Vitamin supplements and mortality in older people

Dear Sir:

Macpherson et al (1) carried out a meta-analysis of multivitamin and multimineral (MVMM) tablet trials and found no effect of MVMMs on average mortality. However, their study may suffer from ecological fallacy. Ecological fallacy means that study-level (group-level) analysis can lead to different conclusions than do corresponding individual-level analyses (2). For this reason, examination of individual-level data is recommended, whenever feasible, to avoid the potential for the ecological fallacy introduced by study-level analyses (2).

Macpherson et al (1) calculated that the average age of the participants in the studies was 62 y. However, ages ranged from 17 to 86 y in the included trials (1). It is probable that the effects of all vitamins and minerals are not identical at the lower and upper ends of such a wide age range. Therefore, pooling diverse trials with young and old people to a single average MVMM effect may camouflage effects of some individual vitamins or minerals, for example, on the oldest people. In the case of vitamin E there is strong empirical evidence of effect modification by age.

In an individual-level analysis of the Alpha-Tocopherol, Beta-Carotene Cancer Prevention (ATBC) Study data, we found that among participants aged 50–62 y at baseline with a dietary vitamin C intake above the median, vitamin E increased mortality by 19% (95% CI: 5%, 35%; based on 1021 deaths). However, among participants aged 66–69 y at baseline with a dietary vitamin C intake above the median, vitamin E decreased mortality by 41% (95% CI: 21%, 56%; based on 195 deaths) (3).

Furthermore, because the follow-up time in the ATBC Study was up to 8 y, the participants became substantially older during the trial so that the baseline age was not a proper way to characterize them over the entire follow-up period. Therefore, the modification of vitamin E effects was also analyzed by using the follow-up age as the time variable (4). Among 10,837 ATBC Study participants who contributed follow-up time past the age of 65 y, the survival curves of the vitamin E and no–vitamin E participants significantly diverged at 71 y. Vitamin E extended life span by ~0.5 y at the upper limit of the follow-up age span (4).

Macpherson et al (1) write that in a meta-regression the estimate of the effect of MVMMs was not associated with the duration of supplementation. In the ATBC Study, the harm from vitamin E in the young participants was restricted to the supplementation period after 3.3 y, indicating that there can be a lag period of several years before the effects of some vitamins appear (3). Macpherson et al used the study-level average durations, which provide a poor basis for analyzing supplementation time–dependent effect modifications. Proper analysis of time-dependent effects requires individual-level data.

It is possible that some vitamins and minerals are beneficial for specific subpopulations. For example, age, sex, smoking, diet, and exercise might modify the effects of some vitamins and minerals, so that some restricted population groups might benefit (and some might be harmed). Such subgroups can be explored by analyzing individual-level data, whereas pooling study-level averages provides no information on relevant narrow subpopulations.

The meta-analysis by Macpherson et al (1) is important in discouraging ordinary middle-aged people from taking MVMMs. Nevertheless, their study should not be interpreted as evidence that none of the vitamins and minerals included in the MVMM tablets have effects on males and females in the age range of 17–86 y. It is possible that some vitamins, such as vitamin E, are useful for restricted groups of older people. Individual-level data analyses are needed for exploring such a possibility.

The author did not declare any conflicts of interest.

Harri Hemilä
Department of Public Health
University of Helsinki
Helsinki, FIN-00014
Finland
E-mail: harri.hemila@helsinki.fi

REFERENCES


Reply to H Hemilä

Dear Sir:

We thank Hemilä for his interest in our article entitled “Multivitamin-multimineral supplementation and mortality: a meta-analysis of randomized controlled trials” (1). Our primary finding was that, across a pooled sample of 91,074 participants, multivitamin-multimineral...
(MVMM) supplementation had no significant effect on the risk of all-cause mortality, mortality due to cancer, or mortality due to cardiovascular disease.

Despite our overall finding, Hemila argues that some vitamins and minerals may be beneficial for specific subpopulations. We concur with his suggestion that variables such as age, sex, and lifestyle factors might modify the effects of some vitamins, such that differential effects may emerge in different subpopulations. However, as pointed out by Hemila, we were unable to perform subanalyses to examine the modifying effect of these different variables given that only trial-level data were available.

If individual-level data were accessible, we could have performed any number of subanalyses. A limitation of this approach is that each subanalysis involves an additional statistical comparison and thus a greater risk of a type I error. Furthermore, subgroup analysis based on post hoc examination of data can lead to erroneous conclusions (2). The findings discussed by Hemila, relating to vitamin E mortality risk across different age groups, still require replication for this reason. To avoid these issues, we used a limited number of prespecified analyses to determine the overall effects of MVMM supplementation in the general population, rather than in specific subpopulations.

