A multimodal intervention for Alzheimer’s disease results in multifaceted systemic effects reflected in blood and ameliorates functional and cognitive outcomes

Authors: Jared C. Roach¹, Lance Edens¹, Daria R. Markewych¹, Molly K. Rapozo², Junko Hara³, Gustavo Glusman¹, Cory Funk¹, Jennifer Bramen⁴, Priyanka Baloni⁵, William R Shankle³,⁶,⁷,⁸, Leroy Hood¹

¹ Institute for Systems Biology, Seattle, WA, USA
² Providence St. Joseph Health, Renton, WA, USA
³ Pickup Family Neurosciences Institute, Hoag Memorial Hospital Presbyterian, Newport Beach, CA, USA
⁴ Pacific Neuroscience Institute Pacific Brain Health Center Santa Monica California USA
⁵ Purdue University, IN, USA
⁶ Department of Cognitive Sciences, University of California, Irvine, CA, USA
⁷ Shankle Clinic, Newport Beach, CA, USA
⁸ EMBIC Corporation, Newport Beach, CA, USA

Corresponding Author:
Jared C. Roach
401 Terry Ave N
Seattle WA 98109, United States
206-732-2108
jared.roach@isbscience.org

Funding: This work was supported by the Alzheimer's Translational Pillar of Providence St. Joseph Health, by the National Science Foundation (NSF) Research Experiences for Undergraduates site "Research Opportunities in Systems Biology", by the Saint John’s Health Center Foundation, by the Pacific Neuroscience Institute Foundation (including the Cary and Will Singleton family), and by National Institutes of Health (NIH) grants R01AG062514, U01AG046139, RF1AG057443, & U01AG061359.

Keywords: diet, exercise, cognitive training, dementia, cognitive decline, Alzheimer’s disease and related disorders (ADRD), Alzheimer’s disease (AD), personalized coaching, remote coaching, COVID, systems biology, multimodal interventions, cognitive impairment, lifestyle, prospective randomized clinical trial (RCT)

Trial registration: ClinicalTrials.gov: NCT03424200 (clinicaltrials.gov/ct2/show/NCT03424200).

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Introduction. Comprehensive treatment of Alzheimer’s disease and related dementias (ADRD) requires not only pharmacologic treatment but also management of existing medical conditions and lifestyle modifications including diet, cognitive training, and exercise. The Coaching for Cognition in Alzheimer’s (COCOA) trial was a prospective randomized controlled trial (RCT) to test the hypothesis that a remotely coached multimodal lifestyle intervention would improve early-stage Alzheimer’s disease (AD). AD results from the interplay of multiple interacting dysfunctional biological systems. Specific causes of AD differ between individuals. Personalized, multimodal therapies are needed to best prevent and treat AD. COCOA collected psychometric, clinical, lifestyle, genomic, proteomic, metabolomic and microbiome data at multiple timepoints across two years for each participant. These data enable systems-biology analyses. We report analyses of the first COCOA data freeze. This analysis includes an evaluation of the effect of the intervention on outcome measures. It also includes systems analyses to identify molecular mediators that convey the effect of personalized multimodal lifestyle interventions on amelioration of cognitive trajectory.

Methods. A total of 55 participants with early-stage AD from Southern California were randomized into two parallel arms. Arm 1 (control; N=24) received standard of care. Arm 2 (intervention; N=31) also received telephonic personalized coaching for multiple lifestyle interventions including diet, exercise, and cognitive training. COCOA’s overarching aim was to gather dense molecular data from an AD cohort to improve understanding of pathophysiology and advance treatment. For the RCT, COCOA’s objective was to test the hypothesis that the Memory Performance Index (MPI) trajectory would be better in the intervention arm than in the control arm. The Functional Assessment Staging Test (FAST) was assessed for a secondary outcome. Assessments were blinded. The nature of the intervention precluded participant blinding.

Results. The intervention arm ameliorated 2.6 ± 0.8 MPI points (p = 0.0007; N = 48) compared to the control arm over the two-year intervention. Top-ranked candidate mediators included: albumin, propionylcarnitine, sphingomyelin, hexadecanedioate, acetylkynurenine, tiglylcarnitine, IL18R1, palmitoyl-sphingosine-phosphoethanolamine, acetyltryptophan, and IL17D. These individual molecules implicated inflammatory and nitrogen/tryptophan metabolism pathways. No important adverse events or side effects were observed.

Conclusions. Clinical trials should include frequent assessment of dense data to maximize knowledge gained. Such knowledge is useful not only in testing a primary hypothesis, but also in advancing basic biological and pathophysiological knowledge, understanding mechanisms explaining trial results, generating synergistic knowledge tangential to preconceived hypotheses, and refining interventions for clinical translation. Data from every trial should allow an intervention to be refined and then tested in future trials, driving iterative improvement. Multimodal lifestyle interventions are effective for ameliorating cognitive decline and may have an effect size larger than pharmacological interventions. Effects may be molecularly idiosyncratic; personalization of interventions is important. Dietary changes and exercise are likely to be beneficial components of multimodal interventions in many individuals. Remote coaching is an effective intervention for early stage ADRD. Remote interventions were effective during the COVID pandemic.
1. Introduction

Complex diseases require complex therapies. The COCOA trial was designed to test the hypothesis that multimodal interventions can slow progression, halt, or reverse — ameliorate — the course of Alzheimer’s disease (AD). Intervention was delivered by telephonic coaching. This coaching leveraged knowledge from prior studies to choose interventions expected to ameliorate cognition, including diet, exercise, and cognitive training.

