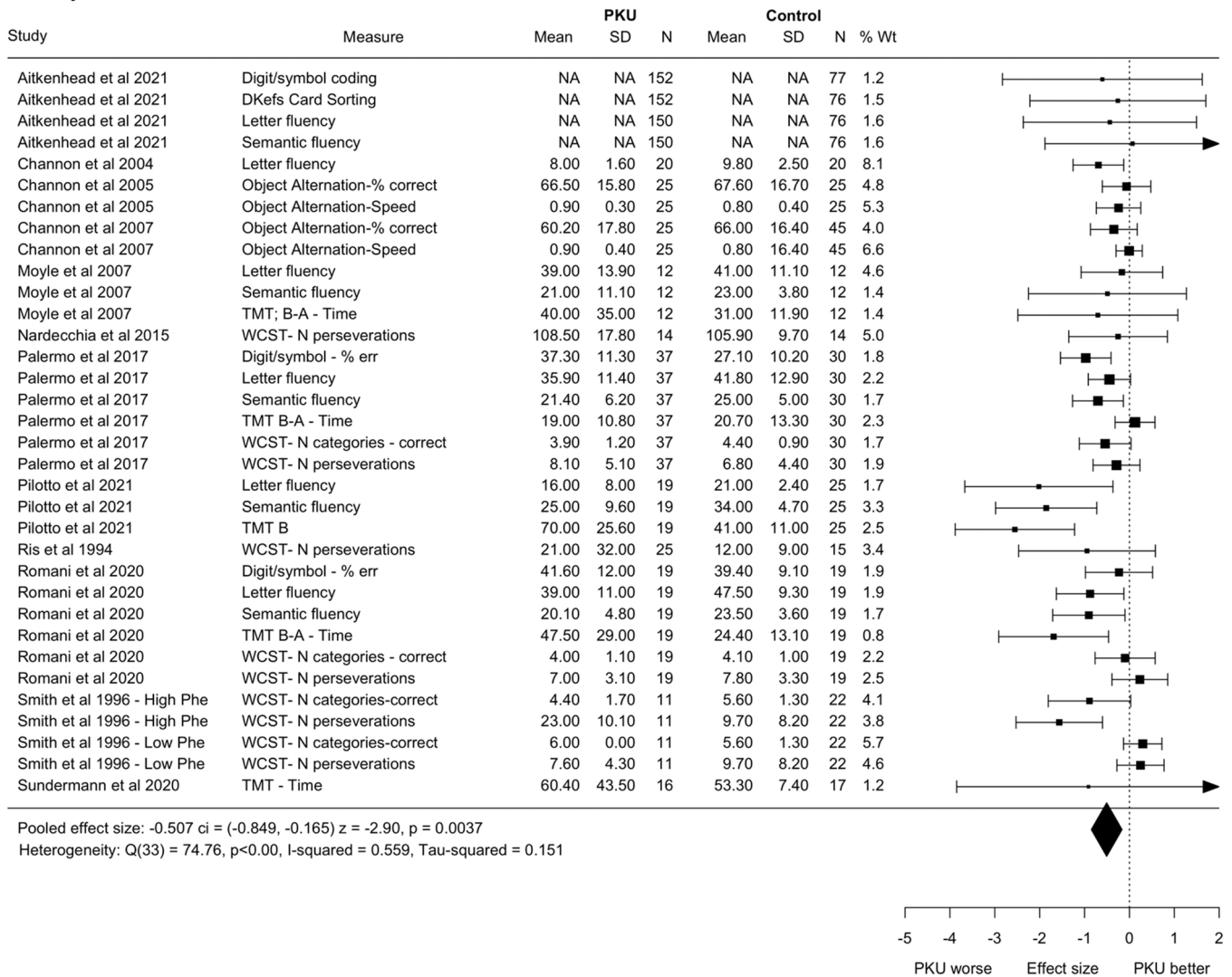
Flexibility

Study	Measure	PKU Mean	SD	N	Control Mean	SD	N	% Wt
Aitkenhead et al 2021	Digit/symbol coding	NA	NA	152	NA	NA	77	1.2
Aitkenhead et al 2021	DKefs Card Sorting	NA	NA	152	NA	NA	76	1.5
Aitkenhead et al 2021	Letter fluency	NA	NA	150	NA	NA	76	1.6
Aitkenhead et al 2021	Semantic fluency	NA	NA	150	NA	NA	76	1.6
Channon et al 2004	Letter fluency	8.00	1.60	20	9.80	2.50	20	8.1
Channon et al 2005	Object Alternation-% correct	66.50	15.80	25	67.60	16.70	25	4.8
Channon et al 2005	Object Alternation-Speed	0.90	0.30	25	0.80	0.40	25	5.3
Channon et al 2007	Object Alternation-% correct	60.20	17.80	25	66.00	16.40	45	4.0
Channon et al 2007	Object Alternation-Speed	0.90	0.40	25	0.80	16.40	45	6.6
Moyle et al 2007	Letter fluency	39.00	13.90	12	41.00	11.10	12	4.6
Moyle et al 2007	Semantic fluency	21.00	11.10	12	23.00	3.80	12	1.4
Moyle et al 2007	TMT; B-A - Time	40.00	35.00	12	31.00	11.90	12	1.4
Nardecchia et al 2015	WCST- N perseverations	108.50	17.80	14	105.90	9.70	14	5.0
Palermo et al 2017	Digit/symbol - % err	37.30	11.30	37	27.10	10.20	30	1.8
Palermo et al 2017	Letter fluency	35.90	11.40	37	41.80	12.90	30	2.2
Palermo et al 2017	Semantic fluency	21.40	6.20	37	25.00	5.00	30	1.7
Palermo et al 2017	TMT B-A - Time	19.00	10.80	37	20.70	13.30	30	2.3
Palermo et al 2017	WCST- N categories - correct	3.90	1.20	37	4.40	0.90	30	1.7
Palermo et al 2017	WCST- N perseverations	8.10	5.10	37	6.80	4.40	30	1.9
Pilotto et al 2021	Letter fluency	16.00	8.00	19	21.00	2.40	25	1.7
Pilotto et al 2021	Semantic fluency	25.00	9.60	19	34.00	4.70	25	3.3
Pilotto et al 2021	TMT B	70.00	25.60	19	41.00	11.00	25	2.5
Ris et al 1994	WCST- N perseverations	21.00	32.00	25	12.00	9.00	15	3.4
Romani et al 2020	Digit/symbol - % err	41.60	12.00	19	39.40	9.10	19	1.9
Romani et al 2020	Letter fluency	39.00	11.00	19	47.50	9.30	19	1.9
Romani et al 2020	Semantic fluency	20.10	4.80	19	23.50	3.60	19	1.7
Romani et al 2020	TMT B-A - Time	47.50	29.00	19	24.40	13.10	19	0.8
Romani et al 2020	WCST- N categories - correct	4.00	1.10	19	4.10	1.00	19	2.2
Romani et al 2020	WCST- N perseverations	7.00	3.10	19	7.80	3.30	19	2.5
Smith et al 1996 - High Phe	WCST- N categories-correct	4.40	1.70	11	5.60	1.30	22	4.1
Smith et al 1996 - High Phe	WCST- N perseverations	23.00	10.10	11	9.70	8.20	22	3.8
Smith et al 1996 - Low Phe	WCST- N categories-correct	6.00	0.00	11	5.60	1.30	22	5.7
Smith et al 1996 - Low Phe	WCST- N perseverations	7.60	4.30	11	9.70	8.20	22	4.6
Sundermann et al 2020	TMT - Time	60.40	43.50	16	53.30	7.40	17	1.2

Pooled effect size: -0.507 ci = (-0.849, -0.165) z = -2.90, p = 0.0037
Heterogeneity: Q(33) = 74.76, p<0.00, I-squared = 0.559, Tau-squared = 0.151

Fig. 2. Examples of forest plots showing effect sizes for differences between adults with PKU and age-matched controls grouped by cognitive function. Glass's Δ effect sizes corrected for small sample size are pooled to calculate an overall effect size (black diamond) that takes into account the number of tasks nested within independent PKU samples. PKU worse= PKU worse than controls; PKU better= PKU better than controls. % Wt= is the relative weight of each study in the calculation of the overall effect size. The study by Aitkenhead et al. (2021) did not report the means and SDs of PKU and control participants, but z scores of the PKU group from the control group. We estimated Glass's Δ effect sizes from these scores.

which showed a marginally significant impairment (p = .07), have a strong monitoring component. Verbal learning/LTM and Visual STM were also borderline (p = .07 and p = .06). As expected, within-function heterogeneity was often significant since the degree of impairment varied depending on measures and PKU groups.