Our results were strengthened by the large number of trials included in our analyses, generating a large pooled sample size. Although there are several advantages to undertaking an individual-level data meta-analysis, such an analysis is not always feasible. For example, we excluded 7 relevant trials from our analysis simply because trial-level data were unobtainable. Given the difficulty in obtaining raw data from chief investigators (especially when many of the trials included in our analysis were more than a decade old), undertaking a patient-level meta-analysis would have further diminished the number of trials included in our analysis.

Hemila states that our meta-analysis is “important in discouraging ordinary middle-aged people from taking MVMMs.” We are not sure how this conclusion was derived from our work given that our meta-analysis did not specifically focus on middle-aged adults. Moreover, whereas we found no effect of MVMMs on mortality across adults of all ages, this does not rule out other possible benefits to health or well-being.

Before our investigation, information on the association of MVMM use and mortality had frequently been obtained from observational studies (3). Our meta-analysis showed that, across randomized controlled trials, MVMM supplementation had no effect on mortality (1). Although we acknowledge that vitamins may have different effects in different subpopulations, it was first necessary to investigate the overall effects of MVMM supplementation in the general population. Identifying a harmful effect of MVMM use across all adults would have shown greater implications than identifying a harmful effect in one of many narrow subgroups. As discussed in our meta-analysis, we call for further research into the effects of MVMM use on all aspects of human health (1). This includes examination of MVMM use in specific subpopulations.

MPP is funded by a Menzies Foundation Scholarship in Allied Health Sciences. AP receives funding from Swisse Wellness Pty Ltd for ongoing research projects. HM holds a Postdoctoral Fellowship, which is funded by Swisse Wellness Pty Ltd. There were no other potential conflicts of interest.

Helen Macpherson
Andrew Pipingas
Matthew P Pase

Centre for Human Psychopharmacology
Swinburne University of Technology
Advanced Technology Building
427–451 Burwood Road
Hawthorn, Victoria 3122
Australia

E-mail: matthewpase@gmail.com

REFERENCES


Limitations to the use of plasma osmolality as a hydration biomarker

Dear Sir:

In some laboratories, plasma osmolality (P_{osm}) is used as the gold standard for detecting dehydration (1), without consideration of its limitations; however, published data dispute this technique (2, 3), which prompts us to write in response to the recent article by Cheuvront et al (4) with regard to quantitative dehydration assessment. This article correctly states that P_{osm} is the key regulated variable in fluid balance, which means that P_{osm} is constantly regulated toward a central set point as the kidneys modify urine concentration and water excretion in response to diet and daily activities. We believe that this controlled regulation limits the efficacy of P_{osm} as an index of hydration change in many experimental designs. This article (4) also states that the "criticisms for adopting P_{osm} as a gold standard for dehydration assessment are minimal" (p 460). We disagree and write to describe several limitations to the use of P_{osm} as a gold standard for dehydration.

First, individuals who lose a large amount of body water (reported as % body mass loss relative to a beginning euhydrated state) may exhibit a decreased P_{osm}, contrary to anticipated hemodilution. For example, a summary of 2 studies (5) reported that the P_{osm} of 6 individuals (out of 39) decreased after they lost 3–8% of body mass. In a different study, men and women who consumed a 500-mL bolus of fluid acutely exhibited an increased P_{osm}, contrary to anticipated hemodilution (1); that is, after 90 min of rest, 4 of 30 P_{osm} measurements increased. These values show that P_{osm} may not reflect widely accepted physiologic principles, and that variance of P_{osm} measurements may be large.

Second, evidence suggests that P_{osm} changes are time- and protocol-specific. Unpublished observations (CX Muñoz, EC Johnson, JK DeMartini, et al, 2012) show that dehydration equivalent to 2% of body mass resulted in P_{osm} changes that were twice as large during mild cycling exercise (2.3 h; ΔP_{osm} of 9 mOsm/kg) compared with a passive exposure (5.0 h; ΔP_{osm} of 4 mOsm/kg); participants consumed no water during either trial in
a 36°C environment. It is likely that this difference occurred because exercise increased intracellular osmolality (6) and increased extracellular fluid tonicity, causing water to move into muscle tissue.

Third, Kenney et al (3) reported that mean (±SE) Posm values in 7 resting, euhydrated young male subjects decreased from 281 ± 3 at baseline to 276 ± 2 mOsm/kg at 60 min after they had consumed 1.9 L of water. However, the mean Posm value returned to baseline (282 ± 2 mOsm/kg) at 90 min postingestion. These findings challenge our understanding of the interactions between intracellular-extracellular fluid shifts (6) and renal compensatory mechanisms; they also suggest that further research into the time course of acute Posm changes is warranted.

