

# Building a clinical platform for response-guided geroscience: safety and feasibility in the randomized ELITE pilot trial

**Bruno Balen**

Ani Biome PBC

**Evelyne Bischof**

bischofevelyne@gmail.com

Sheba Medical Center <https://orcid.org/0000-0002-6034-1742>

**Abigail Goshen**

Sheba Medical Center

**Tzipora Strauss**

Sheba Medical Center

**Charles Chaim Haber**

Sheba Medical Center

**Michael P. Snyder**

Stanford University School of Medicine <https://orcid.org/0000-0003-0784-7987>

**Vadim N. Gladyshev**

Brigham and Women's Hospital; Harvard Medical School <https://orcid.org/0000-0002-0372-7016>

**Steve Horvath**

University of California, Los Angeles <https://orcid.org/0000-0002-4110-3589>

**Brian K. Kennedy**

National University of Singapore <https://orcid.org/0000-0002-5754-1874>

**Nika Pintar**

nika@anibiome.ai

Ani Biome PBC <https://orcid.org/0009-0001-1431-9368>

---

## Research Article

**Keywords:** geroscience, secondary outcomes, safety, feasibility, treatment adherence, multimodal phenotyping, clinical-trial platform

**Posted Date:** October 5th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11209869/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Additional Declarations:** The authors declare potential competing interests as follows: B.B. and N.P. are affiliated with ANI BIOME PBC, which developed the platform, response-state architecture and investigational AgeBiotics products described in this work. B.B. and N.P. are named inventors on patents and patent applications held by or assigned to ANI BIOME PBC relating to multispectral measurement technology, the digital clinical-trial platform and associated computational methods, and compositions related to the AB investigational products. M.P.S. and E.B. serve as scientific advisers to ANI BIOME PBC. V.N.G. served on advisory panels for Retro Biosciences, Dior, LongeVC and Thorne. The Regents of the University of California are the sole owner of patents and patent applications directed at epigenetic biomarkers, including GrimAge and PhenoAge, for which S.H. is a named inventor. S.H. is a founder of the non-profit Epigenetic Clock Development Foundation, which licenses these patents. S.H. was a Principal Investigator at Altos Labs until March 2026 and holds equity in Altos Labs, Inc. The remaining authors declare no competing interests.

---

# Abstract

**Background:** Deeply personalized geroscience requires more than selecting the right intervention. It requires a clinical trial system that connects treatment exposure, longitudinal measurements and clinical oversight to characterize individual responses over time. Establishing whether midlife and older adults can sustain the combined demands of multicomponent treatment and repeated data collection is a foundational step toward this approach. ELITE was an ethics-approved, eight-week, randomized, open-label geroscience trial in 31 adults aged 50–80 without diabetes. It combined repurposed metabolic drugs with two investigational AgeBiotics supplements, AB-01 oral solution and AB-02 capsules, alongside daily app-based data collection and scheduled clinical and biospecimen assessments. This report examines whether participants could sustain this regimen with an acceptable safety profile, presenting the trial's prespecified safety and feasibility secondary outcomes: adverse events, study completion and overall protocol adherence. Feasibility benchmarks were adherence  $\geq 60\%$  and dropout  $< 20\%$ ; overall adherence equally weighted four treatment components and app scanning. Whether such an architecture can be delivered and audited under real clinical conditions remains an important translational question. Intervention-outcome analyses, including the primary MIC-C endpoint and comparisons between randomized arms, are outside the scope of this report and will be presented separately.

**Results:** Thirty of 31 participants completed follow-up (96.8%; 95% confidence interval, 83.8-99.4), with dropout of 3.2%. All 30 completers (100%) met the  $\geq 60\%$  adherence threshold; their mean overall adherence was 91.5% (standard deviation, 7.4). Corresponding means were 89.4% (13.8) across all 31 randomized participants and 93.3% (5.7) among 25 per-protocol participants. Digital capture occurred on 1,816 of 1,978 observation-window participant-days (91.8%); median digital adherence among completers was 97.7% (interquartile range, 89.1-99.6). Nineteen mild adverse events occurred in 17 participants (54.8%); no serious adverse events were reported. Adverse events led to treatment-component withdrawal or interruption in four participants, who continued study follow-up.