3.2.2. Publication bias

Our overall estimate of ES including IQ was − 0.46 (CI= −0.57, −0.35; N = 233). To examine the possibility of a publication bias, we first plotted Glass' Δ effects against the size of the PKU sample as a measure of study precision. The resulting funnel plot showed a moderate but non-significant asymmetry (Egger's test, using PKU sample size as moderator: z = −0.66; p = .51; see Fig. 5 in Appendix 2). To run a sensitivity analysis, we ordered all our individual ESs (N = 233) by precision, based on the number of participants contributing to the estimate, and subdivided the data into 11 bins. The first bin included the first 40 most precise ESs. Each of the following bins included an additional 20 measures (the last one had 13), progressively increasing the proportion of less precise measures. Fig. 3 plots cumulative effect sizes
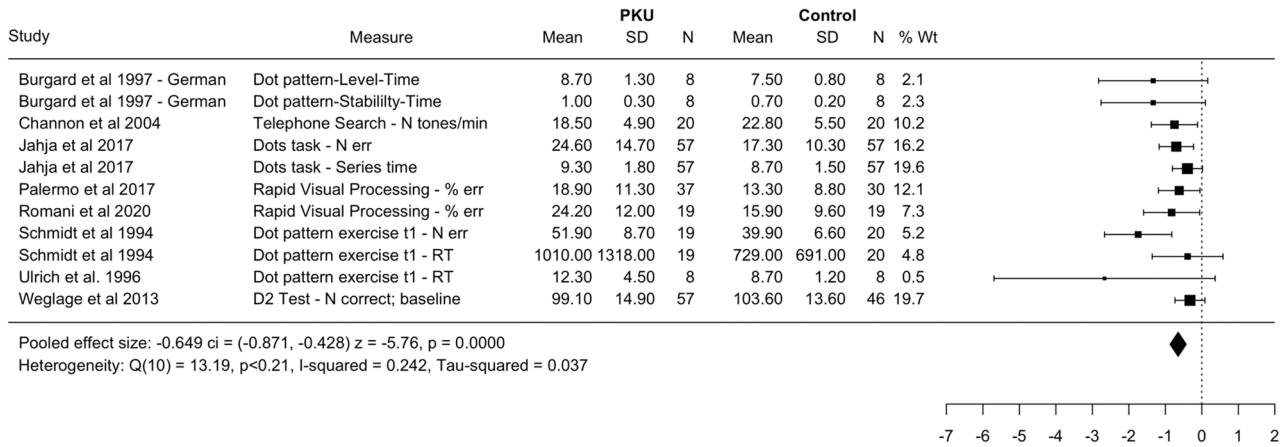
calculated for each of these eleven bins. If a publication bias significantly distorted our estimates, estimates should have increased across the bins. This did not happen. Estimates remained relatively stable around the final estimate of − 0.44 (ranging and between −0.37 and −0.53) and were never close to the "no difference" value of 0.
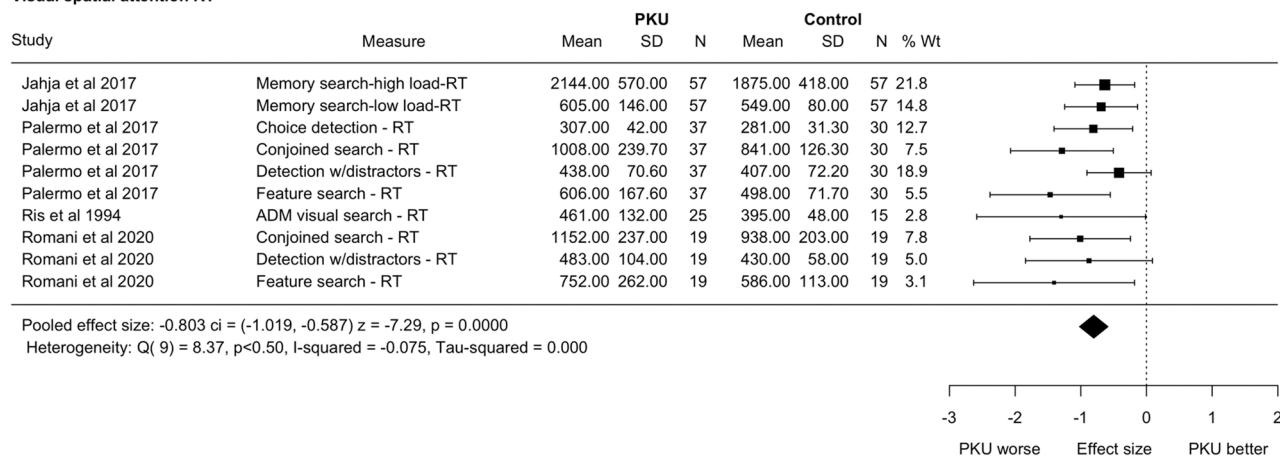
3.2.3. Differences among functions

As noted above, performance was not uniformly affected across functions and measures. Here we assess differences between groups of measures e.g., between tasks measuring speed vs. accuracy and between language vs. visual tasks. Moreover, we assessed whether ESs remained significant even when we considered tasks where speed was not a premium.

To assess a contrast between speed vs accuracy, we considered all tasks where performance was measured both ways. The effect size was much stronger for speed than accuracy measures (N = 39 each; unweighted mean ES = **-0.24** vs **-0.56;** weighted ES **-0.19** vs **-0.47**). We assessed the significance of this difference using a mixed model analysis with '*type of measure*' (RTs vs accuracy) and *current Phe* of the group as

Sustained attention

Study	Measure	PKU Mean	PKU SD	PKU N	Control Mean	Control SD	Control N	% Wt
Burgard et al 1997 - German	Dot pattern-Level-Time	8.70	1.30	8	7.50	0.80	8	2.1
Burgard et al 1997 - German	Dot pattern-Stabililty-Time	1.00	0.30	8	0.70	0.20	8	2.3
Channon et al 2004	Telephone Search - N tones/min	18.50	4.90	20	22.80	5.50	20	10.2
Jahja et al 2017	Dots task - N err	24.60	14.70	57	17.30	10.30	57	16.2
Jahja et al 2017	Dots task - Series time	9.30	1.80	57	8.70	1.50	57	19.6
Palermo et al 2017	Rapid Visual Processing - % err	18.90	11.30	37	13.30	8.80	30	12.1
Romani et al 2020	Rapid Visual Processing - % err	24.20	12.00	19	15.90	9.60	19	7.3
Schmidt et al 1994	Dot pattern exercise t1 - N err	51.90	8.70	19	39.90	6.60	20	5.2
Schmidt et al 1994	Dot pattern exercise t1 - RT	1010.00	1318.00	19	729.00	691.00	20	4.8
Ulrich et al. 1996	Dot pattern exercise t1 - RT	12.30	4.50	8	8.70	1.20	8	0.5
Weglage et al 2013	D2 Test - N correct; baseline	99.10	14.90	57	103.60	13.60	46	19.7

Pooled effect size: -0.649 ci = (-0.871, -0.428) z = -5.76, p = 0.0000
Heterogeneity: Q(10) = 13.19, p<0.21, I-squared = 0.242, Tau-squared = 0.037

Visual spatial attention RT

Study	Measure	PKU Mean	PKU SD	PKU N	Control Mean	Control SD	Control N	% Wt
Jahja et al 2017	Memory search-high load-RT	2144.00	570.00	57	1875.00	418.00	57	21.8
Jahja et al 2017	Memory search-low load-RT	605.00	146.00	57	549.00	80.00	57	14.8
Palermo et al 2017	Choice detection - RT	307.00	42.00	37	281.00	31.30	30	12.7
Palermo et al 2017	Conjoined search - RT	1008.00	239.70	37	841.00	126.30	30	7.5
Palermo et al 2017	Detection w/distractors - RT	438.00	70.60	37	407.00	72.20	30	18.9
Palermo et al 2017	Feature search - RT	606.00	167.60	37	498.00	71.70	30	5.5
Ris et al 1994	ADM visual search - RT	461.00	132.00	25	395.00	48.00	15	2.8
Romani et al 2020	Conjoined search - RT	1152.00	237.00	19	938.00	203.00	19	7.8
Romani et al 2020	Detection w/distractors - RT	483.00	104.00	19	430.00	58.00	19	5.0
Romani et al 2020	Feature search - RT	752.00	262.00	19	586.00	113.00	19	3.1

Pooled effect size: -0.803 ci = (-1.019, -0.587) z = -7.29, p = 0.0000
Heterogeneity: Q(9) = 8.37, p<0.50, I-squared = -0.075, Tau-squared = 0.000

Visuo motor control

Study	Measure	PKU Mean	PKU SD	PKU N	Control Mean	Control SD	Control N	% Wt
Bartus et al 2018	Motor screening test-Err score	10.10	2.10	46	8.50	1.90	31	6.7
Palermo et al 2017	Grooved pegboard - Time	65.10	11.70	37	59.20	5.10	30	4.5
Pietz et al 1998	Fast Video Tracking - Level dist	37.90	15.20	57	26.50	8.60	40	6.0
Pietz et al 1998	Fast Video Tracking - Stability dist	16.60	11.80	57	8.50	6.40	40	5.9
Pietz et al 1998	Long pins/ short pins pegboard-Time	452.00	72.20	57	430.00	77.50	40	7.5
Pietz et al 1998	MLS line following - Duration	14.40	11.80	57	10.70	8.60	40	7.0
Pietz et al 1998	MLS line following - N err	19.60	10.90	57	15.10	9.40	40	7.2
Pietz et al 1998	MLS rotor circle - Duration	101.00	52.00	57	89.00	58.00	40	7.6
Pietz et al 1998	MLS rotor circle - N err	35.00	12.00	57	33.00	12.00	40	7.5
Pietz et al 1998	MLS steadiness - Duration	3.00	4.90	57	1.70	3.50	40	6.9
Pietz et al 1998	MLS steadiness - N err	4.70	6.00	57	3.10	4.70	40	7.1
Pietz et al 1998	Slow Video Tracking-Level dist	17.20	6.70	57	13.10	3.50	40	5.9
Pietz et al 1998	Slow Video Tracking-Stability dist	6.50	3.00	57	4.90	3.60	40	7.6
Pilotto 2021	TMT A	29.00	13.60	19	20.00	5.50	25	2.9
Romani et al 2020	Grooved pegboard - Time	76.70	11.00	19	70.60	9.00	19	5.3
Weglage et al 2013	TMT - Time at baseline	93.20	63.60	57	77.00	19.70	46	4.4