Our overall aim in COCOA is to explore four overarching hypotheses that (1) there are multiple causes of AD, (2) there are multiple interventions to ameliorate AD, (3) different individuals with AD will have different mixes of causes, and (4) the best treatment for any given individual should be personalized to include multimodal interventions that treat the specific causes of their AD (Figure 1). COCOA pre-specified a simple hypothesis to be tested — that the intervention group will have a better cognitive trajectory than the control group. However, proving this hypothesis (lemma #1) does little by itself to advance our four overarching hypotheses. In order to advance these hypotheses, we will also show that (lemma #2) the COCOA intervention causes broad multisystem changes reflected in numerous blood analytes, (lemma #3) several of these analytes are correlated with cognitive improvement and are thus candidate mediators of the effects of the multimodal interventions, (lemma #4) there are mediators & mechanisms shared by multiple individuals that suggest some shared mechanisms may be broadly applicable to many or most AD patients, (lemma #5) there may be some mediators operating in only one or a few COCOA participants, (lemma #6) that the mechanisms identified — both aggregate and individual — are not randomly distributed across all systems, and (lemma #7) that several identified mechanisms have considerable prior evidence for mediating AD interventions. Together these lemmas will bolster support for our four overarching hypotheses. They leverage the epistemological concepts of coherence — combining the ensemble of the data into biological plausibility (Fedak et al., 2015). This synergy is driven by the integration of knowledge from various evidence streams, particularly omics.

2. Methods

Methods for the COCOA trial were published as a separate manuscript (Roach, Hara, Fridman, et al., 2022). In brief, COCOA is a prospective randomized clinical trial (RCT) to test the hypothesis that coached multimodal interventions ameliorate cognitive decline. Main inclusion criteria: age 50 and older; Memory Performance Index (MPI; embic.us/application/clinical) of 65 or below. COCOA recruited, enrolled, and

---

1 Abbreviations: AD, Alzheimer’s disease; ADRD, Alzheimer’s disease and related disorders (or nearly equivalently: Alzheimer’s disease and related dementias), MCI, mild cognitive impairment; MCIS, MCI Screen; RCT, randomized controlled trial; MPI, Memory Performance Index; MoCA, Montreal Cognitive Assessment; FAST, Functional Assessment Staging Test.
randomized 55 participants over 16 months, averaging over 3 enrollments per month (Supplemental Figures 1 & 2). The trial protocol as approved by the Western Institutional Review Board (WIRB; protocol #20172152) is included as Appendix 1. In April 2019, Arivale, our primary provider of trial logistics, ceased operations. Most operations previously contracted through Arivale were then assumed in house. Enrollment ceased due to Arivale ceasing operations. Circa March 2020, the COVID pandemic began. COVID increased risk for in-person activities. The intervention — telephonic coaching — was exceptionally robust (almost presciently so) to COVID disruptions. At least two blood draws were completed by 43 participants (Supplemental Figure 3).

Baseline demographic characteristics are presented in Table 1.

3. Results and Intermediate Logical Steps (Lemmas)

We present our result in seven subsections, each corresponding to a lemma. Lemma #1 is our claim that the COCOA intervention results in cognitive and functional improvement. The other lemmas are: (lemma #2) the COCOA intervention causes broad multisystem changes reflected in numerous blood analytes, (lemma #3)
Figure 2. Aggregate MPI trajectory of COCOA Participants. The COCOA intervention ameliorates aggregate cognitive trajectory. Each MPI measurement from every participant is shown. Cognitive measurements vary considerably between individuals and over time. Change from baseline is graphed. The primary outcome measure is computed with linear regression and graphed in Supplemental Figure 4. However, as the aggregate cognitive trajectories of each arm are most likely non-linear, this figure (illustrated with spline trajectories) may more accurately convey the nature of the trajectories than a linear fit. The baseline is shown as a dark black line and serves as a reference for no change over time. The intervention arm diverges markedly from the control arm during the first several months of the trial, and then maintains a fairly constant improvement over the remaining months. Both trajectories eventually drop below baseline, but with the intervention arm losing relatively little cognition from baseline. Individual trajectories are graphed in Supplemental Figure 5. 90% confidence intervals are shaded.
several of these analytes are correlated with cognitive improvement and are thus candidate mediators of the effects of the multimodal interventions, (lemma #4) there are mediators & mechanisms shared by multiple individuals that suggest some shared mechanisms that may be broadly applicable to many or most AD patients, (lemma #5) there may be some mediators that may be operating in only one or a few COCOA participants, (lemma #6) identified mechanisms implicate specific subsystems, and (lemma #7) these mechanisms are already known to mediate AD interventions.

3.A. Lemma 1: The COCOA intervention causes an amelioration of cognitive and functional trajectories.

Cognition (MPI) is significantly ameliorated in the intervention arm. Linear regression showed an effect size of a 2.6 ± 0.8 MPI point benefit (p = 0.0007; N = 48) to the intervention (coached) arm compared to the control (standard of care) arm over the two-year course of the intervention (Supplemental Figure 4). We performed a similar analysis with the Montreal Cognitive Assessment (MoCA), which showed similar results (Supplemental Figure 6). The MPI benefit was most evidenced as a fixed effect, rather than as an interactive term with time, suggesting that the trajectory of benefit was nonlinear. We fit a spline curve to the data to illustrate this non-linearity (Figure 2). This spline shows that most of the cognitive benefit occurs in the first year of intervention – perhaps during the first several months. Function as measured by the Functional Assessment Staging Test (FAST) is also significantly ameliorated in the intervention arm: linear regression showed a benefit to FAST (p = 0.030; Supplemental Figure 7). Therefore, both a measure of cognition (MPI) and a measure of function (FAST) ameliorated due to the COCOA intervention. This suggests that the COCOA intervention has broad multifaceted benefits across multiple systems – that the COCOA intervention is beneficial for overall brain health and for patient-oriented outcomes relevant to dementia.

3.B. Lemma 2: The COCOA intervention causes systemic molecular changes in the blood.