**Conclusions:** ELITE met its prespecified feasibility benchmarks and demonstrates a practical institutional pathway for integrating repurposed metabolic therapy, sustained longitudinal observation and clinical oversight in response-guided geroscience. More broadly, the study establishes the operational feasibility of an integrated geroscience architecture combining protocolized multidomain intervention delivery, dense longitudinal measurement, provenance-preserving computation, probabilistic response assessment, physician-supervised governance and monitored clinical close-out.

**Trial registration:** ISRCTN84791030. Registered 28 September 2026, retrospectively registered.

## Introduction

Deeply personalized geroscience requires more than selecting the right intervention: it requires a trial system that can observe individual health trajectories, sustain participant engagement, maintain treatment accountability, ensure safety oversight and support auditable computation under real clinical

conditions. ELITE was conducted during the XPRIZE Healthspan Semi-Finals as an eight-week pilot of longitudinal phenotyping and response monitoring. The competition's focus on muscular, immune and cognitive improvement in individual participants made responder identification central to the trial's design. Multicomponent interventions may affect metabolic, functional, immune, cognitive and behavioral domains differently across individuals, yet conventional trials are optimized to estimate an average effect from sparse visits. Identifying credible within-person response trajectories is therefore central to moving geroscience interventions from biological rationale to clinically interpretable use [1–3]. The competition applies personalized response thresholds across muscular, immune and cognitive domains and evaluates the proportion of participants meeting them, making individual response classification, rather than average effect estimation alone, central to the framework.

Repeated within-person observations can reveal the timing, direction and persistence of change that isolated measurements obscure [4, 5]. In ELITE, daily digital collection was intended to capture this temporal detail between scheduled clinical and molecular assessments, without requiring daily clinical testing or biospecimen collection. The periodic assessments supplied biological and functional anchors for interpreting the digital trajectories. Bayesian updating provides a formal way to accumulate evidence and uncertainty for an individual over time [6]. In ELITE, responder assessment was designed as longitudinal decision support, not as a repeated-crossover N-of-1 trial [7].

The central feasibility question was whether adults aged 50–80 years could integrate and sustain the combined treatment and measurement requirements in daily life over the eight-week intervention (Fig. 1). The regimen involved twice-daily AB-02 capsules, three daily administrations of AB-01 oral solution, two pharmacological agents and daily app-based data collection. Completion and adherence tested a practical foundation for personalized medicine: maintaining treatment exposure and longitudinal observation so that evolving individual responses can inform care. This provides the operational basis for the planned longer trial, in which physicians will adapt treatment over time according to each participant's response trajectory.

ELITE was designed to bring these requirements into one clinical pathway: defined intervention exposure, daily participant-generated observations, periodic clinical and molecular anchoring, traceable data processing, physician review and monitored close-out (Fig. 2). The purpose was to preserve the connection between what a participant received, what changed over time and how that change could be interpreted clinically.

ELITE also illustrates how geroscience protocols can be implemented under existing institutional ethics and governance pathways. Ethics approval was obtained to administer repurposed metabolic drugs (metformin, semaglutide and empagliflozin) to adults without diabetes for an explicit healthy-aging objective, within institutional clinical-safety and investigational-product oversight.

The clinical question motivating ELITE is whether multidomain intervention can improve muscular, immune and cognitive function in metabolically at-risk adults without diabetes, and whether individual responders can be identified prospectively. This report addresses the prespecified safety and feasibility

secondary outcomes and the operational evidence supporting their interpretation. For responder identification to be translationally useful, response signals must be attributable, repeatable and clinically anchored; the present report therefore focuses on the feasibility, traceability and governance required for subsequent validation of individual response classification.