Pooled effect size: -0.525 ci = (-0.788, -0.263) z = -3.92, p = 0.0001
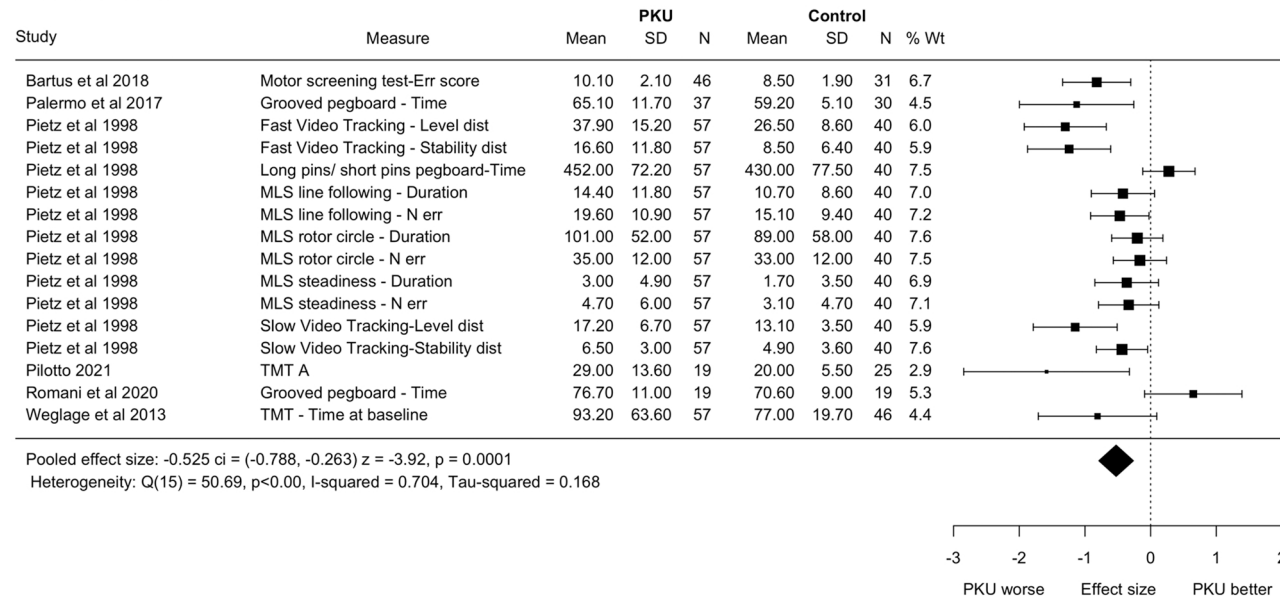Heterogeneity: Q(15) = 50.69, p<0.00, I-squared = 0.704, Tau-squared = 0.168

Fig. 2. (continued).

fixed factors, and 'task nested within group' as a random factor. Results showed a significant main effect of type of measure (coefficient for acc/rt difference = −0.25, z = −3.29, p < .001), a marginally significant, but small main effect of Phe (Phe: coefficient =−0.0007, z = −1.81, p = .07) and a significant interaction type of measure X group (coefficient = 0.001, z = 2.88, p = .004). Differences between speed and accuracy ESs were generally smaller for groups with higher Phe, which were more uniformly impaired.
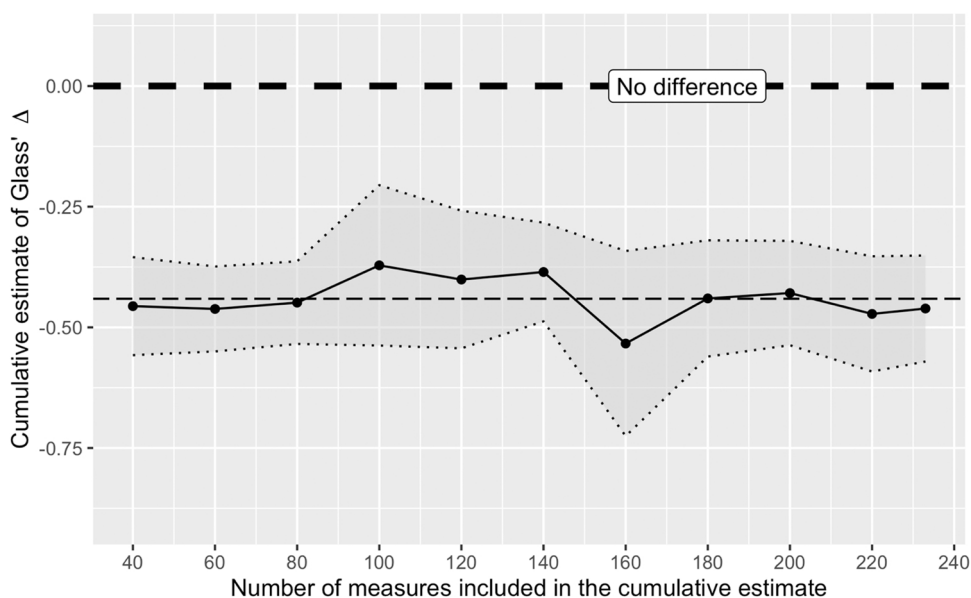
Fig. 3. Glass' Δ effect sizes for differences between PKU and controls based first on the largest studies (first dot on the plot) and then accumulating studies, in order, from the largest samples to the smallest. The solid line is the cumulative estimate of Glass' Δ as studies are added. The shaded area shows 95% confidence limits. 40 measures are included in the first bin because there was very little difference in study precision (number of participants) up to this point.

Table 4
Bivariate Pearson r correlation of raw effect sizes with characteristics of PKU groups.

		Adjusted Glass' Δ	Phe	Age of treatment
Average Phe	r	-.24**		
	Sig. (2-tailed)	< .001		
	N	214		
Age of treatment initiation	r	.13	.38**	
	Sig. (2-tailed)	.06	< .001	
	N	205	187	
Study year	r	.08	-.44**	-.63**
	Sig. (2-tailed)	.20	< .001	< .001
	N	233	215	205

Considering differences between language and visual tasks (excluding accuracy measures showing no impairment), ESs were stronger with visual stimuli (unweighted mean ES= **-0.80** vs **-0.34**; weighted = **-0.54** vs **-0.27**; N = 25 and N = 29). To assess significance, we ran a mixed model analysis with '*type of measure*' (this time: verbal vs visual stimuli) and *current Phe* of the group as fixed factors, and '*task nested into group*' as a random factor. Results showed a significant effect of visual/verbal task-type (difference coefficient = −0.29, z = −2.37, p = .02) but no main effect of Phe (coefficient = −0.0002, z = −0.57, p = .57) or interaction with Phe (coefficient = −0.0001, z = −0.25, p = .81).

Finally, we considered whether the ES for all executive functions taken together (tasks involving flexibility, inhibitory control, reasoning and sustained attention) remained significant when we excluded speed measures and/or tasks with a focus on speed (e.g., complete the task as quickly as possible). The effect size remained significant (N = 50; ES = −0.51, z = −4.11, p < .001) indicating that speed is not the only issue in PKU.

3.3. Moderation of effect sizes by metabolic control

To assess whether and how effect sizes were moderated by

differences in metabolic control among the PKU groups, we carried out correlations where all individual effect sizes for all functions were correlated with three measures:

a) Average Phe level at the time of testing.

b) Age of treatment initiation. Most studies gave some information regarding treatment initiation that we operationalized as follows: soon after birth: score 1; N = 129; within a month: score 2; N = 52; within 2 months: score 3; N = 10; within 3 months: score 4; N = 19.

a) Year of study. It is possible that clinical care and, therefore, metabolic control was better in PKU cohorts reported in more recent publications.

Pearson r correlations among Glass' Δ ESs, average Phe, age of treatment initiation, and year of study are shown in Table 4. There were significant correlations between Glass' Δ ESs, and Phe, with larger impairments in groups with higher Phe. There were also correlations between age of treatment initiation and Phe, with higher current Phe levels associated with later treatment initiation. Finally, year of study was highly negatively correlated with average Phe level and age of treatment initiation. As expected, more recent studies assessed PKU groups with lower Phe and earlier treatment initiation.

We further considered the influence of concurrent Phe average, age of treatment initiation, and year of study on ESs, including these variables as moderators in a mixed model analysis. The adjusted Glass' Δ was the dependent measure, task nested within PKU group was a random factor, and each of the above variables was used as a moderator. Results showed an influence of Phe (coefficient = −0.0004, z = −1.94, p = .05). There was no main effect of year of study (coefficient = 0.0004, z = 0.07, p = .95) or age of treatment initiation (coefficient = −0.02, z = −0.25, p = .80). There was a year of study X Phe interaction (z = 2.83, p = .005) since earlier studies had a larger range of Phe and, therefore, a clearer effect. There was no age of initiation X Phe (z = −71, p = .48) interaction. In sum, when Glass' Δ ESs were weighted by precision (rather than using raw values as in the correlation analyses) any effect other than Phe disappeared. The lack of an effect of Phe initiation is not surprising given that our index had a very limited range and only gave a rough idea of when metabolic control was reached.

Fig. 4 shows a scatter plot of weighted effect sizes by average Phe. A coefficient of − 0.0004 means that for every 100 μmol of Phe the effect size decreases by 0.04 SDs. In a normal distribution, 1.6 % of the distribution lies between that SD and the point which identifies 50 % of the
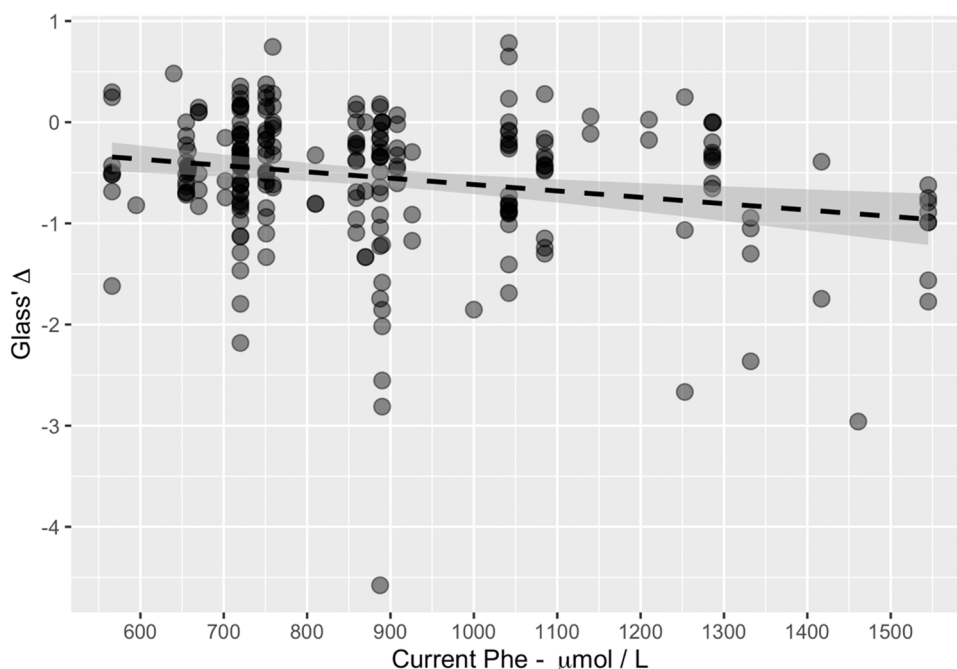
Fig. 4. Weighted Glass' Δ effect sizes as a function of the average current Phe level of the PKU groups.

area (the mean). Therefore, a decrease of 0.04 SD means losing 1.6 places from a mean of 50 for every increase of 100 μmol. Putting this in the context of the studies here, a change in current Phe of around 500 μmol, would mean losing 8 places or, in other words, moving from the 50th percentile to the 42nd. Considering the wider range of Phe values in the older studies and the large estimate of the impact of Phe, a change of around 500 μmol could result in a loss of up to 16 places (from 50th percentile to 34th).

