

Translational Section: Drug Interventions

Editorial

Testing the Geroscience Hypothesis: Early Days

Stephen B. Kritchevsky, PhD^{1,*} and Jamie N. Justice, PhD^{2,⊕}

¹Sticht Center for Healthy Aging and Alzheimer's Prevention and ²Department of Internal Medicine, Section on Gerontology and Geriatric Medicine, Wake Forest University School of Medicine, Winston-Salem, North Carolina.

*Address correspondence to: Stephen B. Kritchevsky, PhD, Sticht Center for Healthy Aging and Alzheimer's Prevention, Wake Forest University School of Medicine, 1 Medical Center Boulevard, Winston-Salem, NC 27157. E-mail: skritche@wakehealth.edu

In this issue, Espinoza and colleagues present the design of a 2-year randomized, double-blind, controlled, clinical trial testing whether 2,000 mg/day of metformin retards the advancement of frailty in prediabetic older adults. Demonstrating that any drug can deflect the progression of frailty would be a major milestone in geriatrics, but this study is also important within the larger frame of geroscience which hypothesizes that human health can be improved by directly targeting the biology of aging (1). Support for this hypothesis is found in studies of model organisms, which show that targeting these pathways in a variety of ways—including the administration of metformin—can increase health span and lifespan (2). The study by Espinoza is one of the few rigorously designed trials to test a drug in a context that has direct implications for evaluating the geroscience hypothesis.

The geroscience hypothesis is relatively new, and there is uncertainty regarding how to optimally design studies to test it. A key uncertainty is end-point selection. Espinoza and colleagues selected the Fried Cardiovascular Health Study (CHS) frailty phenotype as the primary end-point (3). This measure has good face validity with geriatricians and its use is supported by a great deal of observational data underscoring its robust prediction of health outcomes in older adults. However, the frailty phenotype domains were originally operationalized using data from the CHS study, an on-going observational study. The approach which has worked so well for observational research has limitations from a clinical trial perspective.

The Fried CHS frailty phenotype may not be sensitive to change. Participants receive a point for having a value below a threshold cut-point for each of its five components: unintentional weight loss of at least 10 pounds in the past year, slowness, weakness, lack of energy, and low physical activity. This approach makes it relatively insensitive to change because only participants whose measurements actually cross a cut-point during follow-up contribute information. Trials using relatively insensitive measures require more participants or longer follow-up times to detect effects.

A person's classification may change simply due to the passage of time. The phenotype operationalizes the construct of shrinkage as unexpected weight loss over the past year. It is possible that someone who unexpectedly lost weight 11 months previously would no longer meet these criteria at a subsequent follow-up visit even in the absence of any underlying change. Prior weights may not be available when a person enrolls in a trial, so it may be hard to determine if this criterion is met at the baseline of a frailty prevention trial.

Interventions could affect component measures without affecting the underlying physiology of frailty. An exercise intervention is likely to show improvement in the physical activity component even if the underlying frailty pathophysiology is unaffected. Conversely, metformin frequently leads to weight loss so unexplained weight loss may end-up being more common in the treated group even if frailty is not exacerbated.

There are potential options for adapting the Fried frailty construct for clinical trial use. Ceiling effects and sensitivity to change can be addressed by rescaling the component measures. Sanders and colleagues published a Vigor scale (0–10) which scores the frailty components on an expanded ordinal scale, allowing for better discrimination among non- and pre-frail participants (4). One could adopt an approach similar to that used by Simonsick and colleagues when they adapted the Established Populations for Epidemiologic Studies of the Elderly short physical performance battery (SPPB) for the well-functioning Health Aging and Body Composition Study population (5). The SPPB is based on the performance of three lower extremity tasks, each of which is scored on an ordinal scale from 0 to 4. This results in an SBBP score with integer values from 0 to 12. The SPPB is reasonably sensitive for those with poor mobility function but is insensitive in persons with good function. To adapt the battery, the investigators scored the SPPB components on a continuous 0–1 scale assigning scores based on the ratio of the measured value to the best possible score. The rescaling led to a continuous summary score without a ceiling, better distributional characteristics,

and better sensitivity to change (6). The time dependency issue with respect to the shrinkage measure could be addressed in various ways. One might use weight loss since age 50 or 25. This would still be a problem for weight loss interventions, however. One could also consider height loss since age 25. Self-reported adult height is reasonably accurate and height loss would certainly be unintentional and unlikely to be affected by any potential intervention. For trials involving a physical activity intervention, one might consider using accelerometry to quantify sedentary time as an alternative to using a questionnaire to estimate kcals of energy expenditure. These alternatives do not have the depth of data supporting the use of the original measure, but such adaptations could make trials relevant to the underlying physiology of frailty more efficient.

