A framework for selection of blood-based biomarkers for geroscience-guided clinical trials: report from the TAME Biomarkers Workgroup

Jamie N. Justice · Luigi Ferrucci · Anne B. Newman · Vanita R. Aroda · Judy L. Bahnson · Jasmin Divers · Mark A. Espeland · Santica Marcovina · Michael N. Pollak · Stephen B. Kritchevsky · Nir Barzilai · George A. Kuchel

Received: 2 August 2018 / Accepted: 15 August 2018 © American Aging Association 2018

Abstract Recent advances indicate that biological aging is a potentially modifiable driver of late-life function and chronic disease and have led to the development of geroscience-guided therapeutic trials such as TAME (Targeting Aging with MEtformin). TAME is a proposed randomized clinical trial using metformin to affect molecular aging pathways to slow the incidence of age-related multi-morbidity and functional decline. In trials focusing on clinical end-points (e.g., disease diagnosis or death), biomarkers help show that the intervention is affecting the underlying aging biology before sufficient clinical events have accumulated to test the study hypothesis. Since there is no standard set of biomarkers of aging for clinical trials, an expert panel was convened and comprehensive literature reviews conducted to identify 258 initial candidate biomarkers of aging and age-related disease. Next selection criteria were derived and applied to refine this set emphasizing: (1) measurement reliability and feasibility; (2) relevance to aging; (3) robust and consistent ability to predict all-cause mortality, clinical and functional outcomes; and (4) responsiveness to intervention. Application of these

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11357-018-0042-y) contains supplementary material, which is available to authorized users.

J. N. Justice (✉) · S. B. Kritchevsky Internal Medicine Section on Gerontology and Geriatrics, and the Sticht Center for Healthy Aging and Alzheimer’s Prevention, Wake Forest School of Medicine, 1 Medical Center Blvd, Winston-Salem, NC 27157, USA e-mail: jnjustic@wakehealth.edu

L. Ferrucci National Institute on Aging, National Institutes of Health, Baltimore, MD 21224, USA

A. B. Newman Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15260, USA

V. R. Aroda Department of Medicine, Division of Diabetes, Endocrinology, and Hypertension Brigham and Women’s Hospital, Harvard Medical School, Boston, MA 02115, USA

J. L. Bahnson · J. Divers · M. A. Espeland Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem, NC 27157, USA

S. Marcovina Division of Metabolism, Endocrinology, and Nutrition, University of Washington, Seattle, WA 98109, USA

M. N. Pollak Department of Oncology, Jewish General Hospital, McGill University, Montreal, Quebec H3T1E2, Canada

N. Barzilai Department of Medicine, Institute for Aging Research, Albert Einstein College of Medicine, Bronx, NY 10461, USA

G. A. Kuchel UConn Center on Aging, University of Connecticut School of Medicine, Farmington, CT 06030, USA

Published online: 27 August 2018
selection criteria to the current literature resulted in a short list of blood-based biomarkers proposed for TAME: IL-6, TNF-α-receptor I or II, CRP, GDF15, insulin, IGF1, cystatin C, NT-proBNP, and hemoglobin A1c. The present report provides a conceptual framework for the selection of blood-based biomarkers for use in geroscience-guided clinical trials. This work also revealed the scarcity of well-vetted biomarkers for human studies that reflect underlying biologic aging hallmarks, and the need to leverage proposed trials for future biomarker discovery and validation.

Keywords Biomarkers · Aging · Metformin · Randomized controlled trial · Epidemiology · Mortality · Inflammation

Introduction

The geroscience hypothesis holds that a specific set of shared biological mechanisms of aging increases the susceptibility of aged individuals to several chronic diseases and loss of function and that therapies developed to target such shared “drivers” have the potential to delay the onset and progression of multiple chronic diseases and functional decline. The inter-related cellular biologic processes that drive the biology of aging are known as “pillars” or “hallmarks” of aging, and accumulating evidence from mammalian animal models supports the premise that geroscience-guided interventions targeting these processes can extend healthspan and lifespan (Barzilai et al. 2016; Burch et al. 2014; Espeland et al. 2017; Kennedy et al. 2014; Longo et al. 2015; Sierra 2016a). A new generation of clinical trials is being designed to test the geroscience hypothesis in humans (Justice et al. 2016; Newman et al. 2016b; Sierra 2016b). One example is the proposed multicenter clinical trial, Targeting Aging with MEtformin (TAME), which will evaluate whether metformin, a commonly used diabetes drug that also targets the biology of aging, can (1) prevent or delay the incidence of multiple age-related chronic diseases, (2) help maintain function, and (3) influence biological markers of aging in older persons (Barzilai et al. 2016). To accomplish this latter aim, a strategy to select the most appropriate biomarkers of aging to be included in geroscience-guided clinical trials is needed.

Hypothetically, a biomarker of aging should reflect the underlying biology, and a change in biomarker levels should have parallel changes that occur in the susceptibility to disease and loss of function. Thus, interventions targeting aging should result in changes in biomarkers that will eventually delay the incidence, accumulation, clinical evolution, and functional consequences of chronic age-related diseases. One of the critical roles played by biomarkers could be that of surrogate endpoints reflective of risk and progression of several major diseases. In support of this role, the US Food and Drug Administration (FDA) and the Institute of Medicine highlighted the importance of biomarkers as surrogate measures in drug development and trials involving chronic diseases such as cancer and heart disease; however, few existing biomarkers have sufficient clinical evidence of association with the rate of development of aging phenotypes and multi-morbidity (IOM 2010; Biomarkers Definitions Working 2001). The FDA currently does not consider aging an indication for drug development or labeling, which may also impede the emergence of knowledge supporting discovery and validation of geroscience-relevant biomarkers. As a result, there are at this time no approved or commonly accepted biomarkers of aging for clinical trials, nor is there a consensus set of validated biomarkers of the biologic pillars or hallmarks of aging that would be applicable to clinical research. This poses a major translational gap that must be bridged in order to facilitate scientific progress.