4. General discussion

Our study aimed to provide a snapshot of the state of the literature and current cognitive outcomes in adults with PKU (AwPKU) who were born after the introduction, in the late sixties, of new-born screening policies and consequent early treatment in western countries. We wanted to expand previous reviews by providing estimates of performance for a larger set of cognitive functions assessed with more participants and more tasks. For this purpose, we have computed standardized differences (effect sizes=ESs) between the performance of participants with PKU and controls and then used meta-analyses to combine tasks and measures according to function to provide a profile of PKU outcomes.

Our results come from research articles published from 1994, when the first cohorts of early-treated individuals with PKU reached adulthood, to the present. They show that, at present, cognitive outcomes are suboptimal, resulting, on average, in a reduction in cognitive ability of − 0.44 SD (our weighted Glass' Δ ES) for all cognitive skills combined, excluding IQ, of − 0.60 SD for IQ and of − 0.46 SD overall. These effect sizes correspond to the 33rd, 27th and 32nd percentile in a distribution of 100 scores, which means that, on average, AwPKU lose 17, 23 and 18

index points, respectively, from a mean of 50 compared to age-matched controls.[3] Impairments of this size would be expected to impact educational and occupational attainment. Our estimates are similar to the averages by Moyle et al. (2007a), Bilder et al. (2016), and Albrecht et al. (2009), but they were obtained considering more studies, more participants, and more functions. This increases our confidence in the results.

Suboptimal outcomes were related to suboptimal metabolic control. Average Phe levels in our samples were much higher than the levels recommended by current European guidelines (van Spronsen et al., 2017). We found that an average group increase of 500 μmol of current Phe resulted in losing 8 percentile points (Albrecht et al., 2009 did not find a significant effect in adults, but this was likely due to lack of power) and this estimate is likely to be on the low side since it does not consider intra-sample variability which was high and systematically higher than in control groups (see SDs reported in Supplementary materials 2). It is important to stress, however, that we used only a single value of Phe collected at the time of neuropsychological assessment as an index of dietary control. We did not have enough data to consider the lifetime exposure, or Phe exposures at other ages and we did not consider intra-individual variability. Our review was focused on the profile of early-treated AwPKU and results concerning the relationship with metabolic control have clear limitations (see the later "limitations" section).

4.1. Differences among functions

While cognition showed an overall impairment, there were strong differences among functions. The functions which appeared most affected (and significantly affected) were in order: *reasoning* (ES =

---

[3] Effect sizes can be easily converted into percentiles. An effect size measures the deviation from the control mean in number of SD (or z-scores). Normal distribution tables provide values for the % areas of the normal distribution that correspond to given z scores. If one considers that 50% of the distribution lies to the left of the mean, one can calculate the percentage of the normal distribution which is below a z-point by subtracting the % value provided by the table from 50. This will give the percentile score.

−1.02), *visuo-spatial attention speed* (ES = −0.80), *sustained attention* (ES=−0.65*) visuo-motor control* (ES = −0.53), *flexibility and planning* (ES = −0.51), *visual LTM/learning* (ES = −0.50) and *social cognition* (ES = −0.42). Impairments of this size mean dropping between 16 and 35 index points from an average of 50 in a normal distribution of 100 scores. Other functions were impaired, but to a lesser extent (*higher language skills, working memory* and *visuo-spatial skills*) and still other functions (*spelling, visuo-spatial accuracy, language accuracy* and *inhibition*) were unimpaired (see also Brumm et al., 2004 for partially consistent results).

Visuo-spatial attention speed and visuo-motor control were among the most severely affected skills. Visual attention is a skill commonly engaged in daily living when we scan for objects in our visual environment. It is tapped by laboratory tasks where participants are asked search for a target in a visual display. This skill was not separately considered by other reviews (i.e., it was part of the more general category "attention" in Hofman et al., 2018), but impairments in congruent tasks were reported in previous studies (see the Albrecht et al., 2009 review reporting severe impairments in 'choice reaction time' tasks and Huijbregts et al., 2002 reporting impairments in similar tasks in children with PKU). Possibly related to these impairments, we also found a significant difficulty in tasks involving *visuo-motor control*. For example, when participants are requested to insert shaped pegs in consistent holes or track a visual target with a tool (for consistent results from studies not included in our meta-analyses see Griffiths et al., 1995; Luciana et al., 2001; Weglage et al., 1995). This skill is also important for everyday tasks involving manual dexterity.

In terms of executive functions, not all functions in this domain were impaired to the same degree. There were significant difficulties in tasks involving *flexibility* (see also Bilder et al., 2016; Brumm et al., 2004; Hofman, 2018: for the category 'fluency and set shifting'), *sustained attention* (see also Bilder et al., 2016; Hofman et al., 2018 for the categories 'attention' and 'vigilance'; Moyle et al., 2007a), and *reasoning* (see Hofman et al., for the category 'complex executive functions'), but no significant impairment in *inhibition* (see Bilder et al., 2016 and Moyle et al., 2007a for contrary results). It is possible that difficulties in inhibitory control reduce with age. However, contrary to what is sometimes assumed, difficulties of inhibitory control are not systematically reported even in children with PKU (see the review of Christ et al., 2010 for failure to find impairments in the Stroop task). Moreover, child studies have not always assessed performance in interference conditions against performance in baseline/control conditions, which does not allow one to establish whether inhibitory control is disproportionally affected (e.g., see Jahja et al., 2014). Future studies should confirm whether or not there are difficulties in this domain, independent of task complexity and other skills. Finally, we found weak, but still significant impairments in *verbal STM*. It is possible that impairments arise mainly when tasks stress a monitoring, WM component rather than storage (also see Bilder et al., 2016 and Moyle et al., 2007a for no significant STM impairments in AwPKUand Christ et al., 2010 for no impairments in the storage component of working memory in children with PKU).

When we compared speed and accuracy measures in the same tasks, the average effect size was more than two times larger for speed than accuracy. This is consistent with processing speed being a key weakness in PKU (see also Albrecht et al., 2009; Moyle et al., 2007a). Our results, however, showed no speed-accuracy trade-offs in the sense that ESs for accuracy measures were still negative, although not always significant (see Albrecht et al., 2009 for a similar observation). Moreover, impairments in executive functions were significant even when performance was not measured in terms of speed, and speed was not the focus of the task. Together, these results suggest that impairments in AwPKU go beyond a reduction in processing speed (see Janos et al., 2012 for similar results with PKU children; and Romani et al., 2018 for results supporting impairments beyond a generalized reduction in speed of processing). Additionally, we found that performance was generally worse in visuo-spatial than language tasks (memory tasks with spatial vs verbal

stimuli; search tasks vs naming tasks; see Canton et al., 2019; Janzen and Nguyen, 2010 for similar results in child studies). These contrasts among functions point to a cognitive profile which is typical of PKU. We will now turn to discuss the specificity of this profile and its relationship to neurophysiological mechanisms.

4.2. Specificity of the PKU cognitive profile compared to other populations

PKU deficits in executive functions and sustained attention overlap with similar deficits seen in other conditions affecting brain functions such as closed-head injuries, degenerative disorders and aging (for an overlap of deficits between children with PKU and ADHD see Stevenson & McNaughton, 2013; Burton et al., 2015; for an overlap between children with PKU and HIV see Bisiacchi et al., 2018). These similarities are not surprising. Complex functions involve networks of brain areas, so that lesions in different regions and in the white matter connecting these regions can all create difficulties. To understand and monitor PKU, what is more important is to establish how the PKU cognitive profile differs from what is seen in other disorders. We will provide two illustrative examples considering ADHD and developmental dyslexia.

In both PKU and in ADHD there may be an overall reduction in speed of processing. However, in PKU this reduction occurs across all trials (see Romani et al., 2018 for a demonstration). Instead in ADHD, average speed is unduly affected by a relatively small number of very long RTs due to lapses of attention, (see Kofler et al., 2013 for a review and meta-analysis). In fact, in ADHD there is a preference for speed over accuracy (for more errors in the face of preserved speed see Mulder et al., 2010; Wilding et al., 2007), while in PKU it is the opposite: speed is slowed to maintain accuracy (see also De Felice et al., 2018). Differences with developmental dyslexia are also marked. In developmental dyslexia, the main difficulties involve spelling, phonological processing, and verbal learning. There are less marked difficulties in executive functions (e.g., Di Betta and Romani, 2006; Romani et al., 2008, 2015). This is the opposite of what we have seen in PKU. The pattern of impairment in PKU is closer to what is seen in conditions involving dopamine imbalances and white matter damage such as Parkinson's disease, multiple sclerosis, and healthy aging, but direct comparisons are not available and any similarity should be considered with great caution (e.g., for similarity with Parkinson's see Evans et al., 2004; Tufekcioglu et al., 2016; Velema et al., 2015). A comparison with aging is particularly important given possible interactions between high levels of Phe and normal degenerative processes that arise from aging.

In older adults, as well as in PKU, there is a reduction of processing speed, but accuracy is minimally impaired or normal (for results in the visuospatial domain see Hommel et al., 2004; Folk and Lincourt, 1996; Foster et al., 1995). In older adults as well, there are impairments in executive functions, particularly those involving flexibility and monitoring (see Bucur and Madden, 2010; Delaloye et al., 2009), but inconsistent deficits in inhibitory control (for lack of differences see with the Stroop test, see Delaloye et al., 2009; Verhaeghen and De Meersman, 1998; with the Flanker inhibitory control and attention test, see Hamilton et al., 2017; with the Hayling test, see Delaloye et al., 2009). Furthermore, aging also affects visuo-motor coordination (Hamilton et al., 2017; Volkow et al., 1998) and there are indications of stronger impairments in the visuo-spatial than the verbal domain (see Hale and Myerson, 1996; Lawrence, Myerson, and Hale, 1998; Shafto and Tyler, 2014; but also, Park et al., 2002 for contradictory findings). In contrast, however, memory and learning are clearly affected in aging (see Salthouse, 2004), but less affected compared to other functions in PKU.