## Methods

### Translational architecture and intended use

ELITE links intervention exposure, daily longitudinal capture, periodic clinical and molecular anchors, a prespecified Composite Response Score (CRS) and Bayesian decision layer, and physician-supervised trial governance. Each response state can be traced to its source measurements and clinical context.

The protocol-defined intervention was algorithmic polytherapy, integrating the multicomponent treatment regimen with longitudinal digital response monitoring and a physician-supervised decision-support system. In this eight-week pilot, randomized treatment assignments remained fixed: response states were generated and reviewed, but they did not trigger algorithm-guided treatment reassignment.

This pathway forms the decision layer of a broader state-reconstruction architecture. Its observation model is modality-agnostic, treating heterogeneous measurements as time-indexed evidence streams. In ELITE, digital images, multispectral measurements, voice, structured micro-surveys, and clinical, functional and molecular anchors were aligned by participant, source and time. These observations provide complementary constraints for reconstructing a participant-specific latent state from sparse, irregular and partially observed data, with explicit uncertainty and preserved source provenance [8]. The CRS was the prespecified scalar response output evaluated operationally in this report. Reproducible implementation therefore requires prespecification of the observation window, intervention modules, required inputs and timestamps, source rules, CRS construction, missing-data handling, response thresholds, clinical anchors, safety escalation, clinician authority and audit outputs.

### Design, participants and consent

ELITE was a randomized, open-label, two-active-arm phase IIa pilot conducted at Sheba Medical Center, Ramat Gan, Israel, with 1:1 allocation. The Sheba Helsinki Committee / institutional review board approved the study (2361-25-SMC) on 2 September 2025, before enrollment, and all participants provided informed consent. The randomization method, endpoint hierarchy and feasibility benchmarks were specified in the approved protocol before first enrollment on 17 November 2025. Public registration was initially exempted under the regulatory pathway approved by the Helsinki Committee. The study was subsequently registered in the Israeli Ministry of Health MyTrial registry on 21 July 2026 (MOH\_2026-07-21\_016947), retrospectively. The study was also registered with ISRCTN on 28 September 2026 (ISRCTN84791030). Supplementary Table 1 summarizes the protocol outcome hierarchy and identifies the outcomes reported here.

Participants were recruited primarily from the Sheba Healthspan Research Population (SHARP) study, an ongoing randomized study of a personalized healthy-aging intervention in community-dwelling adults older than 50 years at the Sheba Longevity Center, with planned enrollment of approximately 1,500 participants [9]. SHARP participants undergo extensive baseline and 12-month aging, functional and cognitive assessment. All participants recruited from SHARP had completed SHARP before starting ELITE; no minimum interval between SHARP completion and ELITE initiation was required. Further participants were recruited through community and clinical-network outreach. Eligible participants were community-dwelling adults aged 50–80 years who could provide informed consent, use a compatible device, attend required visits and complete home procedures. The study targeted metabolically at-risk adults without diabetes, suited to testing whether low-dose metabolic interventions produce detectable individual responses: key criteria included body-mass index at least 25 kg/m<sup>2</sup> and screening glycated hemoglobin (HbA1c) below 6.5%. Key exclusions included diabetes, body-mass index below 25 kg/m<sup>2</sup>, estimated glomerular filtration rate below 45 mL/min/1.73 m<sup>2</sup>, severe heart failure, recent major cardiovascular events, recent active malignancy, chronic systemic corticosteroid or glucose-lowering medication use, and investigator-determined conditions affecting safety or feasibility.