The purpose of this report is to outline a conceptual framework and evidence-based approach to the prioritization and selection of a panel of blood biomarkers for use in randomized controlled clinical trials of a geroscience therapy. A multi-disciplinary Biomarkers Workgroup convened for a planning workshop and met weekly by phone over an 8-month period. The Biomarkers Workgroup defined biomarker criteria, prioritized selection parameters, and provided guidance for resource development and inclusion of discovery-based platforms. The development of the TAME trial served as an opportunity to apply the selection criteria to putative blood-based biomarkers identified through an exhaustive literature review. The result is an evidence-based short list of proposed biomarkers for the TAME trial, and a framework that could be applied to next-generation clinical trials targeting aging.
Conceptual framework

Biomarker definitions According to consensus definitions by Baker and Sprott (Baker III and Sprott 1988; Sprott 1988, 2010), and the American Federation for Aging Research (AFAR), biomarkers of aging are measures of a biological parameter that, either alone or as a multivariate composite, monitor a biological process underlying aging rather than effects of a specific disease; predict the rate of aging and mortality better than chronological age; can be safely tested across repeat measures in the same organism; and work in humans and in laboratory animals such as mice. The Biomarkers Workgroup adapted these definitions to develop selection criteria tailored to the context of geroscience-guided randomized clinical trials:

1. Measurement reliability and feasibility. The biomarker should be feasible to measure in a clinical trial without incurring undue risk to human subjects, and meet trial specific reliability requirements (see “Trial Context” below), with standardized measurement.

2. Represent biologic aging processes. The biomarker should have face validity such that it represents a process or processes relevant to biologic aging hallmarks, and changes in a measurable and consistent manner with chronological age.

3. Robust and consistent association with risk of death, and clinical/functional trial endpoints. Association with risk. The biomarker should be consistently associated with increased risk of clinical and functional endpoints including all-cause mortality even when controlling for chronological age. Ideally, the biomarker would be involved in the causal pathway, such that direct manipulation of the biomarker changes the associated risk, but at minimum, the biomarker level would move in a direction that predicts the clinical or functional outcome. Robust. The changes in biomarker level with age and associations with risk should be robust across species, datasets, or populations.

4. Responsive to intervention. A biomarker of aging for geroscience-guided trials should be responsive to interventions that affect the biology of aging ideally over a relatively short period of time. A quick response to intervention would allow for shorter trials for fully vetted biomarkers.

The criteria above were developed for research studies in humans. Foundational efforts to identify and categorize the cellular and molecular “hallmarks” of aging introduced a host of potential biomarkers for mammalian aging based on evidence in mouse models and some simpler organisms such as Caenorhabditis elegans and Drosophila melanogaster. Clinical translation of these biomarkers can be problematic, with barriers such as access to tissues, environmental or genetic control, and use of resource and assays that are not feasible in clinical research. However, reports often mix evidence from animal models and humans indiscriminately, and aside from a few thoughtful reviews and studies, relatively, little work has emerged on measures of the biologic “hallmarks” of aging specifically for human research (Burkle et al. 2015; Khan et al. 2017; Rochon et al. 2011). Accordingly, the present framework presented focuses solely on markers that can be measured in humans. Though focused on human research, analogous frameworks to reverse translate for preclinical testing in mammalian species such as rodent, dog, and nonhuman primate can be envisioned.

Trial context This framework should be tailored to the specific clinical trial design or investigational drug being used. It is broad enough to include biochemical assays, clinical measures, imaging, and physiological tests, such as gait speed, grip strength, cognitive assessments, spirometry, and blood pressure. Specific biomarkers selected or types of biomarkers considered depend primarily on the context of the trial. Importantly, the determination of the biomarker must be feasible for the population, size, duration, budget, and logistical constraints of the clinical trial. Trial context dictates:

- Feasibility within trial design:
  - Acceptable additional risk to participants for biomarker determination
  - Inclusion of proposed measures within clinical visits and resource availability

- Reliability of assays:
  - Accuracy of assay and agreement across technical replicates
  - Assay short-term test-retest reliability (correlation $\geq 0.7$)
Reliability and reproducibility of measurement across trial sites or laboratories

- Sensitivity to detect change:
  - Assay detection limits in specific study population
  - Within-subject variation over study duration (e.g., stability over months, years)
  - Estimated intervention effects on biomarker measure.

Changes in biomarker levels should be consistently related to changes in risk of mortality, disease, or functional outcomes and should be reasonably robust to confounders and common medical maneuvers in those recruited. This requires careful consideration of the specific population, including age and sex specific biomarker reference ranges, effects of co-morbid conditions, and concurrent use of common medications. An example is low-density lipoprotein cholesterol (LDL-C), which is a prominent biomarker of atherosclerotic heart disease, and in middle-aged adults, elevated levels of LDL-C are associated with greater risk of cardiovascular related events and mortality. However, at advanced ages, the converse is true, and low levels of LDL-C may be related to higher risk. Moreover, commonly prescribed medications to control lipid levels could result in a change in LDL-C and related biomarkers that are independent of the investigational drug and not reflective of a change in underlying aging biology (High and Kritchevsky 2015).