In PKU, impairments in processing speed may derive from damage to white matter tracts and more pronounced impairments in visuo-spatial than verbal tasks is consistent with matter in posterior regions being more vulnerable (see Pilotto et al., 2021 and Anderson and Leuzzi, 2010 for a review). Impairments in tasks involving executive functions may stem from a combination of impairments in white matter tracts–since efficient and fast connections among brain areas are crucial for complex

tasks–and depletion of neurotransmitters (i.e., dopamine; see Boot et al., 2017). Similarities with aging may derive from an overlap in neurophysiological damage (for demyelination in aging see Eckert et al., 2010; Grueter and Schulz, 2012; for a decline in brain dopamine activity see Backman et al., 2010). Instead, relatively more preserved memory and naming in PKU may derive from preservation of the subcortical structures involved in memory encoding, and of cortical structures involved in storing information.

4.3. Limitations

Our results are important in providing an estimate of outcomes in current cohorts of AwPKU and in highlighting the pattern of spared and impaired abilities. However, it cannot be stressed enough that averages mask strong individual variability (e.g., see Palermo et al., 2017). At an individual level, some AwPKU will display impairments which are much milder or much more severe than the average. Moreover, although cognitive performance is clearly linked to metabolic control, there are indications of exceptions where some individuals are less affected by the toxic effect of high Phe (see Leuzzi et al., 2020, van Vliet et al., 2018). Variability in sensitivity to Phe may be related to environmental factors, but also to individual neurobiological characteristics which may allow some people to minimize the deleterious effects of Phe (see Boot et al., 2017; Dijkstra et al., 2021).

At the group level, several factors may affect our estimates of impairment. Firstly, level of impairment may be underestimated because results are gathered from participants who are in contact with a clinical care team and are willing to engage in research. The metabolic control of AwPKU who are lost to follow-up may be worse. In the opposite direction, the inclusion of older cohorts could potentially overestimate levels of impairment. However, when year of study was entered as a moderator we found no significant effect, showing that impairment was not significantly greater in older cohorts. Lastly, one should consider that future cohorts including older participants may show an increased level of impairment. All our results come from groups of young adults (average 26.5 years old, SD = 4.5) who have generally maintained a PKU diet up to adolescence but have relaxed their diet since. We do not know whether the degree of impairment will remain stable over the years or not. There is little indication of deterioration at present (see Feldmann, 2019; Nardecchia, 2015), but this is the first generation of early-treated people with PKU to reach middle age and there is some evidence that white matter damage increases with age (see Nardecchia et al., 2015; Mastrangelo et al., 2015). Cognitive performance may deteriorate accordingly when the brain has been exposed to high Phe levels for many years (see Vardy et al., 2020 for a discussion; see also Pilotto et al., 2021 for evidence of possible interactions between the effect of high Phe and aging).

Although we had enough power to detect differences among cognitive functions, our power to estimate level of impairment for some individual functions was more limited. An extreme example was language functions, where narrative skills and spelling were assessed only in one sample of AwPKU. Moreover, we have noted that attribution of tasks to functions has an arbitrary element, since tasks are never pure and typically involve more than one ability. However, the fact that we considered a variety of tasks when estimating ESs and effects were estimated considering the heterogeneity of the measures significantly ameliorated this difficulty.

Finally, our estimates of the impact of metabolic control on cognitive impairments were limited. For most studies we could assume that a phenylalanine restricted diet, in some form, was maintained until adolescence, but time of diet discontinuity was not provided by most studies. This is, in fact, difficult to establish since diet adherence may vary in degree at different times. Phe levels at different ages were not systematically available, nor were measures of Phe fluctuations, which could be equally important in determining outcomes (see Romani et al., 2017, 2019).

Ideally, meta-analytic assessments of the effects of Phe on performance should be based on pooling within-group correlations (see use of this methodology by Waisbren et al., 2008 with PKU children and Fonnesbeck et al., 2013 with mixed-age groups) and considering effects of concurrent Phe, historical Phe levels, and Phe fluctuations to partial out effects. Even stronger results will arise from longitudinal studies where the effects of changes in metabolic control on cognition are assessed within participants, thus controlling for differences in socio-economic status, education, and genetic potential that may confound outcomes in between-participant studies. Unfortunately, these studies are at present very limited (see Thomas et al., submitted for publication).

4.4. Conclusions

Our results show that treatment in current PKU cohorts was successful in avoiding the more severe cognitive impairments linked to untreated disease. Still, there remain significant cognitive impairments that vary in magnitude across different functions. The average impairment was close to half a SD below the control mean, but, while some cognitive functions were close to normal (verbal memory, naming, spelling, inhibition, visuo-spatial attention accuracy), others were severely impaired (reasoning, visuo-spatial attention RT, sustained attention, and visuo-motor control) and average group performance implies much more severe impairments in a portion of individuals. These impairments may potentially curtail the aspirations of AwPKU and affect their ability to cope with the commitments of work and family life. Suboptimal outcomes were linked to suboptimal Phe levels suggesting that differences from controls may be further reduced if metabolic control improves with more support, better dietary supplements, and with the increased introduction of pharmacological and genetic treatments. The disease-specific cognitive profile identified by our review will be important to consider when monitoring disease progression and the effectiveness of new treatment interventions (see also the recommendations of the European Guidelines; van Wegberg et al., 2017).

Statements and declarations

Appendix 1

See in Table 5.

Table 5

Demographic and metabolic characteristics of the PKU groups included in the meta-analysis. NF = number of females. SD- standard deviation. Birth=treated close after birth. Ed= Education; SE=socio-economic status; m=month/s. All groups continuously monitored for most groups it is difficult to ascertain will when diet was maintained although it can be assumed that it was maintained till adolescence. * Phe as IDC> 18.

Study	Subgroup/condition	Nation	N	N F	Age Mean	Age SD	Current phe in μmol/L Mean	Current phe in μmol/L SD	N	N F	Age Mean	Age SD	Age of diet treatment		Controls matched for
Abgottspon et al. (2022)	all on diet	Swiss	20	10	31.4	9	703	203	40	20	29.8	**8.2**	< 1 m		Ed
Aitkenhead et al. (2021)*		British	154	85	32.30	9.04	908	359	76	46	36.30	9.15	< 1 m	–	Ed
Bartus et al. (2018)		Hungary	46	21	29.5	–	595	279	31	11	25	–	birth	–	–
Burgard et al. (1997)	on diet	German	8	3	19.7	1.1	870	186	8	3	20.2	2.4	< 1 m	–	IQ
Channon et al. (2004)	on diet	British	20	8	24.6	4.6	859	285	20	8	24	4	< 1 m	–	IQ, Ed, SES
Channon et al. (2005)	on diet	British	25	12	26.7	4.9	759	261	25	12	26.5	5.5	< 1 m	–	IQ, Ed
Channon et al. (2007)	off diet	British	25	12	26.7	4.9	1286	198	45	19	28.8	7.5	< 3 m	25 stopped diet 10 years+	Ed
Dawson et al. (2011)	on diet	British	21	19	19	0.9	640	103	21	–	–	–	–	until adolescence	
	off diet	British	56	31	31	0.6	1461	185	56	–	–	–	–	until adolescence	
De Felice et al. (2018)	Same PKU group different tasks	British	37	25	27.3	8.1	751	319	39	29		8.4	birth	7 not on diet at time of testing	Ed
Palermo et al. (2017)					27.5	7.3	720	343			27.6	7.4			
Palermo et al. (2020)					27.4	8.3	670	319				8.5			
Jahia et al. (2016)	Same PKU group different tasks	Dutch	56	31	28.9	6.5	657	336	53	39	26	5.8	birth	–	–
Jahia et al. (2017)					27.7	6	655	342				5.9			
Moyle et al. (2006)		Australian	9	8	26	3.7	1000	775	9	8	25	3.7	–	–	–
Moyle et al. (2007b)		Australian	12	10	28.5	11.4	–	–	12	–	29.2	11	birth	not on diet at testing	Ed
Nardecchia et al. (2015)	Only adult results (2nd assessment)	Italian	14	12	24.97	1.6	–	–	14	12	23.7	2.6	< 3 m		–
Pietz et al. (1998)		German	57	33	23.6	3.4	1085	303	40	24	23	–	< 3 m	–	–
Ris et al. (1994)		USA	25	12		–	1332	–	15	9	23	–	< 1 m		SES

Table 5 (*continued*)

Study	Subgroup/condition	Nation	N	N F	Age Mean	Age SD	Current phe in μmol/L Mean	SD	N	N F	Age Mean	Age SD	Age of diet treatment		Controls matched for
Romani et al. (2020)	Only Italian group	Italian	19	8	22	4.1	1042	428	19	8	24.7	3.4	birth	8 stopped diet before adolescence –	
Schmidt et al. (1994)	Baseline data time1	German	14	8	25.4	–	1417	345	20	11	20.7	–	birth	–	–
Smith et al. (1996)	Low Phe	Canada	11	4	20.5	5.5	566	217	22	9	24.1	4.5	–	–	IQ
	High Phe	Canada	11	4	25.4	5.5	1545	182					–	–	IQ
Sundermann et al. (2011)	G1 placebo 1st	German	8	8	22.4	5.3	1210	320	15	15	32.1	6.4	birth	until adolescence	–
	G2 placebo 2nd	German	9	9	31	4.7	1140	310					birth	until adolescence	–
Sundermann et al. (2020)		German	16	–	31	4.4	926.2	430	17	–	24.7	3.9	< 3 m	Not specified	
Ullrich et al., (1996)		German	8	–	25.8	3.8	1253	176	8	–	20.1	3.9	–	On average on diet until 10 years old	IQ
Weglage et al. (2013)	baseline data	German	57	37	31.7	5.9	810	–	46	24	34.2	11	< 3 m		SES
Pilotto et al. (2021)		German	19	12	41	3.75	890	459	25	15	34	2.5	< 1 m	10 on diet/9 on diet till 18	Ed
Average			**31.2**	**19**	**26.5**	**5.2**	**954**	**307**	**29.9**	**18.1**	**26.4**	**6.2**			
SD			*30.4*	*17.4*	*4.5*	*2.6*	*289*	*132*	*17.6*	*11.2*	*4.6*	*2.7*			
sum				**757**						**676**					