Fourth, 2 recent publications (7, 8) showed that a single Posm or serum osmolality measurement was a poor predictor of changes in hydration status when a single, fasted morning blood sample is collected. The former article (7) involved modified fluid intake in habitually low-volume drinkers and habitually high-volume drinkers, with the outcome that Posm was constant across days in men and women, whereas urinary biomarkers reflected modified water consumption. The latter publication (8) showed that serum osmolality was a poor predictor (r² = 0.01) of 24-h water retention-clearance by the kidneys. Furthermore, the NHANES (1988–1994) reported that serum osmolality values were constant across a wide range of fluid intakes (9). Men exhibited similar mean Posm values (range: 279–281 mOsm/kg) regardless of total daily fluid intake, which ranged from 1.7 to 7.9 L; women exhibited similar Posm values (range: 276–278 mOsm/kg) across a total daily fluid intake range of 1.3–6.1 L. These studies argue that Posm is not appropriate in clinical settings, in which a single blood sample is collected during an office visit.

Furthermore, Cheuvront et al (4) recommended that a Posm value of 301 ± 5 mOsm/kg be used clinically as the threshold of dehydration (p = 0.01) of 24-h water retention-clearance by the kidneys. The latter publication (8) showed that serum osmolality was a poor predictor (r² = 0.01) of 24-h water retention-clearance by the kidneys. Furthermore, the NHANES (1988–1994) reported that serum osmolality values were constant across a wide range of fluid intakes (9). Men exhibited similar mean Posm values (range: 279–281 mOsm/kg) regardless of total daily fluid intake, which ranged from 1.7 to 7.9 L; women exhibited similar Posm values (range: 276–278 mOsm/kg) across a total daily fluid intake range of 1.3–6.1 L. These studies argue that Posm is not appropriate in clinical settings, in which a single blood sample is collected during an office visit.

Finally, serum samples contain numerous substances (eg, sodium, chloride, potassium, bicarbonate, urea, glucose) that constitute 95% of total osmolality. Even though they are found in small amounts (4–5%), proteins influence total osmolality considerably. Thus, the water content in a serum sample is less per unit volume than in a calibration solution, and to obtain an accurate measurement of osmolality, the empirical value should be mathematically corrected. Furthermore, normal intraindividual differences in serum protein concentration (range: 6.0–8.5 g/dL) and within-individual changes in serum protein concentration induced by factors such as physical training and heat acclimation (11) increase the statistical variance and difficulty of interpreting the meaning of Posm as a hydration index.

We recommend that scientists use Posm as a marker of dehydration cautiously, with careful consideration of experimental protocol (ie, dehydration compared with hyponatremia, exercise compared with rest) and tight control of dietary total osmolar load and fluid volume (2, 8, 10). We recommend that Posm not be used in clinical settings as a gold standard for dehydration assessment (2, 7, 8). The limitations (described above) reflect the dynamic and complex regulation of human fluid-electrolyte balance (2), which does not lend itself to generalizations.

All authors were involved in the writing of this letter, reviewed its content, and approved the final version. None of the authors claimed a conflict of interest.

REFERENCES

Reply to LE Armstrong et al

Dear Sir:

We have great respect for the authors who have expressed interest in our article, and we appreciate the opportunity to reply to their letter; however, we find little convincing evidence for their concerns.

First and foremost we wish to emphasize 2 important points from our article that were left out of the quote taken from page 460 (1). We were very careful in our review to outline why plasma osmolality ($P_{\text{osm}}$) should be considered a gold standard for assessing dehydration, defined as intracellular dehydration (or hypertonic-hypovolemia), and not extracellular dehydration (or isotonic-hypovolemia or volume depletion). We also point out the criticality of considering the dehydration magnitude. With these 2 very important points in mind, the criticisms that we describe as “minimal” on page 460 relate directly to articles that have neglected these important points in their misguided assertions about the limitations of using $P_{\text{osm}}$ for assessing dehydration.

The criticisms of our review on dehydration assessment seem to involve 3 major points: 1) disparate research findings, 2) a $P_{\text{osm}}$ threshold of 301 ± 5 mmol/kg for dehydration, and 3) the contribution of protein to $P_{\text{osm}}$.