**Biomarker categories** The Biomarkers Workgroup identified three primary biomarker categories consistent with models proposed by NIH Biomarker Definitions Working Groups and FDA guidance, markers of the (A) investigational drug, (B) underlying biology, and (C) clinical disease outcomes. Biomarkers of the investigational drug can contribute knowledge about clinical pharmacology and inter-individual variation in responses to treatments, and can include circulating measures of levels of the drug or its relevant metabolites, or known drug-specific effects that could mediate effects on trial outcomes. Markers of underlying biology provide proof of concept and mechanistic insight, and may suggest future therapeutic candidates. Biomarkers of the clinical disease(s) being targeted or population studied may provide early indicators of drug effects on clinical disease and could serve as surrogate trial endpoints. The final selection of biomarkers addressing effects of the investigational drug and clinical outcomes should be matched to the unique features of each individual trial, yet overarching features of biomarkers linking the underlying biology to clinical outcomes should have a degree of consistency across all proposed geroscience-guided clinical trials. The present report is focused on those features of blood-based biomarkers of biologic aging or age-related diseases that may be generalizable to other geroscience-guided trials.

**Case study: Targeting Aging with MEtformin (TAME)**

TAME is a proposed 6-year double blind placebo-controlled randomized trial of metformin involving 3000 nondiabetic men and women aged 65–80 years to be recruited across 14 US-based sites. Metformin was selected based on its effects on biological hallmarks of aging in cells and animal models: metformin inhibits the mitochondrial complex I in the electron transport chain and reduces endogenous production of reactive oxygen species (ROS) (Batandier et al. 2006; Bridges et al. 2014; Kickstein et al. 2010); activates of AMP-activated kinase (AMPK) (Cho et al. 2015; Duca et al. 2015; Zheng et al. 2012), decreases insulin/insulin-like growth factor-1 (IGF-1) signaling (Barzilai et al. 2016; Nair et al. 2014) (Foretz et al. 2010, 2014), reduces DNA damage (Liu et al. 2011); inflammation and the senescence associated secretory phenotype (Lu et al. 2015, 2014; Moiseeva et al. 2013; Saiso 2015). When administered in vivo in rodents, the lifespan effects of metformin alone are either not observed (Smith Jr. et al. 2010; Strong et al. 2016), or relatively modest (~4–6% extension of median lifespan) (Martin-Montalvo et al. 2013). However, the effects on health are substantial, with improvements on tests of physical and cognitive function, cataracts, oral glucose, and insulin tolerance improved by up to 30% (Allard et al. 2015, 2016; Allard et al. 2015; Martin-Montalvo et al. 2013). These findings are coupled by observation that in persons with diabetes, the use of metformin is associated with lower rates of cancer (Landman et al. 2010; Lee et al. 2011; Libby et al. 2009), cardiovascular risk factors and events (Abualsuod et al. 2015; Kooy et al. 2009), dementia (Luchsinger et al. 2015);
and all-cause mortality (Bannister et al. 2014; Johnson et al. 2005; Roussel et al. 2010; Schramm et al. 2011). TAME was conceptualized as a prototype geroscience-guided trial using metformin to target clinical outcomes of aging. Main trial outcomes are the incidence of (1) death or any new age-related chronic disease (myocardial infarction, stroke, hospitalized heart failure, cancer, dementia or mild cognitive impairment, multimorbidity) and (2) major age-related functional outcomes (major decline in mobility or cognitive function, or onset of activities of daily living limitation). Biomarkers of aging comprise an exploratory trial outcome, and we hypothesize that metformin’s beneficial effects, if observed, will be associated with markers of biologic aging.

As there is currently no consensus on what biomarkers of aging should be preferentially addressed in geroscience-guided trials, a Biomarkers Workgroup was convened for a planning workshop (NIA; Baltimore, MD) and met weekly by phone for 8 months (Oct 2017–May 2018). The workgroup consisted of experts in the basic biology of aging, metformin pharmacology, gerontology, biostatistics, epidemiology, endocrinology, and geriatric medicine (see author list for participating workgroup members). The workgroup led defined biomarker parameters and conducted an exhaustive search to pre-specify biomarkers and rigorously apply the trial biomarker criteria. An overview of the process of biomarker identification, prioritization and selection, and proposed list of biomarkers is shown in Fig. 1.

Candidate biomarker identification A total of 258 potential biomarkers of aging were identified by input from individual members of the Biomarkers Workgroup. Additionally, literature was reviewed to identify a set of biomarkers of biological aging: published multi-assay composites (Belsky et al. 2015; Belsky et al. 2017a; Fried et al. 2001; Howlett et al. 2014; Li et al. 2015; Mitnitski et al. 2013, 2015; Mitnitski and Rockwood 2015; Sanders et al. 2014; Sebastiani et al. 2017), consensus-derived panels (Burkle et al. 2015; Engelfriet et al. 2013; Jylhava et al. 2017; Khan et al. 2017; Lara et al. 2015; Wagner et al. 2016; Xia et al. 2017), and large aging studies (Martin-Ruiz et al. 2011; Rochon et al. 2011) were consulted and 229 candidate biomarkers identified. An additional 29 recognized biomarkers of TAME’s clinical cardiovascular, cancer, and cognitive outcomes were identified (Supplement Material 1).