We expect that the COCOA intervention causes many effects in many organs and tissues, and that these are reflected in the blood. We analyze three classes of blood omics data: plasma proteomics, plasma metabolomics, and clinical labs. Collectively, we call all the measured molecules and metrics from these assays “analytes”. Our global analyses encompassed 1461 analytes: 456 protein assays, 952 metabolites, and 53 clinical labs; 1447 of these are distinct. Our analysis of the aggregate cognitive trajectory (Figure 2) suggests that the response to intervention is non-linear, with the greatest divergence between the case and control trajectory happening during the first six months of the intervention. Therefore, we focus on the initial molecular physiological trajectory of participants. In addition to computing all-by-all correlations, we computed two statistics for each analyte. One statistic estimates the analyte’s value as a biomarker for the intervention, and the other statistic estimates the analyte’s value as a biomarker for the outcome. These two “intervention value” and “outcome value” statistics were computed as the significance of the slope of the regression between the analyte value and: (statistic 1) arm assignment or, (statistic 2) the overall MPI slope.
To provide evidence for the causal effect of the COCOA intervention on the blood, Figure 3 is a multidimensionally scaled (MDS) representation of all analytes. The significance of association between the case and control values for that analyte is represented by the size of the icon. The clustering of all significant analytes is astronomically significant by inspection. We also performed an analysis focused solely on
Considering these analyses together, analyte clustering must be due to the intervention. Therefore, the intervention causes manifold changes in analytes in the blood.

Figure 4. Bile acid, xenobiotic, and nitrogen metabolism pathways are likely altered by the COCOA intervention. The top pathways implicated by alterations in metabolites in response to the COCOA intervention are all consistent with prior knowledge that these pathways are impacted by diet and consumption of xenobiotics. Some of these pathways are also known to be affected by exercise and other components of the multimodal intervention. In order to leverage an existing workflow for pathway enrichment (Metaboanalyst, www.metaboanalyst.ca; Pang et al., 2021), for this figure only, we limited our analysis to the metabolite subset of all COCOA analytes. We divided metabolites into two groups, based on overall variance of all measurements of that analyte. High-variance analytes might have non-monotonic and non-linear trajectories; sampling such trajectories a few times a year could abet overconfidence. Because power to detect the effect of the intervention on a pathway depends on how well the components of that pathway can be detected by metabolite mass spectrometry and COCOA sampling frequency, failure to highly rank a pathway is not strong evidence for the absence of effects of the intervention on that pathway. (A) Pie charts indicate the category abundance of metabolites in each classification. (B) Pathway enrichment. (Left panel) Nitrogen metabolism pathways are enriched by low-variance metabolites; (Right panel) Bile acid and xenobiotic pathways are enriched by high-variance metabolites.
We are using a multimodal intervention. Although the desire is for this intervention to be as personalized as possible, many aspects of our interventions will have no effect on cognition. Furthermore, even if we know what node or subsystem we want to target (e.g., blood pressure), few interventions specifically target that element of the system. For example, a suboptimal exercise program might improve blood pressure but degrade sleep. In this manuscript, we seek to identify specific mediators that not only change in response to the COCOA intervention, but also improve cognition. We continue on this epistemological path in the following lemmas.

3.C. Lemma 3: Molecular changes in the blood of COCOA participants are correlated with cognitive improvement.

For this lemma, we do not consider participant arm; that is, we ignore the case or control “label”. We treat the COCOA trial as a cohort study. Individuals in the study differ in lifestyle for many reasons; they differ genetically and environmentally. The analytes most correlated with overall MPI slope are presented in Table 3. A primary “full” analysis considers all data, and a secondary “robust” analysis excludes outliers. Taken individually, the significance of any given analyte is weak. However, patterns emerge when these data are considered as an ensemble.

Xenobiotics, such as aspirin and food components (in some cases, “supplements”), are common among the most significant analytes in Table 3. The top two analytes are salicylate and N-acetylalliin (a xenobiotic metabolite that can be derived from garlic). Aspirin dose is a known predictor of cognitive trajectory in AD, at least in some contexts (Wattmo et al., 2011). Identification of these xenobiotics suggests that interventions other than the COCOA intervention may play a large role in cognitive improvement. The effect of these xenobiotics appears to be independent of the effect of the multimodal lifestyle intervention, and that therefore the best interventions may include combinations of both lifestyle and pharmacological/xenobiotic interventions.

Consumption of particular pharmaceuticals or dietary chemicals that may be correlated with cognitive trajectory and may or may not be causal. Another non-exclusive explanation for the significance of these xenobiotics is that the presence of xenobiotics in the blood is a biomarker of attention and compliance to health advice — these xenobiotics may correlate with cognitive improvement but not cause cognitive improvement. A number of other analytes suggest other mechanisms, including those that may be targeted by the COCOA intervention.

3.D. Lemma 4: Some potential mediators & mechanisms may be shared by multiple individuals.

A potential mediator should (1) be influenced by the COCOA intervention, (2) influence cognition, and (3) the intervention effect should be in the same direction as the cognition effect. Therefore, an analyte must be identified as a top candidate in both lemma 2 and lemma 3 to be a potential mediator. Although our overarching hypothesis is that there are multiple mechanisms that may contribute personalizable therapies for AD, and that these may not be shared between particular individuals, we are confident there will be some that are shared between many individuals diagnosed with AD. If that were not the case, AD would no longer merit distinction as a named disease entity. Therefore, we aimed to detect shared or common molecular mediators of the
Correlations between changes in COCOA analytes from baseline to first follow-up

Figure 5. Candidate mediators of the COCOA intervention loosely cluster into two groups. Absolute correlation distances (edges) are shown between highly correlated COCOA analytes (nodes), based on changes between the first and second COCOA timepoints. One cluster including IL17D and IL18R1 may implicate an immune or inflammatory pathway(s) in mediation; another cluster may indicate metabolic pathway(s) in mediation. The non-random distribution of candidate mediators would be unlikely if they were all false positives (noise). This high-correlation network of all analytes is visualized via the Fruchterman-Reingold algorithm (Fruchterman & Reingold, 1991). This network includes the best edge for each node if that edge has a correlation of at least 0.55, and all edges with correlations of at least 0.85. This graph emphasizes the relationships and resulting correlation paths determined by the strongest most significant correlations between analytes, whereas the visualization in Figure 3 is influenced by subtle effects from hundreds of weak correlations in the all-by-all analysis. Edge thickness ~ correlation; red indicates candidate mediator (Table 5). Labels are shifted radially out from the node they label to avoid obscuring network connectivity. The placement of albumin is arbitrary as it is not highly correlated to any other analyte.