Appendix 2

See in Fig. 5.
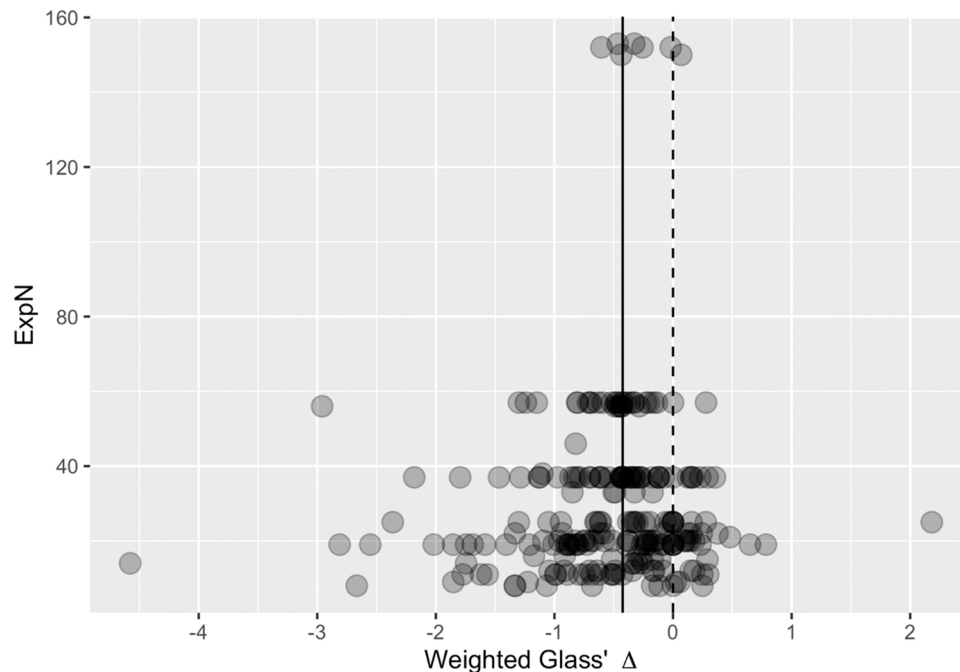
Fig. 5. Funnel plot of individual effect sizes (Glass' Δ) plotted against the size of the PKU sample. The solid line marks the overall estimate of the effect size. The dotted line at 0 marks the "no difference" point between PKU and controls. Most studies show a negative effect. Balanced data should have similar numbers of points to the left and right of the solid line.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.neubiorev.2022.104925.

References

Aitkenhead, L., Krishna, G., Ellerton, C., Moinuddin, Md, Matcham, E., Shiel, L., Murphy, E., 2021. Long-term cognitive and psychosocial outcomes in adults with phenylketonuria. J. Inherit. Metab. Dis. 44 (6), 1353–1368. https://doi.org/10.1002/jimd.12413.

Albrecht, J., Garbade, S.F., Burgard, P., 2009. Neuropsychological speed tests and blood phenylalanine levels in patients with phenylketonuria: a meta analysis. Neurosci. Biobehav. Rev. 33, 414–421.

Anderson, P.J., Leuzzi, V., 2010. White matter pathology in phenylketonuria. Mol. Genet. Metab. 99, S3–S9.

Antenor-Dorsey, J.A.V., Hershey, T., Rutlin, J., Shimony, J.S., McKinstry, R.C., Grange, D.K., White, D.A., 2013. White matter integrity and executive abilities in individuals with phenylketonuria. Mol. Genet. Metab. 109 (2), 125–131.

Backman, L., Lindenberger, U., Li, S.C., Nyberg, L., 2010. Linking cognitive aging to alterations in dopamine neurotransmitter functioning: recent data and future avenues. Neurosci. Biobehav. Rev. 34 (5), 670–677. https://doi.org/10.1016/j.neubiorev.2009.12.008.

(*) Bartus, A., Palasti, F., Juhasz, E., Kiss, E., Simonova, E., Sumanszki, C., et al., 2018. The influence of blood phenylalanine levels on neurocognitive function in adult PKU patients. Metab. Brain Dis. 1–7. https://doi.org/10.1007/s11011-018-0267-6.

Bilder, D.A., Noel, J.K., Baker, E.R., Irish, W., Chen, Y., Merilainen, M.J., et al., 2016. Systematic review and meta-analysis of neuropsychiatric symptoms and executive functioning in adults with phenylketonuria. Dev. Neuropsychol. 41, 245–260.

Bisiacchi, P., Tarantino, V., Mento, G., Burlina, A., 2018. Subclinical executive function impairment in children with asymptomatic, treated phenylketonuria: A comparison with children with immunodeficiency virus. Cognitive Neuropsychology. 35 (3–4), 200–208.

Boot, E., Hollak, C.E.M., Huijbregts's, S.C.J., Jahja, R., van Vliet, D., Nederveen, A.J., Nieman, D.H., Bosch, A.M., Bour, L.J., Bakermans, A.J., et al., 2017. Cerebral dopamine deficiency, plasma monoamine alterations and neurocognitive deficits in adults with phenylketonuria. Psychol. Med. 47, 2854–2865.

Borenstein, M., Hedges, L.V., Higgins, J.P.T., Rothstein, H.R., 2009. Introduction to Meta-Analysis. John Wiley & Sons,, Chichester, U.K.

Bramer, W.M., Rethlefsen, M.L., Kleijnen, J., Franco, O.H., 2017. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. Systematic Reviews. 6 (245).

Brumm, V.L., Azen, C., Moat,s, R.A., Stern, A.M., Broomand, C., Nelson, M.D., et al., 2004. Neuropsychological outcome of subjects participating in the PKU adult collaborative study: a preliminary review. J. Inherit. Metab. Disord. 27, 549–566.

Brunner, R.L., Berry, H.K., 1987. Phenylketonuria and sustained attention: the continuous performance test. Int. J. Clin. Neuropsychol. 9, 68–70.

Bucur, B., Madden, D.J., 2010. Effects of adult age and blood pressure on executive function and speed of processing. Exp. Aging Res. 36 (2), 153–168.

(*) Burgard, P., Rey, F., Rupp, A., Abadie, V., Rey, J., 1997. Neuropsychologic functions of early treated patients with phenylketonuria, on and off diet: results of a cross-national and cross-sectional study. Pediatr. Res. 41, 368–374.

Burlina, A.P., Lachmann, R.H., Manara, R., Cazzorla, C., Celato, A., van Spronsen, F.J., Burlina, A., 2019. The neurological and psychological phenotype of adult patients with early-treated phenylketonuria: a systematic review. J. Inherit. Metab. Dis. 104, S19–S25.

Burton, B., Grant, M., Feigenbaum, A., Singh, R., Hendrene, R., et al., 2015. A randomized, placebo-controlled, double-blind study of sapropterin to treat ADHD symptoms and executive function impairment in children and adults with sapropterin-responsive phenylketonuria. Molecular Genetics and Metabolism 114, 415–424.

Canton, M., Gall, Le, Feillet, D., Bonnemains, C, F., Roy, A., 2019. Neuropsychological profile of children with early and continuously treated phenylketonuria: systematic review and future approaches. J. Int. Neuropsychol. Soc. 25, 624–643.

(*) Channon, S., German, E., Cassina, C., Lee, P., 2004. Executive functioning, memory, and learning in phenylketonuria. Neuropsychology 18 (4), 613–620.

(*) Channon, S., German, E., Cassina, C., Lee, P., 2005. Executive functioning and speed of processing in phenylketonuria. Neuropsychology 19 (5), 679–686.

(*) Channon, S., Goodman, G., Zlotowitz, S., Mockler, C., Lee, P.J., 2007. Effects of dietary management of phenylketonuria on long-term cognitive outcome. Arch. Dis. Child. 92, 213–218.

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 (S22–S32).

Clarke, J.T., Gates, R.D., Hogan, S.E., Barrett, M., MacDonald, G.W., 1987. Neuropsychological studies on adolescents with phenylketonuria returned to phenylalanine-restricted diets. Am. J. Ment. Retard. 92, 255–262.

(*) Dawson, C., Murphy, E., Maritz, C., Lachmann, R.H., 2011. Dietary treatment of phenylketonuria: the effect of phenylalanine on reaction time. J. Inherit. Metab. Dis. 34 (2), 449–454.

(*) De Felice, S., Romani, C., Geberhiwot, T., MacDonald, A., Palermo, L., 2018. Language processing and executive functions in early treated adults with phenylketonuria (PKU). Cogn. Neuropsychol. 35 (3–4), 148–170.

de Groot, M.J., Hoeksma, M., Blau, N., Reijngoud, D.J., van Spronsen, F.J., 2010. Pathogenesis of cognitive dysfunction in phenylketonuria: review of hypotheses. Mol. Genet. Metab. 99, S86–S89.

Delaloye, C., Moy, G., Baudois, S., De Bilbao, F., Remund, C.D., Hofer, F., Giannakopoulos, P., 2009. The contribution of aging to the understanding of the dimensionality of executive functions. Arch. Gerontol. Geriatr. 49 (1), e51–e59.

DeRoche, K., Welsh, M., 2008. Twenty-five years of research on neurocognitive outcomes in early-treated phenylketonuria: intelligence and executive function. Dev. Neuropsychol. 33 (4), 474–504. https://doi.org/10.1080/87565640802101482.

Di Betta, A.-M., Romani, C., 2006. Lexical learning and dysgraphia in a group of adults with developmental dyslexia. Cogn. Neuropsychol. 23 (3), 376–400.