Exclusions and prioritization Sixty-seven biomarkers of aging that were not blood-based (e.g., imaging, physiologic) were omitted from consideration as several functional measures were already considered for determination of secondary trial outcome. Thirty-nine markers were excluded based on participant or resource burden, low feasibility, or assay reliability concerns (Supplemental Material 2). For example, measures that
require access to cells derived from standard blood draw (e.g., CD4/CD8 T cell ratio, T cell p16INK4a expression, mitochondrial respirometry) may be ideal biomarkers of aging, but have low feasibility for large trials due to significant resource burden, need of skilled laboratory personnel and specialized equipment that may not be readily available or easily standardized across clinical trial sites. Other biomarkers demonstrate uncertain assay reliability or validity. These include circulating growth/differentiating factor 11 (GDF11) assays using antisera with cross reactivity issues (Rodgers 2016; Rodgers and Eldridge 2015) or require a highly specific immunoplexed LC-MS/MS assay which is reliable but may not be feasible for large-scale trials such as TAME (Schafer et al. 2016). Additionally, several cytokines (e.g., interleukin-2, interleukin-1β, interferon-γ) demonstrate inconsistent detectability resulting from analyte degradation in long-term storage, or low assay sensitivity (McKay et al. 2017).

The remaining 86 candidate biomarkers were ranked based on frequency of appearance in the literature and weighted by strength of expert opinion and utility in monitoring disease outcomes (see Supplemental Material 1).

1. Frequency of use: appearance in 17 consulted publications was tallied.
2. Utility in diagnosing or monitoring disease: 48 biomarkers of clinical importance for clinical disease evaluation were noted from FDA guidance documents or disease association statements (e.g., American Heart Association). For the CVD endpoints MI, stroke, and CHF, 33 total biomarkers were identified (18 common to aging biomarkers list, 15 new) (Chow et al. 2017; Jickling and Sharp 2015; Thygesen et al. 2012). Two FDA-recognized markers for early AD or MCI were identified (Administration 2018), and 11 for cancer prognosis, staging, or disease monitoring.
3. Expert opinion: markers that appeared in the literature were weighted based on strength of Biomarker Workgroup expert suggestion: suggested exclusion (−), unmentioned (), consideration (+), recommended (++), and strongly recommended (+++).

Biomarker selection The top 20 ranked candidate biomarkers were evaluated according to the Biomarker Workgroup-identified criteria. The Biomarker Workgroup evaluated face validity of biomarker, and considerations related to feasibility and potential confound by common medical conditions or treatments. Literature reviews using Pubmed were conducted for each biomarker to evaluate association with risk, robustness, and responsiveness to relevant interventions (overview Table 1, and full listings in Supplemental Material) with filters for (1) age (≥ 45 years), (2) prospective studies, and (3) human or clinical research. Direction of associations and magnitude of effects across publications, datasets, and populations were tracked.

- **Association with risk of clinical disease or functional decline/disability onset:** PubMed search strategies were used to evaluate association of each individual candidate biomarker with risk of clinical events, disease-related mortality, disability, or functional declines and all-cause mortality (Supplemental Material 3). In addition, separate searches for biomarker with each clinical disease (CVD, MCI/AD, cancer) and functional outcome (mobility, disability, frailty) were also conducted and evidence of associations noted. Studies in populations with acute or severe diseases were excluded. Given wide differences in outcomes, populations, and model adjustments, the estimated effects sizes were not pooled or systematically summarized.

- **Associations with risk of death:** PubMed searches used: <selected biomarker> AND “mortality” OR “all-cause” OR “death” OR “lifespan.” Publication reference lists and the website MortalityPredictors.org were consulted to identify additional relevant publications. A detailed listing of studies is included in Supplemental Material 3. Estimated effect sizes of biomarker’s association with all-cause death were summarized as age-adjusted hazard ratios (HR), with range HRs, and number of studies age-adjusted models considered listed (Table 1).