An adapted frailty phenotype measure is by no means the only option for geroscience-relevant trials. Potential outcomes fall on a continuum from those that directly test the hypothesis (e.g. emergence of age-related diseases or life-span) to indirect biomarker-based assessments (Figure 1). The National Institute on Aging's (NIA) Intervention Test Program evaluates pharmacological agents in mice for their effects at the far right of this continuum (e.g. median lifespan) (7). This end-point is a translational challenge. For example, the Targeting Aging with Metformin (TAME) study, a randomized trial designed to test metformin's ability to retard the incidence of multiple age-related diseases and death, will require 3,000 people followed for more than 4 years (8). Studies such as TAME will be necessary to definitively test the geroscience hypothesis. But we can anticipate many trials like the one described by Espinoza and colleagues that evaluate interventions relevant to the biology of aging, and which, if promising, might be scaled up. What kinds of measures might be considered for such trials?

Reliable biomarkers reflecting changes to fundamental aging processes would be valuable. Currently, NIA is supporting an initiative to develop and validate such biomarkers for clinical trials, but at this point, the data related to options for human studies is insufficient to support their suitability for use in trials. There is a rapidly growing literature on indirect measurements which may have value in the clinical trials context, and some broad strategies have emerged: (a) surrogate markers of individual biological hallmarks or pillars; (b) multivariable biomarker composites; and (c) deficit accumulation indices.

Epigenetic age estimators are leading examples of this approach. There are age-related patterns of accumulating DNA

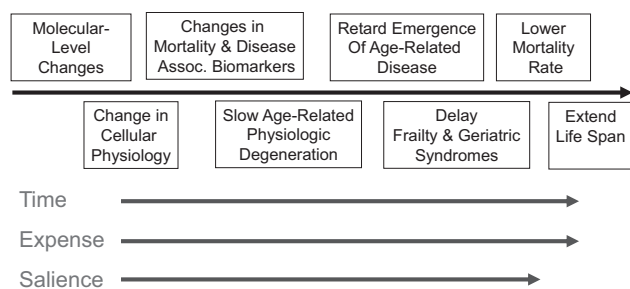


Figure 1. Evaluation continuum for clinical trials in geroscience. Clinical trials are being designed to test the geroscience hypothesis. Potential end-points for such trials exist on a continuum, with biomarker-based measures of the underlying aging biology on one end, and rate of occurrence of clinical disease, geriatric syndromes like frailty, and mortality on the other. With increasing levels of assessment from biomarker-based to hard clinical outcomes, the duration required to observe change in the trial endpoint (time), number of research subject and costs to run the trial (expense), and salience of the endpoint to geroscience hypothesis also increases.

methylation. It is possible to calculate a “biological age” based on these patterns. Levine and colleagues extended this idea to calculate a score based on an apparent phenotypic age which was calculated using levels of various age-related blood chemistries (9). Much work is needed to understand the reliability and sensitivity to change of such measures. Recent preliminary reports suggest that epigenetic age estimators are sensitive to short-term administration of a thymotropic drug combination (recombinant human growth hormone, dehydroepiandrosterone, and metformin) and vitamin D (10,11).

Biomarker composites are scores based on age-related biomarkers which can be used to calculate a summary score which may or may not be referenced to an expected age for an individual with that biomarker profile (12). The behavior of these composites in intervention settings may indicate the pace of the aging process. For example, the Comprehensive Assessment of Long-term Effects of Reducing Intake of Energy trial tested the effects of 2 years of 25% caloric restriction in healthy, nonobese adults (13). In a post hoc analysis, Belsky and colleagues used a collection of clinical chemistries to calculate a “biological age” score and showed that caloric restriction slowed apparent age-advancement (14).

Cumulative deficit indices calculate the proportion of the sum of the number of abnormal clinical signs, symptoms, diseases, or biomarker measures an individual has (numerator) relative to the total number of items included in the index (denominator) to create a score ranging from 0 to 1 (also called Frailty Indices) (15). The trial described by Espinoza and colleagues includes a frailty index as a secondary end-point. Such indices strongly predict mortality and disease outcomes independent of age, and recent data published from the Look AHEAD (Action for Health in Diabetes) trial show that a frailty index is responsive to an intensive lifestyle intervention (16). When the index includes a large number of items the influence of any single component is dampened. This lowers the susceptibility of this approach to items which may be responsive to the intervention for reasons that are not related to the aging process.

Being the early days of the geroscience hypothesis, no one is entirely certain what the best end-point might be. The choice might depend on the biological target (e.g. senolytics vs mechanistic target of rapamycin [mTOR] inhibitors) among many other factors. Early guesses may be off-base and as knowledge emerges, investigators may kick themselves for getting it wrong. Until we have a solid empiric foundation laying out which end-points are valid, clinically meaningful, reproducible, and sensitive to change, readers, peer-reviewers, and editors should be open-minded and pay attention to not only the prespecified primary end-point but also to secondary and exploratory end-points. A secondary or exploratory end-point may turn out to be the one that will ultimately be found to be the most useful.