Dijkstra, A.M., van Vliet, N., van Vliet, D., Romani, C., Huijbregts, S.C.J., van der Goot, E., Hovens, I.B., van der Zee, E.A., Kema, I.P., M. Rebecca Heiner-Fokkema, M.R., van Spronsen, F., 2021. Correlations of blood and brain biochemistry in phenylketonuria: Results from the Pah-enu2 PKU mouse. Mol. Genet. Metab. 134 (3), 250–256. https://doi.org/10.1016/j.ymgme.2021.09.004.

Donlon, J., Sarkissian, C., Levy, H.L., Scriver, C.R., 2015. Hyperphenylalaninemia: phenylalanine hydroxylase deficiency. Scriver'S. Online Metab. Mol. Bases Inherit. Disord. https://doi.org/10.1036/ommbid.97.

Eckert, M.A., Keren, N.I., Roberts, D.R., Calhoun, V.D., Harris, K.C., 2010. Age-related changes in processing speed: unique contributions of cerebellar and prefrontal cortex. Front. Hum. Neurosci. 4 (10) https://doi.org/10.3389/neuro.09.010.2010.

Egger, M., Davey Smith, G., Schneider, M., Minder, C., 1997. Bias in meta-analysis detected by a simple, graphical test. Br. Med. J. 315 (7109), 629–634.

Evans, A.H., Costa, D.C., Gacinovic, S., Katzenschlager, R., O'Sullivan, J.D., Heales, S., Lees, A.J., 2004. L-Dopa-responsive Parkinson's syndrome in association with phenylketonuria: in vivo dopamine transporter and D2 receptor findings. Mov. Disord.: Off. J. Mov. Disord. Soc. 19, 1232–1236. https://doi.org/10.1002/mds.v19:10.

Feldmann, R., Osterloh, J., Onon, S., Fromm, J., Rutsch, F., Weglage, J., 2019. Neurocognitive functioning in adults with phenylketonuria: report of a 10-year follow-up. Mol. Genet. Metab. 126, 246–249.

Folk, C.L., Lincourt, A.E., 1996. The effects of age on guided conjunction search. Exp. Aging Res. 22 (1), 99–118.

Fonnesbeck, C.J., McPheeters, M.L., Krishnaswami, S., Lindegren, M.L., Reimschisel, T., 2013. Estimating the probability of IQ impairment from blood phenylalanine for phenylketonuria patients: a hierarchical meta-analysis. J. Inherit. Metab. Dis. 36, 757–766.

Foster, J.K., Behrmann, M., Stuss, D.T., 1995. Aging and visual search: generalized cognitive slowing or selective deficit in attention? Aging, Neuropsychol., Cogn. 2 (4), 279–299.

Gonzalez, M.J., Gassio, R., Artuch, R., Campistol, J., 2016. Impaired neurotransmission in early-treated phenylketonuria patients. Semin. Paediatr. Neurol. 23 (4), 332e40.

Griffiths, P., Paterson, L., Harvie, A., 1995. Neuropsychological effects of subsequent exposure to phenylalanine in adolescents and young adults with early-treated phenylketonuria. J. Intellect. Disabil. Res. 39 (Pt 5), 365–372. https://doi.org/10.1111/j.1365-2788.

Grueter, B.E., Schulz, U.G., 2012. Age-related cerebral white matter disease (leukoaraiosis): a review. Postgrad. Med. J. 88 (1036), 79–87.

Hale, S., Myerson, J., 1996. Experimental evidence for differential slowing in the lexical and nonlexical domains. Aging, Neuropsychol., Cogn. 3, 154–165.

Hamilton, L.D., Thomas, E., Almuklass, A.M., Enoka, R.M., 2017. A framework for identifying the adaptations responsible for differences in pegboard times between middle-aged and older adults. Exp. Gerontol. 97, 9–16. https://doi.org/10.1016/j.exger.2017.07.003.

Hofman, D.L., Champ, C.L., Lawton, C.L., Henderson, M., Dye, L.A., 2018. systematic review of cognitive functioning in early treated adults with phenylketonuria. Orphanet J. Rare Diseases 13, 150.

Hommel, B., Li, K.Z., Li, S.C., 2004. Visual search across the life span. Dev. Psychol. 40, 545–558.

Huijbregts, S., de Sonneville, L., van Spronsen, F.A., Licht, R., Sergeant, J., 2002. The neuropsychological profile of early and continuously treated phenylketonuria: orienting, vigilance, and maintenance versus manipulation-functions of working memory. Neurosci. Biobehav. Rev. 26, 697–712.

Ialongo, C., 2016. Understanding the effect size and its measures. Biochem. Med. 26 (2), 150–163.

Jahja, R., Huijbregts, S.C.J., de Sonneville, L.M.J., van der Meere, J.J., van Spronsen, F.J., 2014. Neurocognitive evidence for revision of treatment targets and guidelines for phenylketonuria. J. Paediatr. 164 (4), 895–899.

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

(*) Jahja, R., van Spronsen, F.J., de Sonneville, L.M.J., van der Meere, J.J., Bosch, A.M., Hollak, C.E.M., Huijbregts, S.C.J., 2016. Social-cognitive functioning and social skills in patients with early treated phenylketonuria: a PKU-COBESO study. J. Inherit. Metab. Dis. 39, 355–362.

Janos, A.L., Grange, D.K., Steiner, R.D., White, D.A., 2012. Processing speed and executive abilities in children with Pheylketonuria. Neuropsychology 26, 735–743.

Janzen, D., Nguyen, M., 2010. Beyond executive function: non-executive cognitive abilities in individuals with PKU. Mol. Genet. Metab. 99, S47–S51.

Koch, R., Burton, B., Hoganson, G., Peterson, R., Rhead, W., Rouse, B., Scott, R., Wolff, J., Stern, A.M., Güttler, F., Nelson, M., de la Cruz, F., ColdwellJ Erbe, R., Geraghty, M.

T., Shear, C., Thomas, J., Azen, C., 2002. Phenylketonuria in adulthood: a collaborative study. J. Inherit. Metab. Dis. 25 (5), 333–346.

Kofler, M.J., Rapport, M.D., Sarver, D.E., Raiker, J.S., Orban, S.A., Friedman, L.M., Kolomeyer, E.G., 2013. Reaction time variability in ADHD: a meta-analytic review of 319 studies. Clin. Psychol. Rev. 33, 795–811.

Lawrence, B., Myerson, J., Hale, S., 1998. Differential decline of verbal and visuospatial processing speed across the adult life span. Aging, Neuropsychol., Cogn. 5, 129–146.

Leuzzi, V., Chiarotti, F., Nardecchia, F., van Vliet, D., van Spronsen, F.J., 2020. Predictability and inconsistencies of cognitive outcome in patients with phenylketonuria and personalised therapy: the challenge for the future guidelines. J. Med. Genet. 57, 145–150.

Lezak, M.D., Howieson, D.B., Bigler, E.D., Tranel, D., 2012. Neuropsychological Assessment, fifth ed. Oxford University Press,, New York.

Luciana, M., Sullivan, J., Nelson, C.A., 2001. Associations between phenylalanine-to-tyrosine ratios and performance on tests of neuropsychological function in adolescents treated early and continuously for phenylketonuria. Child Dev. 72 (6), 1637–1652. https://doi.org/10.1111/1467-8624.00370.

Luciana, M., Hanson, K.L., Whitley, C.B., 2004. A preliminary report on dopamine system reactivity in PKU: Acute effects of haloperidol on neuropsychological, physiological, and neuroendocrine functions. Psychopharmacology, 175, 18–25.

MacDonald, A., Van Wegberg, A., Ahring, K., Beblo, S., Bélanger-Quintana, A., Burlina, A., Giżewska, M., 2020. PKU dietary handbook to accompany PKU guidelines. Orphanet J. Rare Dis. 15 (1), 171.

Mastrangelo, M., Chiarotti, F., Berillo, L., Caputi, C., Carducci, C., Di Biasi, C., Manti, F., Nardecchia, F., Leuzzi, V., 2015. The outcome of white matter abnormalities in early treated phenylketonuric patients: a retrospective longitudinal long-term study. Mol. Genet. Metab. 116 (3), 171–177. https://doi.org/10.1016/j.ymgme.2015.08.005.

(*) Moyle, J.J., Fox, A.M., Bynevelt, M., Arthur, M., Burnett, J.R., 2006. Event-related potentials elicited during a visual go-Nogo task in adults with phenylketonuria. Clin. Neurophysiol. 117, 2154–2160.

Moyle, J.J., Fox, A.M., Arthur, M., Bynevelt, M., Burnett, J.R., 2007a. Meta-analysis of neuropsychological symptoms of adolescents and adults with PKU. Neuropsychol. Rev. 17, 91–101.

(*) Moyle, J.J., Fox, A.M., Bynevelt, M., Arthur, M., Burnett, J.R., 2007b. A neuropsychological profile of off-diet adults with PKU. J. Clin. Exp. Neuropsychol. 29, 436–441.

Mulder, M.J., Bos, D., Weusten, J.M.H., van Belle, J., van Dijk, S.C., Simen, P., van Engeland, H., Durston, S., 2010. Basic impairments in regulating the speed-accuracytradeoff predict symptoms of attention-deficit/hyperactivity disorder. Biol. Psychiatry 68, 1114–1119.

(*) Nardecchia, F., Manti, F., Chiarotti, F., Carducci, C., Leuzzi, V., 2015. Neurocognitive and neuroimaging outcome of early treated young adult PKU patients: a longitudinal study. Mol. Genet. Metab. 115, 84–90.

(*) Palermo, L., Geberhiwot, T., MacDonald, A., Limbac,k, E., Hall, S.K., Romani, C., 2017. Cognitive outcomes in early treated adults with phenylketonuria (PKU): a comprehensive picture across domains. Neuropsychology 31 (3), 255–267.

(*) Palermo, L., MacDonald, A., Limback, E., Robertson, L., Howe, S., Geberhiwot, T., Romani, C., 2020. Emotional health in early-treated adults with phenylketonuria (PKU): relationship with cognitive abilities and blood phenylalanine. J. Clin. Exp. Neuropsychol. 42 (2), 142–159. https://doi.org/10.1080/13803395.2019.1696753.