COCOA intervention. To this end, we compared the top-ranked “robust” rankings of analytes that ranked highly in both (1) significance of response to the COCOA intervention (listed in Table 2, from lemma 2), and (2) significance of association with MPI slope (listed in Table 3, from lemma 3). These include both “full”-data ranking and “robust” rankings. The intersection between these lists is shown in Table 4. Three of the top seven
analytes (kynurenate, N-acetylkynurenine, N-acetyltryptophan) are associated with tryptophan metabolism. This coherence suggests that certain subsystems (in addition to individual analytes) might mediate the intervention.

There are four analytes that are nominally significant in both Table 2 and Table 3 (full analysis): propionylcarnitine, kynurenate, gamma-glutamyl-alpha-lysine, and albumin. This overlap of 4 is about the same as would be expected solely due to the simplest prediction from type I error (0.05 x 0.05 x 1461 = 3.6). Even though our p-values are approximate — and we likely lack power to detect changes in many of the analytes due to analyte idiosyncrasies (e.g. sparse consumption of certain xenobiotics), lack of much greater intersection than might be expected by chance weakens our argument that there is true signal in this list; our argument must rely on coherence and consistency considerations, which we describe in lemmas 6 & 7, below.

3.E. Lemma 5: Some potential mediators may be operating in only one or a few COCOA participants.

It is possible that some individuals may respond particularly well to a specific component of the COCOA intervention and/or via a specific molecular mediator. To this end, we compared the top-ranked “full” rankings of analytes that ranked highly in both (1) significance of response to the COCOA intervention (Table 2), and (2) significance of association with MPI slope (Table 3). Focus on the “full” rankings can identify analytes that have an importance driven primarily by a single or a few individuals. The intersection between these lists is shown in Table 4. Identification of these “personal” mediators in COCOA could lead to therapeutic targets available to a broader range of individuals. For example, there may be aspects of the intervention that were difficult to comply with, and only benefitted a few individuals able to comply. If we can better identify such intervention aspects, we may be able to drop compliance barriers or use alternative interventions to broaden the individuals that could leverage one or more of these personal mediators to drive cognitive benefit.

3.F. Lemma 6: Potential mediators and mechanisms are not random.

Random distribution of mechanistic attributes of potential mediators would be predicted if they all arose by chance. Non-random distribution argues against our list of potential mediators all being the result of experimental noise. To make our claim that COCOA mediators cluster cohesively, we computed correlations between all possible pairs of analytes. We then portrayed them with dimensionality reduction (Figure 5). The best candidate mediators cluster more than expected by chance, suggesting that there is a specific set of systems mediating the COCOA intervention’s effect on cognition, and that the analytes within a system preferentially correlate with each other.

3.G. Lemma 7: Several identified mechanisms have considerable prior evidence for mediating AD interventions.

There are two lines of prior arguments that suggest that there are likely several if not many molecular mediators in the blood that might be identified by the COCOA study. First, a large number of mediators have been suggested in the literature, with varying degrees of evidence (e.g., Niranjan, 2013; Verdile et al., 2015). Although most of these devolve from studies of causes of AD, many are also likely to mediate ameliorative
treatments, presumably with an opposite direction of effect. Second, since we observe an effect of the COCOA intervention on cognition (lemma #1), there must be some mediator(s). Since the main COCOA modalities are diet (including xenobiotics) and exercise, and the vast majority of biological information conveyed by these modalities is conducted through the blood (after transduction through gut, muscle, heart, and other organs), there are likely several if not many mediators of their effects to be found in the blood. These arguments establish a high prior probability that a number of molecular mediators should be identifiable in the blood. We list all decent analyte mediator candidates from any of the above analyses (Table 5).

One might expect candidate mediators for ameliorating AD to be on a list of “well known suspects”, or closely linked to such suspects in knowledge graphs. For example, one might expect a gene implicated by AD GWAS to have its gene product identified, or an analyte functionally associated with that gene. None of the candidate COCOA mediators intersect with a list of GWAS-implicated genes (Seto et al., 2021). Most mediators of multimodal interventions identified by COCOA target pathways are distinct from the pathways conveying genetic causes of AD. COCOA interventions might target (1) environmental causes of AD, and (2) pathways that treat but do not cause AD.

**Subsystems.** Chen et al. leveraged two-sample Mendelian randomization and GWAS summary statistics to explore the causal association between 486 metabolites and five neurodegenerative diseases (NDDs) including AD (Chen et al., 2021). They suggestively associated 164 metabolites with the risk of at least one neurodegenerative disease. They found metabolic pathways involved in NDDs, including “urea cycle” and “arginine biosynthesis”, “purine metabolism”, and “D-Glutamine and D-glutamate metabolism” for AD. They also found “phenylalanine, tyrosine and tryptophan biosynthesis” for ALS and MS. “Carnitine synthesis” was associated with FTD. Leao et al. (2021) show that the acylcarnitine subsystem differs between controls and Alzheimer’s patients. Particular metabolites implicated in AD included urea and hexadecanedioate. Hexadecanedioate is also implicated in ischemic stroke (Sun et al., 2019), possibly mediated by its effects on blood pressure and/or its hypolipidemic, anti-obesity, or anti-diabetogenic properties. These pathways markedly overlap the pathways implicated by our current analysis. More broadly in the literature, all of these pathways have been implicated in ADRD. Some may potentiate AD by a vascular dementia diasthesis, and some are affected by APOE genotype.