- **Responsiveness to interventions:** PubMed search and published data from the Diabetes Prevention Program (DPP) were consulted to determine whether the biomarker of interest was sensitive to change in less than 6 years when exposed to interventions of interest. Geroscience-identified interventions were searched (e.g., metformin, caloric restriction). If data were available, the percent change in the candidate biomarker with metformin treatment was evaluated compared to reference group or placebo control (Supplemental Material 4).
<table>
<thead>
<tr>
<th>Rank</th>
<th>Candidate biomarker</th>
<th>Underlying process</th>
<th>Aging biology</th>
<th>Robust</th>
<th>Mortality HR [range], # studies</th>
<th>Clinical</th>
<th>Functional</th>
<th>Treatment response</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Meets criteria</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>IL-6</td>
<td>Inflammation</td>
<td>+++</td>
<td>+++</td>
<td>1.66 [1.0–2.8], 18</td>
<td>CVD, cancer MCI</td>
<td>Phys, cog</td>
<td>+++</td>
</tr>
<tr>
<td>2</td>
<td>CRP</td>
<td>Inflammation</td>
<td>++</td>
<td>+++</td>
<td>1.63 [1.0–2.6], 14</td>
<td>CVD, cancer</td>
<td>Phys, cog</td>
<td>++</td>
</tr>
<tr>
<td>3</td>
<td>TNFα</td>
<td>Inflammation</td>
<td>+++</td>
<td>+++</td>
<td>1.54 [1.2–2.5], 7</td>
<td>CVD, cancer, MCI/dementia</td>
<td>Phys, cog</td>
<td>+++</td>
</tr>
<tr>
<td>4</td>
<td>Cystatin C</td>
<td>Kidney aging</td>
<td>+++</td>
<td>++</td>
<td>1.84 [1.3–4.4], 13</td>
<td>CVD</td>
<td>Phys, cog</td>
<td>+</td>
</tr>
<tr>
<td>5</td>
<td>IGF-1</td>
<td>Nutrient sign</td>
<td>+++</td>
<td>U</td>
<td>High: 1.30 [0.7–1.7], 10</td>
<td>Cancer</td>
<td>Phys.</td>
<td>–</td>
</tr>
<tr>
<td>6</td>
<td>NT-proBNP</td>
<td>CardioVasc.</td>
<td>+</td>
<td>+++</td>
<td>1.38 [1.2–1.4], 3</td>
<td>CVD</td>
<td>Phys.</td>
<td>+</td>
</tr>
<tr>
<td>7</td>
<td>Insulin</td>
<td>Nutrient sign</td>
<td>++</td>
<td>++</td>
<td>1.21 [1.0–1.5], 5</td>
<td>CVD</td>
<td>–</td>
<td>++</td>
</tr>
<tr>
<td>8</td>
<td>GDF15</td>
<td>Stress respn</td>
<td>+++</td>
<td>+++</td>
<td>2.24 [1.5–4.0], 11</td>
<td>CVD</td>
<td>Phys.</td>
<td>Metformin</td>
</tr>
<tr>
<td>13</td>
<td>IGFBPs</td>
<td>Nutrient sig.</td>
<td>+++</td>
<td>U</td>
<td>High: 1.36 [0.6–3.0], 6</td>
<td>CVD, cancer</td>
<td>Phys.</td>
<td>++</td>
</tr>
<tr>
<td><strong>Does not meet criteria</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>DHEAS</td>
<td>Endocrine</td>
<td>+</td>
<td></td>
<td>1.54 (men), 1.18 (all)</td>
<td>CVD</td>
<td>Phys. (M)</td>
<td>–</td>
</tr>
<tr>
<td>10</td>
<td>Telomere length</td>
<td>Telomeres</td>
<td>+++</td>
<td>+/-</td>
<td>(+/-) 1.07 [0.7–1.4], 10</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>11</td>
<td>Adiponectin</td>
<td>Obesity</td>
<td>–</td>
<td>+, U</td>
<td>1.33 [1.1–1.9], 14</td>
<td>CVD, MCI/dem</td>
<td>Phys, cog</td>
<td>–</td>
</tr>
<tr>
<td>12</td>
<td>Isoprostanes**</td>
<td>Ox. stress</td>
<td>++</td>
<td>–</td>
<td>HR 1.1; n = 1,</td>
<td>CVD (+/-)</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>14</td>
<td>Thyroid hormone</td>
<td>Endocrine</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>++</td>
</tr>
<tr>
<td>15</td>
<td>CMV antibody</td>
<td>Immune</td>
<td>–</td>
<td>–</td>
<td>1.26 [1.1–1.4], 3</td>
<td>–</td>
<td>Frailty</td>
<td>–</td>
</tr>
<tr>
<td>16</td>
<td>Leptin</td>
<td>Obesity</td>
<td>–</td>
<td>–</td>
<td>0.98 [0.7–1.2], 4</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>17</td>
<td>CCL11 (eotaxin)</td>
<td>Inflammation</td>
<td>++</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>18</td>
<td>Fibrinogen</td>
<td>CardioVasc</td>
<td>+</td>
<td>+/-</td>
<td>1.20 [1.1–1.4], 3</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>19</td>
<td>AGE/rAGE</td>
<td>CardioVasc</td>
<td>++</td>
<td>+/-</td>
<td>+/-</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>20</td>
<td>tPA</td>
<td>CardioVasc</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>CVD</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td><strong>Additional consideration</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Epigenetic clocks^</td>
<td>Epigenetics</td>
<td>+++</td>
<td>+++</td>
<td>1.39 [1.1, 2.2], 6</td>
<td>Cancer, CVD</td>
<td>Frailty (+/-)</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>HbA1c</td>
<td>Metabolic</td>
<td>+++</td>
<td>++</td>
<td>1.45 [1.1, 2.5], 12</td>
<td>CVD</td>
<td>–</td>
<td>++</td>
</tr>
</tbody>
</table>

**Stronger evidence may exist for isoprostanes measured in urine**

^Includes multiple algorithms for DNA methylation scores and age acceleration
Based on this systematic process, 8 of the 258 pre-specified blood-based biomarkers remained as candidate markers to use as an exploratory outcome:

- Inflammation: interleukin-6 (IL-6), tumor necrosis factor α receptor II (TNFRII), high sensitivity c-reactive protein (CRP)
- Stress response and mitochondria: growth differentiating factor 15 (GDF15)
- Nutrient signaling: fasting insulin, insulin-like growth factor 1 (IGF-1)
- Kidney aging: cystatin C
- Cardiovascular: N-terminal B-type natriuretic peptides (NT-proBNP)
- Metabolic aging: hemoglobin A1c

Each biomarker and its role are briefly explained in a graphical table (Fig. 2). A few specific comments on selections are provided here: TNFα and TNF receptors (I, II) were examined; however, TNFα receptors (e.g., TNFRII) were workgroup recommended to minimize analytic variability, given the fact that serum TNFα serum levels tend to be low and unstable with storage at −80 °C (Barron et al. 2015; Cesari et al. 2003; Marti et al. 2014). Moreover, IL-6, CRP, and TNFα are commonly used and are independently associated with mortality risk (Bruunsgaard et al. 2003; Lio et al. 2003; Penninx et al. 2004; Reuben et al. 2002; Roubenoff et al. 2003; Stork et al. 2006), but the combination of IL-6 and TNF receptor levels has been shown to perform particularly well when combined as a pro-