## Randomization and interventions

Participants were randomized 1:1 to Arm A (n = 16) or Arm B (n = 15) for an approximately eight-week active intervention. The protocol specified automated allocation through a centralized, secure, web-based system, using a permuted block design stratified by sex and by age decade (50–59, 60–69 and 70–80 years), with the central system concealing each allocation until assignment; block size was not specified. Randomization occurred at Day 0 once all baseline procedures were complete, after which the site pharmacy dispensed the assigned products. Participants, investigators, site staff and outcome assessors were aware of allocation, and no placebo was used. Blinding was not feasible because the arms differed in administration route and schedule. Arm A received metformin 425 mg orally once daily plus semaglutide 0.25 mg by subcutaneous injection once weekly. Arm B received metformin 425 mg orally once daily plus empagliflozin (Jardiance) 10 mg orally once daily. These modules paired biguanide activity with GLP-1 receptor or SGLT2 pathways, respectively, as mechanistically distinct low-dose interventions. Metformin has been advanced as a geroscience repurposing candidate, while GLP-1 receptor agonists and SGLT2 inhibitors have substantial human outcome evidence across obesity, cardiovascular, renal and metabolic disease populations [10–12]. In this feasibility setting, the randomized regimens served as clinically governable, mechanistically distinct interventions that imposed realistic attribution, tolerability and safety demands on the response-monitoring architecture.

In addition to the two assigned metabolic drugs, participants in both randomized arms received the same prescribed regimen of two investigational AgeBiotics supplements: AB-01, an oral solution given as three 19-mL sachets daily, and AB-02, a capsule formulation given as five capsules twice daily. Each AB-01 sachet was diluted in 300 mL of water and taken in the morning, before lunch and before dinner; AB-02 was taken after meals. The pro-metabolic rationale was to perturb the host-microbiome-immune-metabolic interface with AB-01 and support intracellular cofactor availability, redox handling and tissue

resilience with AB-02 [13–15]. The AgeBiotics formulations and metformin were administered in both randomized arms according to their respective dosing schedules. The randomized difference between arms was the arm-specific pharmacologic agent: semaglutide in Arm A and empagliflozin in Arm B. Prescribed treatment changes and discontinuations were subject to physician supervision, with pharmacy accountability and protocol-defined safety review. This was the protocol-level pro-metabolic rationale, referring to support for substrate handling and oxidation, energy transfer, inter-system coordination, recovery and preserved function rather than targeting an isolated biomarker.

Participants received app onboarding and training in daily digital scanning and adherence logging. The app guided data collection, provided medication reminders and allowed participants to confirm product-specific intake or select a skip status, storing dose quantities, reminder times and update timestamps. Treatment-adherence analyses used the site medication-compliance and accountability records described below.

### Measurement schedule and data provenance

Daily guided app sessions were designed to capture tongue, left-eye, right-eye, face and wrist images, a structured micro-survey and a voice recording. Participants also reported contextual information about events immediately preceding the scan and during the previous day. A dedicated device acquired multispectral arrays for the tongue, face and wrist. Sessions were timestamped and linked to source records. Periodic clinical visits provided functional assessments and molecular sampling. A digital scan-day required at least one unique source-linked scan document on a participant-date. Expected days were counted inclusively from baseline to the final assessment (TP3) or documented study withdrawal. Optional post-TP3 scans were analyzed as exploratory continuation and excluded from required-window denominators. Capture availability, processing success and required spectral completeness were evaluated as distinct operational metrics rather than collapsed into a single measure of digital feasibility.

Scheduled visits paired daily digital observations with body-composition, cardiorespiratory, pulmonary, vascular/autonomic, muscular, mobility/frailty, cognitive, psychological, endocrine and clinical-laboratory assessments. Affect, cognition, sleep and symptom measures were included to capture regulatory and experiential dimensions of response alongside objective physiological and molecular measures [15]. Biological sampling included blood-derived specimens at baseline (TP1) and the final assessment (TP3), and stool, oral (tongue) swabs and vaginal swabs at TP1, the intermediate sampling time point (TP2) and TP3; vaginal sampling applied to female participants. Participant identifiers and collection time points linked these specimens to the longitudinal digital records. Supplementary Table 2 summarizes the collection schedule and specimen types. The completed molecular program included Nightingale nuclear magnetic resonance metabolomics (105 metabolites), Olink Reveal proteomics (1,034 proteins), NULISaseq CNS Disease Panel 120 (120 proteins), Genos IgG N-glycan capillary-gel-electrophoresis profiling (27 chromatographic peaks and six derived traits, with GlycanAge also calculated), metagenomics, whole-genome bisulfite sequencing (WGBS), and RNA sequencing (RNA-seq). Assay-

specific sample and quality-control denominators were retained by assay and time point so that each completed molecular layer could be linked to the participant-level clinical and digital trajectories.