Park, D.C., Lautenschlager, G., Hedden, T., Davidson, N.S., Smith, A.D., Smith, P.K., 2002. Models of visuospatial and verbal memory across the adult life span. Psychol. Aging 17 (2), 299.

(*) Pietz, J., Dunckelmann, R., Rupp, A., Rating, D., Meinck, H.M., Schmidt, H., et al., 1998. Neurological outcome in adult patients with early-treated phenylketonuria. Eur. J. Pediatr. 157, 824–830.

Pilotto, A., Zipser, C.M., Leks, E., Haas, D., Gramer, G., Freisinger, P., Berg, D., 2021. Phenylalanine effects on brain function in adult phenylketonuria. Neurology 96, e399–e411. https://doi.org/10.1212/WNL.0000000000011088.

Pustejovsky, J.E. , 2016. Alternative formulas for the standardized mean difference. https://www.jepusto.com/alternative-formulas-for-the-smd/.

Report of Medical Research Council Working Party on Pheylketonuria, 1993. Recommendations on the dietary management of Phenylketonuria. Arch. Dis. Child. 68, 426–427.

(*) Ris, M.D., Williams, S.E., Hunt, M.M., Berry, H.K., Leslie, N., 1994. Early-treated phenylketonuria: adult neuropsychologic outcome. J. Pediatr. 124, 388–392.

Romani, C., Tsouknida, E., Olson, A., 2015. Encoding order and developmental dyslexia: a family of skills predicting different orthographic components. Q. J. Exp. Psychol. 68 (1), 99–128.

Romani, C., Betta, Di, Tsouknida, E A.M., Olson, A., 2008. Lexical and nonlexical processing in developmental dyslexia: a case for different resources and different impairments. Cogn. Neuropsychol. 25 (6), 798–830.

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. Cogn. Neuropsychol. 35 (3–4), 171–198. https://doi.org/10.1080/02643294.2017.1320278.

Romani, C., Palermo, L., Macdonald, A., Limback, E., Hal, l S.K., Geberhiwot, T., 2017. The Impact of phenylalanine levels on cognitive outcomes in adults with phenylketonuria: effects across tasks and developmental stages. *Neuropsycholog*y 31, 242–254.

Romani, C., Manti, F., Nardecchia, F., Valentini, F., Fallarino, N., Carducci, C., De Leo, S., Macdonald, A., Palermo, L., Leuzzi, V., 2019. Adult cognitive outcomes in phenylketonuria: explaining causes of variability beyond average Phe levels. Orphanet J. Rare Dis. 14, 1–17.

(*) Romani, C., Manti, F., Nardecchia, F., Valentini, F., Fallarino, N., Carducci, C., De Leio, S., MacDonald, A., Palermo, L., Leuzzi, V., 2020. Cognitive outcomes and

relationships with phenylalanine in phenylketonuria: a comparison between italian and english adult samples. Nutrients 12, 3033. https://doi.org/10.3390/nu12103033.

Salthouse, T.A., 2004. Localizing age-related individual differences in a hierarchical structure. Intelligence 32 (6), 541–561.

(*) Schmidt, E., Rupp, A., Burgard, P., Pietz, J., Weglage, J., Sonneville, L.D., 1994. Sustained attention in adult phenylketonuria: the influence of the concurrent phenylalanine-blood-level. J. Clin. Exp. Neuropsychol. 16 (5), 681–688. https://doi.org/10.1080/01688639408402681Smith.

Shafto, M., Tyler, L., 2014. Language in the aging brain: the network dynamics of cognitive decline and preservation. Science 346, 583–587.

Shoraka, H.R., Haghdoost, A.A., Baneshi, M.R., Bagherinezhad, Z., Zolala, F., 2020. Global prevalence of classic phenylketonuria based on neonatal screening program data: systematic review and meta-analysis. Clin. Exp. Pediatr. 63 (2), 34–43. https://doi.org/10.3345/kjp.2019.00465.

(*) Smith, M.L., Klim, P., Mallozzi, E., Hanley, W.B., 1996. A test of the frontal-specificity hypothesis in the cognitive performance of adults with phenylketonuria. Dev. Neuropsychol. 12, 327–341.

van Spronsen, F.J., Huijbregts, S.C.J., Bosch, A.M., Leuzzi, V., 2011. Cognitive, neurophysiological, neurological and psychosocial outcomes in early-treated PKU-patients: a start toward standardized outcome measurement across development. Mol. Genet. Metab. 104, S45–S51.

van Spronsen, F.J., van Wegberg, A.M, Ahring, K., Bélanger-Quintana, A., Blau, N., Bosch, A.M., et al., 2017. Key European guidelines for the diagnosis and management of patients with phenylketonuria. Lancet Diabetes Endocrinol. 5 (9), 743–756. Doi.org/10.1016/S2213-8587(16)30320-5.

van Spronsen, F.J., Balu, N., Harding, C., Burlina, A., Longo, N., Bosch, A.M., 2021. Phenylketonuria. Nat. Rev. Dis. Prim. 7, 36.

Stemerdink, N.B.A., Van Der Molen, M.W., Kalverboer, A.F., Van Der Meere, J.J., Huisman, J., de Jong, L.W., et al., 1999. Prefrontal dysfunction in early and continuously treated phenylketonuria. Dev. Neuropsychol. 16, 29–57.

Stevenson, M., McNaughton, N., 2013. A comparison of phenylketonuria with attention deficit hyperactivity disorder: Do markedly different aetiologies deliver common phenotypes? Brain Research Bulletin 99, 63–83.

(*) Sundermann, B., Garde, S., Dehghan, Nayyeri, M., et al., 2020. Approaching altered inhibitory control in phenylketonuria. Eur. J. Neurosci. 52, 3951–3962. https://doi.org/10.1111/ejn.147.

(*) Sundermann, B., Pfleiderer, B., Möller, H.E., Schwindt, W., Weglage, J., Lepsien, J., Feldmann, R., 2011. Tackling frontal lobe–related functions in PKU through functional brain imaging: a Stroop task in adult patients. J. Inherit. Metab. Dis. 34 (3), 711–721. https://doi.org/10.1007/s10545-011-9318-4.

Thomas, L., Olson, A., Romani, C. (submitted for publication). The impact of metabolic control on cognition, neurophysiology, and well-being in PKU: A systematic review and meta-analysis of the within-participant literature.

Trefz, F., Maillot, F., Motzfeldt, K., Schwarz, M., 2011. Adult phenylketonuria outcome and management. Mol. Genet. Metab. 104, S26–S30.

Tufekcioglu, Z., Cakar, A., Bilgic, B., Hanagasi, H., Gurvit, H., Emre, M., 2016. Adult-onset phenylketonuria with rapidly progressive dementia and parkinsonism. Neurocase 22 (3), 273–275. https://doi.org/10.1080/13554794.2016.1142567.

(*) Ullrich, K., Weglage, J., Pietsch, M., Fünders, B., Oberwittler, C., Von Eckardstein, H., Colombo, J.P., 1996. Effect of L-DOPA on visual evoked potentials and neuropsychological tests in adult phenylketonuria patients. Eur. J. Pediatr. 155 (S1), S74–S77. https://doi.org/10.1007/pl00014256.

Vardy, E.R.L.C., MacDonald, A., Ford, S., Hofman, D.L., 2020. Phenylketonuria, co-morbidity, and ageing: a review. J. Inherit. Metab. Dis. 43, 167–178. https://doi.org/10.1002/jimd.12186.

Velema, M., Boot, E., Engelen, M., Hollk, C., 2015. Parkinsonism in phenylketonuria: a consequence of dopamine depletion? J. Inherit. Metablic Dis. Rep. 20, 35–38. https://doi.org/10.1007/8904_2014_386.

Verhaeghen, P., De Meersman, L., 1998. Aging and the stroop effect: a meta-analysis. Psychol. Aging 13 (1), 120.

van Vliet, D., Wegberg, A.M.J., Ahring, K., Bik-Multanowski, M., Blau, N., Bulut, F.D., Spronsen, F.J., 2018. Can untreated PKU patients escape from intellectual disability? A systematic review. Orphanet J. Rare Dis. 13, 149. https://doi.org/10.1186/s13023-018-0890-7.

Vockley, J., Andersson, H.C., Antshel, K.M., Braverman, N.E., Burton, B.K., Frazier, D.M., et al., 2014. Phenylalanine hydroxylase deficiency: diagnosis and management guideline. Genet. Med. 16 (2), 188–200, 10.1038/ gim.2013.157.

Volkow, N.D., Gur, R.C., Wang, G.J., Fowler, J.S., Moberg, P.J., Ding, Y.S., Logan, J., 1998. Association between decline in brain dopamine activity with age and cognitive and motor impairment in healthy individuals. Am. J. Psychiatry 155 (3), 344–349.

Waisbren, S.E., Noel, K., Fahrbach, K., Cella, C., Frame, D., Dorenbaum, A., Levy, H., 2007. Phenylalanine blood levels and clinical outcomes in phenylketonuria: a systematic literature review and metaanalysis. Mol. Genet. Metab. 92 (1–2), 63–70.

van Wegberg, A.M.J., MacDonald, A., Ahring, K., Bélanger-Quintana, A., Blau, N., Bosch, A.M., et al., 2017. The complete European guidelines on phenylketonuria: diagnosis and treatment. Orphanet J. Rare Dis. 12 (1), 162. https://doi.org/10.1186/s13023-017-0685-2.

Weglage, J., Pietsch, M., Fünders, B., Koch, H., Ullrich, K., 1995. Neurological findings in early treated phenylketonuria. Acta Paediatr. 84 (4), 412–415.

(*) Weglage, J., Fromm, J., Van Teeffelen-heithoff, A., Möller, H.E., Koletzko, B., Marquardt, T., 2013. Neurocognitive functioning in adults with phenylketonuria: results of a long term study. Mol. Genet. Metab. 110 (Suppl 44–8).

Wilding, J., Pankhania, P., Williams, A., 2007. Effects of speed and accuracy instructions on performance in a visual search task by children with good or poor attention. Br. J. Psychol. 98, 127–139.