### Blood-based biomarkers for geroscience-guided trials

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Underlying Biologic Process &amp; Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL-6, CRP,</td>
<td>Inflammation &amp; Intercellular Signaling Interleukin 6 (IL-6) is a proinflammatory cytokine and Tumor Necrosis Factor-α Receptor II (TNFRII) is a TNF-α receptor involved in acute-phase response. C-Reactive Protein (CRP) is an acute phase protein produced in response to inflammation. Cytokine dysregulation is a driver of pathophysiologic processes leading to disease, functional decline, frailty, and death.</td>
</tr>
<tr>
<td>TNFRII</td>
<td></td>
</tr>
<tr>
<td>GDF15</td>
<td>Stress Response &amp; Mitochondria Growth Differentiating Factor 15 (GDF15) is a member of the TGF-β superfamily robustly associated with mortality, cardiovascular events, cognitive decline and dementia. GDF15 is increasingly recognized in mitochondrial dysfunction, and as a biomarker of aging.</td>
</tr>
<tr>
<td>IGF-1, Insulin</td>
<td>Nutrient Signaling Disruption of the insulin/insulin-like growth factor (IGF-1) signaling pathway is implicated in longevity in animal models. In humans, IGF-1 and fasting insulin are responsive to caloric restriction, and low IGF-1 in growth hormone receptor deficiency conveys disease protection.</td>
</tr>
<tr>
<td>Cystatin-C</td>
<td>Kidney Aging Cystatin C, an extracellular inhibitor of cysteine proteases, is a marker of renal disease and aging. It is an independent risk factor for all cause and CVD-related mortality, and multi-morbidity, and higher levels are consistently associated with poor physical function and cognition.</td>
</tr>
<tr>
<td>NT-proBNP</td>
<td>Cardiovascular Health B-type natriuretic peptides (BNP, NT-proBNP) are secreted in response to cardiomyocyte stretching to decrease vascular resistance. NT-proBNP has a greater-half life and accuracy compared with BNP and is used to diagnose and establish prognosis for heart failure.</td>
</tr>
<tr>
<td>HGBA1c</td>
<td>Metabolic Aging Glycated hemoglobin (hemoglobin A1c; HGBA1c) is formed in a non-enzymatic glycation pathway and is a marker for 3-mo average plasma glucose. High HGBA1c reflects poor glucose control, and in older nonobestics is strongly associated with death, chronic disease, and functional decline.</td>
</tr>
<tr>
<td>Molecular</td>
<td>Epigenetic, Interdependent, Multi-Omic Data intensive molecular platforms can explore global changes in epigenetic, transcriptomic, proteomic and proteostasis, and small metabolite signatures. These approaches may better capture complex and multifactorial processes underlying aging.</td>
</tr>
<tr>
<td>Signature</td>
<td></td>
</tr>
</tbody>
</table>

inflammatory cytokine score predicting the risk of mortality and mobility disability is elevated (Varadhan et al. 2014). While the biomarkers above generally meet the workgroup-derived criteria, IGF-1 does not: the relationship between IGF-1 and mortality/frailty is U-shaped with both high and low levels associated with all-cause mortality and adverse health outcomes (Andreasen et al. 2009; Burgers et al. 2011; Cappola et al. 2003; Doi et al. 2016; Friedrich et al. 2009; Hu et al. 2009; Kaplan et al. 2007; Laughlin et al. 2004; Leng et al. 2009; Maggio et al. 2007; Roubenoff et al. 2003; Saydah et al. 2007; van der Spoel et al. 2015), which could complicate the interpretation of change with intervention. However, ample evidence indicates its importance to biological aging, including use as a target for intervention to improve healthspan and lifespan in mice (Mao et al. 2018). Hemoglobin A1c (HbA1c, HGBA1c) was selected as a marker of metformin’s potential glycemic effects in TAME and therefore excluded from consideration as a biomarker of aging. In TAME, diabetes was excluded as an inclusion criterion or outcome measure, and even in a non-diabetic population, the effects of metformin on glucose metabolism may be difficult to separate from those on biological aging. Nevertheless, HbA1c is a prominent biomarker of aging and generally meets workgroup-identified criteria; therefore, in other studies, its use should be considered as a marker of metabolic aging (Dubowitz et al. 2014; Palta et al. 2017; Pani et al. 2008). Innovative “omics”-based approaches and epigenetic markers did not meet selection criteria, but the Biomarkers Workgroup noted unique advantages as biomarkers of aging and represent an area ripe for development for future trials. For example, the stress response and mitochondrial marker GDF15 did not appear in literature searches for biomarkers of aging, but was instead identified by experts on the workgroup based on proteomic analyses, and ample evidence as a marker of metformin (Gerstein et al. 2017), association with risk of mortality (Wiklund et al. 2010), cardiovascular disease, and heart failure, and biology of aging as a senescence-associated secretory protein and marker of systemic energy homeostasis (Chung et al. 2017; Kim et al. 2018). This example underscores the importance of sustained vigilance for developments from various platforms for proteomics, metabolics, and other techniques and store biologic samples for future assays work in this area.