4. Conclusions

We have shown that personalized multimodal lifestyle interventions can ameliorate cognitive trajectories in individuals on the AD spectrum. These interventions cause many changes in blood analytes. Several of these analytes are also correlated with cognitive amelioration. Analytes that are altered following the intervention and that are correlated with cognitive amelioration are candidate mediators. A mediator mechanistically (causally) conveys the influence of the intervention, propagating its information through biological pathways, to the brain, where it influences the final common pathways of ADRD pathophysiology, presumably on neurons and their
synapses. The mediators identified by COCOA implicate inflammatory and metabolic pathways that respond to changes in diet and possibly exercise. A number of candidate mediators (Tables 4 & 5) are anticorrelated with cognitive improvement. Reductions in blood levels of these mediators appear beneficial to cognition. Several of these are associated with food consumption, particularly highlighting nitrogen metabolism. Therefore, reducing caloric intake, possibly including protein restriction, may be a beneficial component of multimodal interventions in some individuals.

Our results have a number of broad implications. First, COCOA served well as a pilot trial. More trials should be modeled after COCOA’s trial design. Such designs can generate new information, both testing well defined hypotheses and expanding the scope of inquiry into areas of pathophysiology that are less explored. Second, ADRD can be treated using mediators that are not directly related to the most prominent pathways (e.g., amyloid hypothesis) targeted by many pharmaceutics. Pathways to treat ADRD may be different than pathways that cause ADRD. Lifestyle treatments are likely to complement, not replace, pharmaceutical treatments. Both are likely to be part of many optimal personalized multimodal therapies. Third, lifestyle interventions should be emphasized in current treatment for all individuals on the ADRD spectrum. More funding for lifestyle interventions should be applied both to research and via payers for medical care.

There are numerous limitations to the COCOA trial. Briefly, we surveyed a large number of variables, and although some results were nominally statistically significant, these results would not be convincing on their own. The general conclusions are solid; specific conclusions about specific analytes are not. Correlation is not always causation. Correlation due to prospective randomization in an RCT is a very strong argument for causation. Thus we can be somewhat confident that the COCOA intervention causes broad molecular changes in the blood, and that the COCOA intervention improves the cognitive trajectory of AD. Although we provide some evidence for specific blood analytes that mediate this causation, there are alternative explanations for mediation. The same individuals who are compliant with diet and exercise advice may also be compliant with cognitive training, causing their blood omics to shift through their careful diet and exercise behaviors. Such shifts in blood omics could in theory be a red herring — the actual cause of cognitive improvement could be via neurocognitive mechanisms driven by the cognitive training elements of the multimodal intervention. However, biological plausibility — discussed below — suggests that blood does indeed convey at least some causality.

**Biological plausibility & the rising importance of epistemology for the future of clinical trials.** An important component of systems epistemology is biological plausibility: placing current research coherently and consistently in the context of previous research. More generally, new observations should be interpreted in the light of previous observations. Much of the argument we present today would not have been as strong a few years ago. In particular, at the time of trial design prior to beginning enrollment of participants, we could not have known the best methods of analysis or computed an exact power for the trial. These uncertainties create practical challenges for launching trials designed to take advantage of future knowledge. Such challenges must
be overcome for the full potential of research on complex diseases to flourish. Although not used extensively in our COCOA analyses to date, artificial intelligence (AI) is likely to potentiate the value of dense-data studies (Bughin et al., 2017; Naylor, 2017).

**Final thoughts.** A large amount of contextual knowledge concerning the roles, associations, mechanisms, and relationships of once-obscure metabolites is becoming available. Many of these implicate or highlight the roles of particular metabolites in response to diet and exercise. Dense-data trials are designed to be increasingly valuable as global knowledge accumulates. Such knowledge, particularly for metabolites, is increasing rapidly; the value of dense trials (both newly conducted and re-analyses of previously released data) will also increase rapidly.

**Supplemental Table Legends**

**Supplemental Table 1.** Data for all analytes. Each value is the difference between measurement at the baseline and first follow-up timepoint for an individual. Units are as described in Methods.

**Supplemental Table 2.** Population-level measures for specific analytes may be causally influenced by variation in specific genes. Many of the causal associations between genes and metabolites have been compiled (pheweb.org; Gagliano Taliun, et al., 2020). These associations relate some of the candidate mediators identified by COCOA. For example: KMO is a top influencer for both kynurenate and acetylkynurenine; SLCO1B1 is a top influencer for both hexadecanedioate and pregnenolone.

**Contributions**

Planning: LH, WS, JR; Analysis: JR, LE, DM, WS; Coaching: MR; Manuscript: All authors.

**Acknowledgements**

Deborah Fridman, Jennifer C. Lovejoy, Kathleen Jade, Laura Heim, Rachel Romansik, Adrienne Swietlikowski, Sheree Phillips, Maria Fischer, Dan Fischer, Lauren Dill, Michael Brant-Zawadzki, Sophiya Rajbhandari, Dale Bredesen, Mary Kay Ross, Amanda Wikan, Alex Lewis, Maia Kurnik, Christie Davidson, and Jennifer Eklund have inspired, managed, mooted, conceived, planned, participated, organized, supported, discussed, and/or analyzed aspects of the COCOA project or its precursors. COCOA would not have been possible without the participants, for whom we reserve the greatest acknowledgement and appreciation. Institutional support from Hoag Memorial Hospital Presbyterian was vital. Oregon Health & Science University’s Bioinformatics and Computational Biomedicine program supported interns. Data used in preparation of this article were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A
complete listing of ADNI investigators can be found at adni.loni.usc.edu. Statue images (Figure 1): Folco Masi & Ziad Al Halabi (unsplash.com).

Data Availability
Data is available in Supplemental Table 1.

Declaration of interests
Dr. Shankle is an employee of EMBIC Corporation. Dr. Hara owns stock in EMBIC Corporation. There are no other conflicts of interest for any of the authors.

References


abnormalities and potential biomarkers associated with treatment of GLUT1 deficiency with ketogenic diet.


Hendrix, SB et al. Not sexy enough: the uphill battle faced by non-traditional treatments in the U.S. despite having the most compelling AD results to date. AAIC 2022.


Wood; Amy K. Glen; Lindsey G. Kvarfordt; Finn Womack; Liliana Acevedo; Timothy S. Yoon; Chunyu Ma; Veronica Flores; Meghamala Sinha; Yodsawalai Chodpathumwan; Arash Termehchy; Jared C. Roach; Luis Mendoza; Andrew S. Hoffman; Eric W. Deutsch; David Koslicki; Stephen A. Ramsey. RTX-KG2: a system for building a semantically standardized knowledge graph for translational biomedicine. BMC Bioinformatics. Submitted 2022.