To accelerate progress in evaluating the geroscience hypothesis, it will be important that those conducting such studies bank appropriately collected biological samples, obtain potentially useful measures in a standardized way, and be prepared to share data and results in common formats. Sharing how a variety of measures behave across intervention types and settings will help to minimize missteps and inefficiencies. In recognition of the value of this goal, the NIA recently funded the Translational Geroscience Network (R33 AG061456) which has a goal to identify and promulgate standard protocols and data collection instruments to facilitate data sharing across geroscience efforts. Its first work products are expected to appear later in 2020.

Funding

This work was partially supported by National Institute on Aging Grants: P30 AG21332 and K01..

References

- Kennedy BK, Berger SL, Brunet A, et al.. Geroscience: linking aging to chronic disease. *Cell*. 2014;159:709–713. doi: 10.1016/j.cell.2014.10.039
- Barzilai N, Crandall JP, Kritchevsky SB, Espeland MA. Metformin as a tool to target aging. *Cell Metab*. 2016;23:1060–1065. doi: 10.1016/j.cmet.2016.05.011
- Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. *J Gerontol Ser A*. 2001 56: M146–M157. doi: 10.1093/gerona/56.3.M146
- Sanders JL, Singh J, Minster RL, et al.; Long Life Family Study Research Group. Association between mortality and heritability of the scale of aging vigor in epidemiology. *J Am Geriatr Soc*. 2016;64:1679–1683. doi: 10.1111/jgs.14190
- Simonsick EM, Newman AB, Nevitt MC, et al.; Health ABC Study Group. Measuring higher level physical function in well-functioning older adults: expanding familiar approaches in the Health ABC study. *J Gerontol A Biol Sci Med Sci*. 2001;56:M644–M649. doi: 10.1093/gerona/56.10.m644
- Houston DK, Leng X, Bray GA, et al.; Action for Health In Diabetes (Look AHEAD) Movement and Memory Ancillary Study Research Group. A long-term intensive lifestyle intervention and physical function: the look AHEAD Movement and Memory Study. *Obesity (Silver Spring)*. 2015;23:77–84. doi:10.1002/oby.20944
- Warner HR, Ingram D, Miller RA, Nadon NL, Richardson AG. Program for testing biological interventions to promote healthy aging. *Mech Ageing Dev*. 2000;115:199–207. doi: 10.1016/s0047-6374(00)00118-4
- Justice JN, Niedernhofer L, Robbins PD, et al. Development of clinical trials to extend healthy lifespan. *Cardiovasc Endocrinol Metab*. 2018;7:80–83. doi: 10.1097/XCE.0000000000000159
- Levine ME, Lu AT, Quach A, et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging-Us*. 2018;10:573–91. doi: 10.18632/aging.101414
- Fahy GM, Brooke RT, Watson JP, et al. Reversal of epigenetic aging and immunosenescent trends in humans. *Aging Cell*. 2019;18:e13028. doi: 10.1111/accel.13028
- Chen L, Dong Y, Bhagatwala J, Raed A, Huang Y, Zhu H. Effects of vitamin D3 supplementation on epigenetic aging in overweight and obese African Americans with suboptimal vitamin D status: a randomized clinical trial. *J Gerontol A Biol Sci Med Sci*. 2019;74:91–98. doi: 10.1093/gerona/gly223
- Justice JN, Ferrucci L, Newman AB, et al. A framework for selection of blood-based biomarkers for geroscience-guided clinical trials: report from the TAME Biomarkers Workgroup. *Geroscience*. 2018;40:419–436. doi: 10.1007/s11357-018-0042-y
- Ravussin E, Redman LM, Rochon J, et al.; CALERIE Study Group. A 2-year randomized controlled trial of human caloric restriction: feasibility and effects on predictors of health span and longevity. *J Gerontol A Biol Sci Med Sci*. 2015;70:1097–104. doi: 10.1093/gerona/glv057
- Belsky DW, Huffman KM, Pieper CF, Shalev I, Kraus WE. Change in the rate of biological aging in response to caloric restriction: CALERIE biobank analysis. *J Gerontol A Biol Sci Med Sci*. 2017;73:4–10. doi: 10.1093/gerona/glx096
- Rockwood K, Mitnitski A. Frailty in relation to the accumulation of deficits. *J Gerontol A Biol Sci Med Sci*. 2007;62:722–727. doi: 10.1093/gerona/62.7.722
- Simpson FR, Pajewski NM, Nicklas B, et al.; Indices for Accelerated Aging in Obesity and Diabetes Ancillary Study of the Action for Health in Diabetes (Look AHEAD) Trial. Impact of multidomain lifestyle intervention on frailty through the lens of deficit accumulation in adults with type 2 diabetes mellitus. *J Gerontol A Biol Sci Med Sci*. 2019;pii:glz197. doi: 10.1093/gerona/glz197 [epub ahead of print].