Discussion

Geroscience research has shown that aging has a distinct biology and that this biology can be targeted in order to extend health and longevity in animal models. As a result, geroscience-guided clinical trials such as TAME designed to test the geroscience hypothesis in humans are being planned. This report presents a conceptual framework for inclusion of biomarkers in such clinical trials and uses TAME as the test case to apply this framework. Exhaustive efforts were undertaken to arrive at a concise list of well-justified biomarkers for use in a clinical trial (Fig. 2), including markers of inflammation (IL-6, TNFRII, CRP), stress response and mitochondrial health (GDF15), nutrient signaling (insulin, IGF-1), multisystem disease markers of kidney aging (cystatin C), cardiovascular health (NT-proBNP), and metabolic aging (hemoglobin A1c). The present review and discussion highlight the paucity of markers reflecting the biologic hallmarks of aging validated for use in clinical research, and the importance of including data-intensive approaches to drive biomarker discovery and uncover new potential therapeutic targets for next-generation geroscience trials.

Notable biomarker exclusions The Biomarker Workgroup initially proposed to include a biomarker from each of the 7 to 9 geroscience-identified biologic “hallmarks” or “pillars” of aging, but soon concluded that this approach was not practical or feasible at this time. Many markers require access to tissues or resources that are not feasible for large clinical trials. Relatively, few markers can be directly measured from samples obtained from blood draw, or have validated and reliable blood-based surrogates. Several other putative biomarkers of aging were not proposed as candidate biomarkers due to failure to meet workgroup-identified selection criteria. For example, telomere length is frequently used as a biomarker of aging, with in vitro and in vivo research implicating telomere length in cellular senescence and oxidative stress (Blackburn 2000; Blasco 2007; Lopez-Otin et al. 2013; von Zglinicki 2002). However, telomere length is inconsistently or not associated at all with clinical or functional outcomes (Sanders and Newman 2013).

The epigenetic clock, a biomarker of aging strongly associated with chronological age, consistently predicts risk of mortality and key clinical outcomes (Chen et al. 2016; Horvath 2013; Horvath et al. 2016; Levine et al. 2002).
Preclinical evidence suggest epigenetic aging signatures in liver may be altered by interventions such as caloric restriction and rapamycin (Wang et al. 2017), which lends promise of epigenetic clocks for clinical trials in the future. Though epigenetic clocks may detect cross-sectional differences in diet and lifestyle, currently, responsiveness to an intervention like metformin is not well supported in epidemiologic literature (Kim et al. 2017; Quach et al. 2017). This coupled with relative expense to quantify, tipped the balance of cost-benefit. This may change in the future: costs will likely decline, and ongoing efforts are underway calibrate epigenetic clocks on physiologic age scores that may prove particularly useful as a biomarker for geroscience-guided trials (Levine et al. 2018). Continued refinement and validation of these clocks in human clinical trials are needed, and geroscience-guided trials provide excellent validation tools for biomarker development. Likewise, hypothesis-free, data-intensive molecular approaches do not meet criteria for pre-specified trial biomarkers but represent an important area for scientific development. The molecular signatures derived from interdependent processes provide insight into the complex and multifactorial processes underlying biological aging and responses to intervention. These platforms are increasingly more attractive for clinical trials as new technologies emerge and assay costs decrease. Such biomarker discovery requires carefully planned collections of sample materials and investigator engagement for ancillary studies using new techniques to evaluate novel, putative biomarkers.

Analytic considerations This review proposes a variety of a priori suggestions for biomarkers for consideration in the design of randomized controlled geroscience-guided clinical trials, yet the analytic plan is intentionally minimal. This open approach is in accord with recent FDA guidance for clinical trials on age-related diseases without consensus biomarkers (e.g., early Alzheimer’s disease). When inadequate information exists for hierarchical structuring of a series of biomarkers as a supporting outcome, FDA encourages trial sponsors to “analyze the results of these biomarkers independently, though in a prespecified fashion, with the understanding that these findings will be interpreted in the context of the state of the scientific evidence” (Administration 2018). This is particularly relevant to TAME, which was designed in a collaborative and consensus building process with FDA guidance. This open analytic strategy is also intended to encourage use of the biomarker data to develop, refine, and validate multi-assay biomarker composites for next-generation geroscience-guided trials. Ultimately, the validation of biomarkers of biological aging requires longitudinal data because, owing to compensatory and resilience mechanism; there is a temporal gap between the changes that occur in biomarkers and the changes in clinical outcomes, either disease related or functionally related. Implementation of multiple biomarkers of the “pillars” or “hallmarks” of aging is a tantamount strategy to develop the best set of measure to use in clinical trials.

Composites can be developed using combinatorial techniques such as factor analysis or principal component analysis (Karaski et al. 2012; Sebastiani et al. 2017) or weighted and summed based each markers’ prediction of death (Sanders et al. 2018). Such composites should be demonstrated to predict aging outcomes, including death and disability, independently of chronologic age and should also be superior to age itself in the strength of the association with these outcomes (Sanders et al. 2012a). Moreover, multi-assay composites and deficit accumulation indices are likely to outperform single-biomarkers as an index of biological age (Belsky et al. 2017b; Cohen et al. 2017; Evans et al. 2014; Li et al. 2015; Mitnitski et al. 2002; Sanders et al. 2014, 2018; Sebastiani et al. 2016; Sebastiani et al. 2017). Limitations of individual clinical biomarkers in epidemiological research are well known, including risk of conditional associations, misidentification of aging biomarkers, oversimplification of the basic biology/physiology of the aging organism, and poor generalizability and reproducibility (Cohen et al. 2017). Of interest, a systematic evaluation of 11 markers of aging, including telomere length, epigenetic clocks, and clinical biomarker composites, confirms that biomarker composite measures were most consistently related to physical and cognitive function compared with any one biomarker (Belsky et al. 2017b; Newman et al. 2008). Additionally, multi-assay biomarker indices are responsive to interventions targeting the biology of aging, such as caloric restriction (Belsky et al. 2017a). For example, the Healthy Aging Index (HAI), a composite score of physiologic aging that predicts mortality (O’Connell et al. 2018; Sanders et al. 2014, 2018; Wu et al. 2017), cardiovascular disease (McCabe et al. 2016), and disability (Sanders et al. 2012b), is improved with weight loss by caloric restriction in older adults (Shaver et al.
The effect of caloric restriction on the HAI was greater than any individual component or candidate biomarker, and reflects a meaningful difference: the net reduction of HAI by 0.63 points translates to an approximate 9% reduction in mortality risk. In sum, accumulating evidence supports the utility of multicomponent biomarker scores and deficit accumulation indices for future geroscience-guided clinical trials.