<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Control (n=22)</th>
<th>Coaching (n=31)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female, n (%)</td>
<td>8 (36%)</td>
<td>10 (32%)</td>
</tr>
<tr>
<td>Age, mean years (range)</td>
<td>73.5 (58-89+)</td>
<td>75.5 (58-89+)</td>
</tr>
<tr>
<td>MPI score, mean (range)</td>
<td>42.5 (13-64)</td>
<td>40.2 (15-63)</td>
</tr>
<tr>
<td>FAST score, mean (range)</td>
<td>3.0 (2-4)</td>
<td>3.2 (2-4)</td>
</tr>
<tr>
<td>Race, fraction White (%)</td>
<td>22/22 (100%)</td>
<td>27/31 (84%)</td>
</tr>
<tr>
<td>BMI, mean (range)</td>
<td>26.7 (21, 34)</td>
<td>25.6 (20, 34)</td>
</tr>
</tbody>
</table>

Maximum education completed
- High school/GED: 2 2
- Some college/tech school: 1 6
- Associate and/or bachelor’s: 11 14
- Master’s/doctorate/professional: 5 6
- No response: 3 3

Table 1. Baseline participant characteristics. Demographic attributes were balanced between the two arms of the COCOA trial (N=53 completed baseline assessments).
Table 2. List of analytes most significantly altered by the COCOA intervention. These analytes have the most significant differences in analyte change between the two arms of the trial. They are not necessarily associated with outcomes. Analyte changes are computed as the difference in value between the first two timepoints, approximately four months apart — but with a few outliers due to the COVID pandemic ([Supplemental Figure 3](#)). The primary “full” analysis included all data points. A secondary “robust” analysis excluded outlier analyte values. For a comprehensive overview, consult [Supplemental Table 1](#), and sort by the column(s) of interest.