Disparate research findings

Six published articles or reports were used when trying to refute our review. Curiously, only 2 of those studies were designed to produce dehydration and only one directly described the potential for using $P_{\text{osm}}$ to quantify dehydration (2). Although the remaining studies referenced do describe the normal, and extremely well-documented, physiologic response to both normal and overconsumption of water (water intake ≥ water losses), when carefully read they do not in any way refute the perspectives presented in our article. As a matter of interpretation, we would also suggest that the composite figure from Sawk et al (2) shows that $P_{\text{osm}}$ responded to dehydration exactly as expected in 33 of 39 volunteers (85%). In a recent study from our laboratory and Senay’s pioneering research have shown that plasma protein can be added by heat exposure as well as lost with dehydration. We acknowledge that some fluxes are already part of the observed variance and diagnostic error. Any acute influence of protein flux due to exercise protein fluxes are already part of the observed variance and diagnostic error. Any acute influence of protein flux due to exercise would also be remedied by allowing proper recovery (1). In other words, the potential for plasma protein to confound the appropriate use of $P_{\text{osm}}$ for assessing dehydration is marginal at best.

In our review article (1), we carefully described the true limitations of using $P_{\text{osm}}$ for dehydration assessment on page 460. The concerns expressed in the letter by Armstrong et al are clearly but curiously misplaced. We must therefore regard the limitations inferred by the title of their letter as “false.”

All of the authors were involved in the writing of this letter, reviewed its content, and approved the final version. The opinions or assertions contained herein are the private views of the authors and should not be construed as official or reflecting the views of the US Army or the US Department of Defense. None of the authors claimed a conflict of interest.

Samuel N Cheuvront
Robert W Kenefick
Nisha Charkoudian

US Army Research Institute of Environmental Medicine
Thermal and Mountain Medicine Division
Kansas Street, Building 42
Natick, MA 01760
E-mail: samuel.n.cheuvront.civ@mail.mil

Michael N Sawka

Georgia Institute of Technology
Atlanta, GA

P$_{\text{osm}}$ threshold of 301 ± 5 mmol/kg

A full appreciation for the genesis of the 301 ± 5-mmol/kg threshold for dehydration requires knowledge of biological variation and diagnostic decision making, which goes well beyond the scope of this letter. We encourage interested readers to seek Cheuvront et al (1, 4, 5) for details. Briefly, the nosological sensitivity of $P_{\text{osm}}$ is modest but superior to all other common body fluids used to assess dehydration. When the variance term for $P_{\text{osm}}$ is properly considered, the range of $P_{\text{osm}}$ values that indicate dehydration (≥2% body mass) agree extremely well with many independently published observations and commonly accepted clinical thresholds for dehydration (4). Change values are better when it is practical to make 2 measures, but here again $P_{\text{osm}}$ does extremely well (4, 5). The Δ$P_{\text{osm}}$ remains sensitive even when water loading is used (urine osmolality: $P_{\text{osm}}$ < 1.5). This practice is often adopted in research where “assurance” of euhydration is desired; however, it is important to recognize that it also decreases the nosological sensitivity of the 301 ± 5-mmol/kg threshold (4). Under said circumstances, a + 5-mmol/kg change in $P_{\text{osm}}$ still affords 80% probability that intracellular dehydration has occurred (4, 5), which is remarkably consistent with the well-taught osmotic change threshold (~2% or +6 mmol/kg) for renal compensation and water acquisition (thirst) (1).
No and low alcohol intake may have differential effects on risk of overall and cause-specific mortality

Dear Sir:

We read with great interest the article by Vergnaud et al (1) on the relation between adherence to the World Cancer Research Fund (WCRF)/American Institute for Cancer Research (AICR) guidelines and risk of death in Europe. This well-crafted, large-scale study conducted in participants in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort offers valuable data regarding the impact of the WCRF/AICR recommendations on reducing total and cause-specific mortality and suggests that the utility of these guidelines may extend beyond the scope of cancer prevention. We are, however, keen on gaining additional understanding of the presented results in their Table 4: namely, the risk of death associated with alcohol consumption.

The authors found that adherence to the WCRF/AICR recommendation for daily alcohol intake (≤2 drinks for men and 1 drink for women) was protective against all-cause mortality in men but not in women. This result was based on a scoring system that operationalized this alcohol-specific guideline into 3 categories of ethanol intake: ≤20, >20 to ≤30, and >30 g/d for men and ≤10, >10 to ≤20, and >20 g/d for women. Among the 257,421 male study participants, the men whose ethanol intake was >20 to ≤30 g/d had a significantly reduced risk of death compared with men whose consumption exceeded 30 g/d (HR: 0.89), as did men who limited their intake to ≤20 g/d compared with the same referent (HR: 0.89). Significant associations between risk of death and the alcoholic drinks component of the WCRF/AICR recommendations were not observed among the 121,443 female study participants.