**Biomarker discovery and resource development** A panel of a priori defined biomarkers has inherent utility for geroscience-guided trials but may not capture the complex and multifactorial processes underlying aging. This gap can be addressed by high-throughput “omics” technologies. Technologies used in sequencing are dramatically reshaping research and drug development, and powerful discovery and screening technologies permit the assessment of biological parameters to permit the molecular and cellular basis for variation in response to therapy and to explain the clinical response to intervention in clinical trials (Biomarkers Definitions Working 2001). Data-intensive technologies for biomarker discovery are increasingly common and include combinatorial chemistry, mass spectrometry, high-throughput screening, cell- and tissue-based microarray, proteomic, and microbiome platforms. Biomarker strategies for geroscience-guided clinical trials are recommended to leverage such technologies as part of the overall biomarker plan. Given the cost, specialization, and invasiveness often required, it may not be feasible to include within large multicenter or multi-year trials, but opportunities for smaller sub-studies or ancillary investigations involving biomarker discovery and cell/tissue collections could be conducted at specialized sites or with lower collection frequency.

Rigorous procedures for biospecimen collections and repository curation are central to future biomarker discovery and resource development for next-generation geroscience-guided trials. The specifics of the plan are almost entirely dependent on the context of the trial, but at a minimum serum and plasma (EDTA and citrated) should be aliquoted and archived, and DNA/RNA isolated if resources warrant. To gain mechanistic insight, isolated and cryopreserved human peripheral blood mononuclear cells (PBMCs) and specialized collections are often required. Specialized collections that pose minimal additional risk to study participants should be prioritized (e.g., urine, stool, peripheral blood cells, saliva), while more invasive tissue biopsies (e.g., adipose, muscle, skin) should only be considered if benefit of knowledge to be gained is warranted and efforts to minimize risks are acceptable. Biospecimen repositories provide a unique for emerging basic research and biomarker studies. For example, the endogenous peptide apelin has recently been identified as a biomarker associated with risk of age-related functional decline and sarcopenia, and a potential drug target to prevent or restore physical function, at least in mice (Vinel et al. 2018). Apelin and other markers of interest emerging from basic research could be examined using repository samples and underscore the utility of a well-cultivated repository and importance of continued surveillance for emerging biomarkers.

**Conclusion**

With recent growth in the field of geroscience, consensus definitions and conceptual frameworks for biomarkers of aging to be used as part of a supporting outcome in clinical trials are imperative. The Biomarkers Workgroup report presents a conceptual framework for inclusion of biomarkers in geroscience-guided clinical trials and uses TAME as the test case. Conservative application of selection criteria using clinical and epidemiological literature returns a concise list of blood-based biomarkers. The Biomarkers Workgroup also identified several key considerations and areas in need of development. First, selection criteria are designed for inclusion of individual biomarkers, yet analytic approaches based on multi-assay composites or deficit accumulation/frailty indices are likely to outperform any one individual marker as trial biomarker outcomes. Second, currently, there is a striking paucity of markers that could adequately reflect the biologic hallmarks of aging in the context of human research in general and large clinical trials in particular. There is a critical need for biomarker and resource development to discover novel biomarkers and therapeutic targets, and to validate existing biomarkers, including composites of several markers for use in clinical trials on aging. Finally, data-intensive and multi-omic approaches are likely to drive biomarker discovery and uncover new potential therapeutic targets for next-generation geroscience trials.
Funding disclosures This study received financial support from the American Federation for Aging Research (AFAR); the Glenn Center for the Biology of Human Aging (Paul Glenn Foundation for Medical Research); National Institutes of Health: K01 AG059837-01 (JNJ), P30 AG021332 (SKB, JNJ); R01 AG048023, R01 AG052608, R35 GM124922 (GAK); P30 AG038072 (NB); R01 AG023629 (ABN), and supported in part by the Intramural Research Program of the National Institute on Aging, National Institute of Health.

Compliance with ethical standards

Conflicts of interest JNJ, NB, JB, JD, MAE, LF, SBK, SM, ABN, MP, and GAK report no conflicts of interest to this work. VRA was the MedStar Health Research Institute’s principal clinical trial investigator for studies involving Elcelyx (delayed release metformin).

References

(IOM) (2010) Evaluation of biomarkers and surrogate endpoints in chronic disease, 2101 CONSTITUTION AVE, WASHINGTON, DC 20418 USA


Cesari M et al (2003) Inflammatory markers and onset of cardio-


Cesari M et al (2003) Inflammatory markers and onset of cardio-

Cesari M et al (2003) Inflammatory markers and onset of cardio-

Cesari M et al (2003) Inflammatory markers and onset of cardio-

Chen BH et al (2016) DNA methylation-based measures of bi-


Foster MC et al (2013) Novel filtration markers as predictors of all-