<table>
<thead>
<tr>
<th>Analyte ID</th>
<th>Name</th>
<th>Symbol or Pathway or System or Category</th>
<th>p-value (full)</th>
<th>p-value (robust)</th>
<th>Rank (full)</th>
<th>Rank (robust)</th>
<th>Effect Size (full)</th>
<th>Effect Size (robust)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVD2_P40225</td>
<td>thrombopoietin</td>
<td>THPO</td>
<td>0.0002</td>
<td>0.0007</td>
<td>1</td>
<td>5</td>
<td>1.10</td>
<td>0.89</td>
</tr>
<tr>
<td>100001540</td>
<td>pyroglutamine</td>
<td>Glutamate Metabolism</td>
<td>0.0003</td>
<td>0.0008</td>
<td>2</td>
<td>6</td>
<td>-1.23</td>
<td>-0.95</td>
</tr>
<tr>
<td>100008993</td>
<td>1-palmityo1-2-arachidonoyl-GPI (16:0/20:4)</td>
<td>Phospholipid Metabolism</td>
<td>0.0004</td>
<td>0.0014</td>
<td>3</td>
<td>8</td>
<td>-0.91</td>
<td>-0.74</td>
</tr>
<tr>
<td>100020219</td>
<td>3-formylindole</td>
<td>Food Component/Plant</td>
<td>0.0007</td>
<td>0.0004</td>
<td>4</td>
<td>1</td>
<td>-0.98</td>
<td>-0.94</td>
</tr>
<tr>
<td>100009181</td>
<td>1-stearyo1-2-oleoyl-GPI (18:0/18:1)</td>
<td>Phospholipid Metabolism</td>
<td>0.0007</td>
<td>0.0007</td>
<td>3</td>
<td>3</td>
<td>-1.33</td>
<td>-1.33</td>
</tr>
<tr>
<td>100001026</td>
<td>galactonate</td>
<td>Fructose, Mannose and Galactose Metabolism</td>
<td>0.0011</td>
<td>0.0007</td>
<td>6</td>
<td>4</td>
<td>-6.76</td>
<td>-6.39</td>
</tr>
<tr>
<td>566</td>
<td>valine</td>
<td>Leucine, Isoleucine and Valine Metabolism</td>
<td>0.0014</td>
<td>0.0014</td>
<td>7</td>
<td>7</td>
<td>-0.43</td>
<td>-0.43</td>
</tr>
<tr>
<td>98</td>
<td>kynurenate</td>
<td>Tryptophan Metabolism</td>
<td>0.0018</td>
<td>0.0018</td>
<td>8</td>
<td>9</td>
<td>-0.64</td>
<td>-0.64</td>
</tr>
<tr>
<td>CVD2_Q9NQ25</td>
<td>SLAM family member 7</td>
<td>SLAMF7</td>
<td>0.0021</td>
<td>0.0026</td>
<td>9</td>
<td>11</td>
<td>0.71</td>
<td>0.64</td>
</tr>
<tr>
<td>100001577</td>
<td>N-acetylcitrulline</td>
<td>Urea cycle; Arginine and Proline Metabolism</td>
<td>0.0027</td>
<td>0.0091</td>
<td>10</td>
<td>27</td>
<td>-2.22</td>
<td>-1.65</td>
</tr>
<tr>
<td>NEX_O14944</td>
<td>epieregulin</td>
<td>EREG</td>
<td>0.0034</td>
<td>0.0115</td>
<td>11</td>
<td>36</td>
<td>2.57</td>
<td>2.09</td>
</tr>
<tr>
<td>100001162</td>
<td>propionylcarbinicute</td>
<td>Fatty Acid Metabolism ( &amp; BCAA Metabolism)</td>
<td>0.0035</td>
<td>0.0044</td>
<td>12</td>
<td>15</td>
<td>-0.89</td>
<td>-0.78</td>
</tr>
<tr>
<td>CVD2_Q14763</td>
<td>TNF receptor superfamily, member 10b</td>
<td>TNFRSF10B</td>
<td>0.0038</td>
<td>0.0038</td>
<td>13</td>
<td>12</td>
<td>0.59</td>
<td>0.59</td>
</tr>
<tr>
<td>100008984</td>
<td>1-palmityo1-2-palmitoleoyl-GPC (16:0/16:1)</td>
<td>Phospholipid Metabolism</td>
<td>0.0039</td>
<td>0.0116</td>
<td>14</td>
<td>38</td>
<td>-0.91</td>
<td>-0.60</td>
</tr>
<tr>
<td>CVD2_Q95523</td>
<td>sortilin 1</td>
<td>SORT1</td>
<td>0.0043</td>
<td>0.0043</td>
<td>15</td>
<td>14</td>
<td>0.68</td>
<td>0.68</td>
</tr>
<tr>
<td>100002784</td>
<td>2-oxoarginine</td>
<td>Tryptophan Metabolism</td>
<td>0.0054</td>
<td>0.0054</td>
<td>16</td>
<td>17</td>
<td>-1.77</td>
<td>-1.77</td>
</tr>
<tr>
<td>100000015</td>
<td>xanthurenamat</td>
<td>Tryptophan Metabolism</td>
<td>0.0056</td>
<td>0.0056</td>
<td>17</td>
<td>18</td>
<td>-1.26</td>
<td>-1.26</td>
</tr>
<tr>
<td>1539</td>
<td>1-palmityo1-2-oleoyl-GPC (16:0/18:1)</td>
<td>Phospholipid Metabolism</td>
<td>0.0057</td>
<td>0.0057</td>
<td>18</td>
<td>19</td>
<td>-0.30</td>
<td>-0.30</td>
</tr>
<tr>
<td>CVD2_P01241</td>
<td>growth hormone 1</td>
<td>GH1</td>
<td>0.0059</td>
<td>0.0051</td>
<td>19</td>
<td>16</td>
<td>3.66</td>
<td>3.40</td>
</tr>
<tr>
<td>999947671</td>
<td>X - 18921</td>
<td></td>
<td>0.0065</td>
<td>0.0065</td>
<td>20</td>
<td>21</td>
<td>1.17</td>
<td>1.17</td>
</tr>
<tr>
<td>100000276</td>
<td>maltotriose</td>
<td>Glycogen Metabolism</td>
<td>0.0068</td>
<td>0.0230</td>
<td>21</td>
<td>54</td>
<td>8.84</td>
<td>6.72</td>
</tr>
<tr>
<td>100000040</td>
<td>ribitol</td>
<td>Pentose Metabolism</td>
<td>0.0078</td>
<td>0.0154</td>
<td>22</td>
<td>42</td>
<td>-0.63</td>
<td>-0.38</td>
</tr>
<tr>
<td>INF_O43557</td>
<td>TNF (ligand) superfamily, member 14</td>
<td>TNFSF14</td>
<td>0.0080</td>
<td>0.0248</td>
<td>23</td>
<td>59</td>
<td>1.91</td>
<td>1.27</td>
</tr>
<tr>
<td>NEX_O94732</td>
<td>FK506 binding protein 7</td>
<td>FKBP7</td>
<td>0.0082</td>
<td>0.0082</td>
<td>24</td>
<td>23</td>
<td>1.18</td>
<td>1.18</td>
</tr>
<tr>
<td>100009066</td>
<td>1-palmityo1-2-oleoyl-GPI (16:0/18:1)</td>
<td>Phospholipid Metabolism</td>
<td>0.0082</td>
<td>0.0005</td>
<td>25</td>
<td>2</td>
<td>-1.41</td>
<td>-1.51</td>
</tr>
<tr>
<td>391</td>
<td>citrulline</td>
<td>Urea cycle; Arginine and Proline Metabolism</td>
<td>0.0084</td>
<td>0.0084</td>
<td>26</td>
<td>24</td>
<td>-0.51</td>
<td>-0.51</td>
</tr>
<tr>
<td>1002</td>
<td>allantoin</td>
<td>Purine Metabolism, Xanthine/Inosine</td>
<td>0.0084</td>
<td>0.0060</td>
<td>27</td>
<td>20</td>
<td>-0.47</td>
<td>-0.43</td>
</tr>
<tr>
<td>407</td>
<td>lysine</td>
<td>Lysine Metabolism</td>
<td>0.0088</td>
<td>0.0088</td>
<td>28</td>
<td>25</td>
<td>-0.21</td>
<td>-0.21</td>
</tr>
<tr>
<td>100001272</td>
<td>1-oleoyl-GPC (18:1)</td>
<td>Lysolipid</td>
<td>0.0090</td>
<td>0.0090</td>
<td>29</td>
<td>26</td>
<td>-0.43</td>
<td>-0.43</td>
</tr>
<tr>
<td>999933132</td>
<td>X - 11787</td>
<td></td>
<td>0.0094</td>
<td>0.0094</td>
<td>30</td>
<td>29</td>
<td>-0.31</td>
<td>-0.31</td>
</tr>
<tr>
<td>CVD2_P22004</td>
<td>bone morphogenetic protein 6</td>
<td>BMP6</td>
<td>0.0097</td>
<td>0.0309</td>
<td>31</td>
<td>78</td>
<td>1.04</td>
<td>0.81</td>
</tr>
<tr>
<td>533</td>
<td>urea</td>
<td>Urea cycle; Arginine and Proline Metabolism</td>
<td>0.0099</td>
<td>0.0099</td>
<td>32</td>
<td>31</td>
<td>-0.72</td>
<td>-0.72</td>
</tr>
<tr>
<td>999947802</td>
<td>X - 16397</td>
<td></td>
<td>0.0101</td>
<td>0.0101</td>
<td>33</td>
<td>33</td>
<td>0.83</td>
<td>0.83</td>
</tr>
<tr>
<td>100008904</td>
<td>1-stearyo1-2-oleoyl-GPC (18:0/18:1)</td>
<td>Phospholipid Metabolism</td>
<td>0.0108</td>
<td>0.0108</td>
<td>34</td>
<td>35</td>
<td>-0.49</td>
<td>-0.49</td>
</tr>
<tr>
<td>10000263</td>
<td>imidazole lactate</td>
<td>Histidine Metabolism</td>
<td>0.0108</td>
<td>0.0353</td>
<td>35</td>
<td>88</td>
<td>-0.35</td>
<td>-0.26</td>
</tr>
<tr>
<td>CVD2_Q14116</td>
<td>interleukin 18</td>
<td>IL18</td>
<td>0.0109</td>
<td>0.0370</td>
<td>36</td>
<td>92</td>
<td>0.86</td>
<td>0.60</td>
</tr>
</tbody>
</table>