We are highly curious both to learn whether making the distinction between no and low ethanol intake would alter the results of this analysis and to see the stratification of HRs by cause of death. Whereas it is widely acknowledged that, unlike in cardiovascular disease, the lowest alcohol-related cancer risk is in fact conferred in the absence of alcohol consumption (2), there remains uncertainty regarding whether the protective effect of abstinence on cancer risk translates to survival outcomes. The most current estimate of alcohol-attributable cancer mortality in the United States to our knowledge suggests that alcohol consumption at any level not only increases cancer risk but, more critically, is a major factor behind cancer-related death in men and women (3). Interestingly, the number of alcohol-attributable deaths was highest for female breast cancer in this investigation. A meta-analysis by Bagnardi et al (4) that included 222 articles concerning alcohol consumption and cancer found that light alcohol drinking (≤1 drink/d) was associated with breast cancer death. In contrast and illustrative of the ambiguity related to drinking and cancer mortality, another recent study reported that any alcohol consumption either before or after breast cancer diagnosis had no adverse impact on survival from breast cancer, cardiovascular disease, or other cause, and that moderate consumption may even have a survival benefit (5).

The robust data set of Vergnaud et al presents an opportunity for additional analyses that could shed further light on the advantages or lack thereof of teetotaling in the prevention of cancer or other chronic diseases. As such, we appreciate the authors’ consideration of our request that they both reoperationalize the alcohol-specific WCRF/AICR score such that 0 g/d of ethanol intake is assigned its own category and evaluate alcohol-specific mortality by cause of death and share these results.

Support for this letter was provided by the University of Alabama at Birmingham Cancer Prevention and Control training grant R25 CA047888. The authors had no conflicts of interest to disclose.

Emily Falk Libby
Department of Nutrition Sciences
University of Alabama at Birmingham
VH G017, 1670 University Boulevard
Birmingham, AL 35294-0019
E-mail: elibby@uab.edu

Michelle S Williams
Department of Health Behavior
University of Alabama at Birmingham
Birmingham, AL

Will L Tarver
Department of Health Care Organization and Policy
University of Alabama at Birmingham
Birmingham, AL

Wendy Demark-Wahnefried
Department of Nutrition Sciences
University of Alabama at Birmingham
Birmingham, AL

REFERENCES


Reply to E Falk Libby et al

Dear Sir:

We thank Falk Libby et al for their interest in our article. We acknowledge the need for more detailed analysis of the association between individual components of the World Cancer Research Fund/American Institute for Cancer Research (WCRF/AIRC) score, including alcohol consumption and cause-specific mortality. The association between pattern of lifetime alcohol use and cause of death in the European Prospective Investigation into Cancer and Nutrition (EPIC) study has been addressed in detail by Manuela M Bergmann et al in a manuscript currently under submission. Results cannot be displayed before publication, so we encourage Falk Libby et al to pay attention to the release of this article, which will provide a comprehensive answer to their requests.

None of the authors had a conflict of interest.

Anne-Claire Vergnaud
Dora Romaguera
Petra HM Peeters

Department of Epidemiology and Biostatistics
School of Public Health
Imperial College London
London
United Kingdom


The challenge of complexity and arginine metabolism

Dear Sir:

Chapeau! to Mariotti et al (1) for their attempt to put order to complexity by giving a dimension to arginine fluxes in its metabolism. But, complexity is both a challenge and a burden. An important question relates to the lack of computation of the possible effects that arginine-derived and naturally produced inhibitors of enzymes dealing with arginine metabolism, such as asymmetric-di-methyl-arginine (ADMA), may have on peripheral tissue activity of arginases (2). Do the authors have data on acute effects of arginine ingestion on ADMA? Indeed, it has been reported that long-term ingestion of arginine supplements increases ADMA (3) and inhibition of arginases was efficient in maintaining nitric oxide (NO) production and in preventing damage related to impaired NO production in peripheral tissues (4).

Also, the expression and activity of arginases, and thus their contribution to plasma and urea by red blood cells, were not sufficiently stressed by Mariotti et al in their text or in the supplemental data. Peculiarly, in capillaries red blood cells may dramatically control and blunt arginine concentrations in plasma (5, 6) and this should also be included in a model that focuses on clusters of peripheral needs, even if the said model groups together sums of activities by different compartments. Moreover, habitual dietary arginine intake by controlling arginase expression may rule fluxes of arginine toward availability for protein syntheses or catabolism producing urea. Urea production may become misleading in evaluating adequate nitrogen intake if this is calculated on the basis of urinary urea excretion (7).

The author did not declare any conflicts of interest.
Because the parsimony principle was applied when developing the metabolites of arginine entry into the NO synthase and arginase pathways. In the model, both urea (estimated mainly as “urea synthesis from plasma dietary arginine,” may in part be ascribed to this specific compartment. However, once again, any contribution of red blood cells to the dynamics of postprandial arginine metabolism is both embedded in the data and solved by the model. Of course, our model, like all models, remains a simplification of the system but has proved to be the simplest way to understand the dynamic behavior of the arginine nutritional system.

To answer the direct question posed by Dioguardi with regard to plasma asymmetric-di-methyl-arginine (ADMA), we do have these data on effects after the ingestion of arginine in this setting, and we did not observe that plasma ADMA changed after ingestion (4). Of note, Dioguardi cited a reference that reported an increase in plasma ADMA with long-term arginine supplementation, whereas our results, and those of other groups, indicated no increase in different populations and at different doses (eg, 5–9).

However, from a general standpoint, we agree that little is known about the possible changes in arginine metabolism with regard to NO compared with urea in individuals given large amounts of arginine over the long term, and that changes in arginase activity have emerged as a critical determinant of arginine-NO homeostasis and vascular health (10). Our study was not designed to address these potential long-term effects or to analyze the related underlying possible mechanisms. By using the integrative methodology detailed here, future studies may be able to investigate whether, and to what extent, the key parameters of the system are affected by a long-term increase in arginine intake and should also be able to determine how the system is altered in prepathological conditions (such as with the metabolic syndrome) and in different dietary and nutritional situations.

The authors declared no conflicts of interest.

François Mariotti
Jean-François Huneau

Reply to FS Dioguardi

Dear Sir:

We appreciated the congratulations and comments received from Dioguardi regarding our recently published article, which was the first attempt to delineate the metabolism of dietary arginine, including its bioavailability and utilization for the competitive pathways that are arginase and nitric oxide (NO) synthase (1). The objective of model development was to determine the minimal structure for this nutritional system that could solve the isotopic metabolic data at hand and provide an insight into the key metabolic compartmental structuring that explains how the body deals structurally with arginine intake.

According to the design and process of this modeling study, the effects of any potential changes in arginase or NO synthase activity during the postprandial phase (the potential existence of which was suggested by Dioguardi) are embedded in the isotopic (urea and nitrate) metabolic data and are therefore “computed” in the model predictions for the fluxes of urea and NO production. In the model, both urea and NO production indeed originate from both a plasma compartment and another compartment that aggregates all other possible sources of arginine entry into the NO synthase and arginase pathways. Because the parsimony principle was applied when developing the model, we selected the minimum structure that would include just the main features of the system to reduce model complexity to a manageable level (2, 3), and we did not represent all of the compartments of physiologic interest, such as the red blood cells mentioned by Dioguardi. In other words, a higher-order model with a more detailed structure was not required to analyze the data and the main features of the system. As Dioguardi will understand, this does not mean that red blood cells are not physiologically important with respect to arginase activity, and, as he suggested, peripheral arginase activity, which we estimated mainly as “urea synthesis from plasma dietary arginine,” in part may be ascribed to this specific compartment. However, once again, any contribution of red blood cells to the dynamics of postprandial arginine metabolism is both embedded in the data and solved by the model. Of course, our model, like all models, remains a simplification of the system but has proved to be the simplest way to understand the dynamic behavior of the arginine nutritional system.

REFERENCES


The authors declared no conflicts of interest.
2. Cobelli C, Caumo A. Using what is accessible to measure that which is not: necessity of model of system. Metabolism 1998;47:1009–35.

Describing a taxonomy of cognitive processes for clinical trials assessing cognition

Dear Sir:

Stonehouse et al (1) reported that DHA supplementation improved both memory and reaction time in healthy, young adults. This randomized, placebo-controlled, double-blind clinical trial had many strengths and was, for the most part, technically sound. However, we question the atheoretical manner in which the cognitive tests were grouped into broader cognitive abilities.

In an accompanying editorial, Dangour and Allen (2) questioned the applicability of the cognitive tests used by Stonehouse et al (1). They stated that considerable variability exists in the cognitive tests used between clinical trials and that this significantly hampers comparisons between studies (2). Dangour and Allen proposed that experts in the field should urgently agree on a set of cognitive tests to be used consistently across clinical trials (2). We agree that efforts need to be made to facilitate cross-study comparisons. Yet, consensus as to a standardized set of cognitive tasks is unlikely to be agreed on given the plethora of cognitive tests available and the fact that individual preferences for specific cognitive tests vary greatly. Moreover, because different cognitive tests are suited to different populations and interventions, cognitive tests are often appropriately selected on a case-by-case basis. We propose a less radical solution to aid cross-study comparisons in this area.

Even if researchers cannot agree on the cognitive tests used, consens should be reached on the types of cognitive functions that exist. This would then enable reviewers and readers of published studies to better understand the scope of the tests chosen against the full spectrum of cognitive processes that have been reliably discovered. At present, many clinical trials combine cognitive tests into broader cognitive abilities without justification from existing literature or factor analytic investigation. This appears to be the case in the study by Stonehouse et al (1), whereby cognitive tests are combined into cognitive domains of episodic memory, working memory, attention, and processing speed without explicit justification for this grouping. This significantly hampers comparisons between studies because the cognitive composites are seemingly arbitrary and may never be created again in the same way. We suggest that a standardized and evidence-based approach to grouping cognitive test data will aid comparisons between studies. An empirically supported model for grouping cognitive test data already exists but seems to be ignored by the field of clinical nutrition.