Table 3. List of analytes most significantly correlated with MPI slope. Top analytes in this table may or may not necessarily be affected by the COCOA intervention. Variations in these analytes may be influenced by many factors driving differences between participants in both arms, including genetics and lifestyle choices not impacted by coaching. Analyte changes are computed as the difference in value between the first two timepoints, approximately four months apart — but with a few outliers due to the COVID pandemic (Supplemental Figure 3). The primary “full” analysis included all data points. A secondary “robust” analysis excluded outlier analyte values. For a comprehensive overview, consult Supplemental Table 1, and sort by the column(s) of interest.
## Table 4. Candidate molecular mediators (rough ranking)  
Top analyte candidates for mediating the effect of the COCOA intervention on cognitive amelioration. Two sort orders for candidates are reflected in this table. The first list is derived from the primary “full” analysis including all data points. The second list is derived from a secondary “robust” analysis excluding outlier analyte values. The full analysis detects mediators that may be present in a minority of individuals but could play a big role in those individuals. Analytes with nominally significant effects for both the COCOA intervention and MPI correlation are listed first; following are those with at least a trend in both analyses, followed by the remaining analytes sorted by significance. The rough rankings in this table do not consider direction of effect or baseline differences between arms, which are both considered in the rankings in Table 5. For a comprehensive overview, consult Supplemental Table 1, and sort by the column(s) of interest.
Table 5. Candidate molecular mediators for multimodal interventions for individuals in the early stages of the AD spectrum. Candidates from Table 4 were pruned to produce this list of “best candidates”. First, candidates from the “full” and “robust” lists were combined. Then, two filters were employed: (1) analytes with an effect direction on MPI opposite to the COCOA intervention effect direction were excluded; (2) analytes with a higher likelihood of inflated significance for the effect of the intervention due to regression towards the mean were excluded (i.e., those with p-values < 0.05 for a difference between arms at baseline).
Supplemental Figure 1. CONSORT Flow Diagram for the COCOA Trial. Although there was some attrition during the trial, most participants stayed active long enough to contribute molecular data to at least two timepoints.
Supplemental Figure 2. Participation intervals for each participant. Attrition rates were comparable between the intervention (coaching) and control (standard of care) arms. 24-month completion rates rose slightly for individuals recruited later into the trial, possibly due to subtle effects of increased trial staff experience facilitating logistics and participant experience in the trial. The first vertical dotted line indicates the cessation of Arivale operations, and consequently the cessation of recruitment. The second vertical dotted line indicates the onset of the COVID pandemic. Neither of these events significantly increased the attrition (dropout) rate.
Supplemental Figure 3. Data collection timing for blood draws. Each participant had blood draws at up to 6 timepoints. Blood draws paused temporarily after Arivale ceased operations, as logistics were assumed in house. Blood draws also became infrequent during peak COVID pandemic waves. Some blood draws were skipped during these periods. Metabolomics, proteomics, and clinical labs were assayed from these blood draws. The original trial protocol aimed to collect blood at six timepoints: 0, 4, 8, 12, 18, and 24 months.
Supplemental Figure 4. Linearized MPI trajectories of COCOA Participants. The coaching arm has less decline in cognition than the standard-of-care arm. Although a linear fit does not fully capture the dynamics of the trajectories, this standard approach had been selected to be the primary test of the primary outcome measure. Some final MPI assessments were delayed past 24 months due to the COVID pandemic. Trajectories are computed with linear regression. Change from baseline is graphed. 90% confidence intervals are shaded.
Supplemental Figure 5. Individual MPI trajectories of COCOA Participants. Individual trajectories are highly variable. Although the COCOA intervention shows a benefit in aggregate in this population, any particular individual may not have an overall benefit, and any benefit may wax and wane.
Supplemental Figure 6. Linearized MoCA trajectories of COCOA Participants. The coaching arm has less decline in cognition than the standard-of-care arm. Trajectories are computed with linear regression. 90% confidence intervals are shaded.
Supplemental Figure 7. Linearized FAST trajectories of COCOA Participants. The coaching arm has less decline in function than the standard-of-care arm. Trajectories are computed with linear regression. 90% confidence intervals are shaded.
Supplemental Figure 8. Individual FAST trajectories of COCOA Participants. There are a greater number of intervals showing improvement (decrease of a point) in the coaching arm than in the standard-of-care arm. The COCOA intervention may occasionally reverse the process of disease in some individuals. Points are jittered to avoid overprinting.
Supplemental Figure 9. MoCA trajectory in ADNI. The ADNI MoCA trajectory can be compared to the MPI trajectory of the COCOA control arm, using a conversion factor. The ADNI MoCA trajectory is linear for the first two years, but becomes somewhat nonlinear beyond that (not shown). 90% confidence interval shaded.
Supplemental Figure 10. Relationship of kynurenate change to MPI change in COCOA. A single individual with one of the largest drops in kynurenate shows the largest increase in cognition. The remaining individuals show a smaller (and much less significant) anticorrelation suggesting that kynurenate may play a large role in a single individual, and a small role in many individuals (sex-adjusted p-value with outlier = 0.025; without outlier = 0.17), highlighting the multimodal nature of Alzheimer’s disease (AD) portrayed in Figure 1. It is possible that better targeting of kynurenate-implicated pathways with an improved intervention could better leverage this mediator in a larger number of affected individuals.