On the basis of 70 y of factor analytical work on cognition, Carroll (3) published a seminal book on human cognitive abilities. Through extensive factor analysis of >460 data sets, his work provides a solid empirical and science-based approach to better understanding the structure of cognition. Such is the significance of this publication to the area of applied psychometrics that it has been compared in importance to Sir Isaac Newton’s Mathematical Principles of Natural Philosophy (4).

---

![Diagram](https://example.com/diagram.png)

**FIGURE 1.** The structure of cognitive abilities based on the work of Carroll (3). Note that the figure is designed to give a snapshot of the model and only some of the 69 narrow cognitive abilities are shown. Adapted with permission from Cambridge University Press.
Carroll’s work provides an empirically verified taxonomy of human cognitive abilities (4). In essence, Carroll (3) outlined a 3-strata hierarchical model of cognitive ability (Figure 1). At the broadest level, stratum 3 consists of a general intelligence factor, which subsumes the following 2 strata. The second stratum includes 8 broad cognitive abilities. Stratum 1 includes a group of 69 narrow, well-defined abilities. All of the cognitive abilities can be classified as belonging to one of the following domains: language, reasoning, memory and learning, visual perception, auditory perception, idea production, cognitive speed, knowledge and achievement, and miscellaneous abilities (3). These cognitive abilities can also be broken down into additional narrow abilities. For example, memory and learning can be further broken down into associative memory, meaningful memory, free recall memory, visual memory, and learning abilities. It is easy to group cognitive test scores into these “true” cognitive abilities because the taxonomy was derived through extensive factor analysis of existing cognitive tests used throughout the past century. Carroll also provides descriptions of each cognitive ability. We therefore suggest that researchers use this taxonomy to group cognitive test score data or at least report how their measures map onto this framework. This will allow significantly better comparison across clinical studies assessing cognition.

The findings reported by Stonehouse et al (1) are of great interest, but as pointed out by others, heterogeneity in cognitive outcomes between studies is significantly limiting advancements in this field. It is surprising that researchers continue to group cognitive tasks into seemingly arbitrary cognitive abilities when a comprehensive evidence-based approach exists. Carroll’s work provides “a common nomenclature for professional communications” (4). From a practice perspective, this nomenclature allows for comparison and grouping of cognitive tests across studies. This cognitive taxonomy is widely accepted and used in the field of psychology, and we suggest that it also be appropriately applied in clinical trial research.

Neither of the authors had a conflict of interest.

Matthew P Pase  
Con Stough

Centre for Human Psychopharmacology  
Swinburne University of Technology  
400 Burwood Road  
Hawthorn, Victoria, 3122  
Australia  
E-mail: matthewpase@gmail.com

REFERENCES


Reply to MP Pase and C Stough

Dear Sir:

In an editorial commenting on our recently published study (1), which showed beneficial cognitive effects as a consequence of 6 mo of supplementation with DHA in healthy, young adults, Dangour and Allen (2) expressed major concerns over the heterogeneity of the tests being used to assess the cognitive function of adults in clinical trials. They illustrated their point by noting that a wide selection of cognitive tests had been used across the 10 relevant studies published in the Journal in 2011–2012, and that no 2 studies had adopted the same primary endpoints. What they failed to mention was that only one of these studies used any form of computer-administered cognitive testing. The other studies collected data in written or verbal form. In our own case, we used a sophisticated computerized cognitive assessment system (COM-PASS; Northumbria University) that was purpose designed to deliver multiple, parallel versions of a wide range of classic, standard, and bespoke cognitive tasks. The tasks used in the study were then chosen with reference to the recommendations of the European Food Safety Authority’s recent guidance on the cognitive tests that are suitable for assessing the effects of nutritional interventions (3) and previous work in this area by an expert panel under the auspices of the International Life Sciences Institute (4). The potential benefits of assessing cognitive function with a computer are self-evident and include the collection of accurate information on the speed of performing tasks and responding to stimuli. This information represents a fundamental measure of brain function and is always either equally informative or complementary to information on the accuracy of task performance. Beyond this, on a purely practical level, computerized testing also allows the standardized presentation of properly randomized stimuli, it removes the person-to-person interactions with a researcher that can bias and obfuscate data, and it allows the closely controlled collection of a large amount of data within a short period of time. We are literally surrounded in our everyday lives by powerful personal computers, and computerized cognitive testing can be readily adopted both in the laboratory and in more ecologically valid environments. Given the above, it is somewhat baffling that our own study was picked out for the editorial observations on the heterogeneity of testing across the field.

Dangour and Allen concluded their editorial by suggesting that experts in cognitive testing urgently need to reach a consensus on a small set of outcomes to use across future trials. Pase and Stough, in response, suggest that because consensus in this regard is unlikely, Carroll’s “Three Stratum Theory” (5) could provide a taxonomy for cognitive processes that could then inform a “standardized and evidence based approach to grouping cognitive test data.” By Pase and Stough’s account, our own “atheoretical” collapsing of task outcomes into “arbitrary” composite scores (which, in reality, were based on a previous factor analysis of a similar group of tasks) could be replaced by simply grouping or describing the task outcomes from a study with reference to the 8 broad cognitive ability domains and 69 narrow, well-defined abilities in Carroll’s model. Whereas this seems, on the face of it, to be a plausible suggestion, there are actually several major obstacles standing in the way of adopting this approach. From a purely practical perspective, a major problem would be deciding how a given task outcome maps onto one or more of Carroll’s factor analysis–derived “abilities.” Presumably, this process would require further factor analysis of multiple data sets. From a more theoretical perspective, Carroll’s model could also best be described as a “work in progress.” As he himself noted in his preface, the model was merely a starting point for future investigators and was formulated by looking backward. Carroll also acknowledged the inadequacy of some of the data that he had to work with. For instance,
he noted that the literature on memory and learning “leaves much to be desired” and listed the many gaps in the data that would need to be filled to arrive at a complete picture of this domain. In consequence, Carroll’s model has not been the fixed and stationary taxonomy that Pase and Stough would seem to be suggesting. Rather, it has been in a continuous state of modification since its initial publication. More recently, it has, for instance, been integrated with other models and has been modified and added to as new data and analytic techniques have become available (6). As an example, up to 6 new broad cognitive ability domains have been suggested as additions to Carroll’s original 8 domains (6). It is also notable that Carroll started work on his opus magnum in 1979 and worked on it for 14 y, synthesizing the findings of factor analyses from a vast body of data. Although he himself was a pioneer in the application of computer technology to his complex analyses, the data that he worked with were collected without the benefit of any such technology.

As McGrew noted recently (6), Carroll’s work represented a “ tipping point that provided the first working map of the human cognitive ability terrain, a terrain warranting additional exploration and refined cartographic efforts.” McGrew went on to urge the integration of current and future research into the emerging taxonomy. However, in this task we still seem to be laboring, certainly within the clinical trials field, with the astrolabes, quadrants, and verniers of the early map makers. Simply adopting the ubiquitous technology of our own age would necessarily make for much more accurate mapping tools, and therefore better maps. Although I applaud the ambition of Pase and Stough’s suggestion, I think the necessary first step toward their ultimate goal, and indeed greater standardization of cognitive tests, is the wider adoption of sensitive computerized testing techniques within the clinical trials field. The resulting data can then contribute to the factor-analytic process of further refining the map of human cognitive ability.

The author had no conflicts of interest.

David O Kennedy

Brain, Performance and Nutrition Research Centre
Northumbria University
Newcastle, NE1 8ST
United Kingdom
E-mail: david.kennedy@northumbria.ac.uk

REFERENCES

Erratum


Because of a copyediting error, data are missing in Table 3 for “Distribution” under “Scenario 3.” In the first 2 columns, under “Combination of the 4 countries,” the “Mean ± SD” value should be 0.18 ± 0.04, and the “95th Percentile” value should be 0.24.


Erratum


On page 693, the second sentence in the third paragraph of the Results section contains a copyediting error in which the word “or” was mistakenly used: “15 or 17 subjects” should read “15 of 17 subjects” instead.

Erratum


The supplemental data for this article were inadvertently missed during production and were therefore not posted online. The supplemental data file (Table 1) is now available online.


Erratum


On page 1039, an error appears in the legend to Figure 5. The solid circle line should represent skim milk, and the open circle line should represent the soy-protein beverage. The first sentence of the figure legend should read as follows: “Mean (±SEM) total amino acid (TAA) chemical net balance (NB) after consumption of a nonfat milk-protein beverage (●) or an isonitrogenous, isoenergetic, macronutrient-matched (750 kJ, 18.2 g protein, 1.5 g fat, and 23 g carbohydrate) soy-protein beverage (○).”


Erratum


On page 1053, footnote 2 should include the following additional funding information: “The study was also supported by CP07/00095 from the ISCIII, and MdMR-R was a recipient of a fellowship from ISCIII (Rio Hortega CM11/00030), Spanish Ministry of Economy and Competitiveness, Madrid, Spain.”