### Prespecified outcome hierarchy and reporting scope

The protocol prespecified mean change from baseline to Week 8 in the Muscular-Immune-Cognitive Composite (MIC-C) score as the primary endpoint. This report presents the prespecified safety and feasibility secondary outcomes. Safety outcomes comprise the incidence, severity and relatedness of adverse events (AEs) and serious adverse events (SAEs). Feasibility outcomes comprise study completion rate and overall protocol adherence rate. These outcomes were assessed from baseline through Week 8. The primary MIC-C outcome and other intervention-outcome and validation analyses will be reported separately. Supplementary Table 1 summarizes the outcome hierarchy and reporting status.

The protocol specified participant adherence of at least 60% and dropout below 20% as feasibility benchmarks. Recruitment, data completeness and logistical viability were additional descriptive feasibility measures without numerical thresholds.

Study completion was defined as documented follow-up completion through the end-of-intervention visit (TP3), using all 31 randomized participants as the denominator. Per-protocol inclusion followed the canonical cohort registry, which excluded one study discontinuation, four participants with intervention-component discontinuations and one participant with a late final visit.

Treatment adherence covered metformin, the assigned arm-specific GLP-1 receptor agonist or SGLT2 inhibitor, and AgeBiotics AB-01 and AB-02. The site-provided consolidated medication-compliance extract was reconciled with dispensing/accountability records and participant source records. The overall protocol adherence endpoint and  $\geq 60\%$  benchmark were prespecified. For this analysis, overall protocol adherence was operationalized as the arithmetic mean of each participant's available adherence percentages for the four assigned treatment components and app scanning, with equal component weights and no imputation. Each of the 30 completers contributed all five components (20% each), and cohort summaries weighted participants equally. Analyses included all randomized participants, the 30 completers and the 25-participant per-protocol population. Treatment-only summaries and the stopped-component sensitivity analysis are reported in Supplementary Table 5. App-scanning adherence was the percentage of expected days with an observed scan-day in the observation window defined above.

### Composite Response Score and responder framework

The protocol defined the CRS, termed the ADAPT Score in the protocol document, as a governed, high-frequency projection of within-person change derived from daily digital biomarker data. The platform combined tongue, face, left-eye, right-eye and wrist images, multispectral data, voice recordings and structured micro-survey responses to update the participant's trajectory in near real time. The CRS

estimated change from the individual's own baseline and supported posterior probabilities of benefit and deterioration. The prespecified decision states were mutually exclusive: KEEP when  $\Pr(\Delta\text{CRS} \geq 0)$  was at least 0.80, AUGMENT when the probability was 0.30 to below 0.80, SWITCH when it was 0.10 to below 0.30, and STOPP when it was below 0.10 or a safety trigger occurred (Supplementary Table 3). The prespecified thresholds represented operational decision criteria: 0.80 was set as posterior confidence in benefit sufficient to continue unchanged, 0.30 as the boundary below which alternative strategies should be considered, and 0.10 as evidence of futility or harm warranting discontinuation.

Responder validation was designed across two timescales. The CRS supplied the near-real-time trajectory, while the Week 8 Muscular-Immune-Cognitive Composite (MIC-C) supplied a multidomain clinical anchor derived from clinical, functional and multi-omic assessments across muscular, immune and cognitive domains collected at baseline and Week 8. These domains (muscular, immune and cognitive function) were selected as competition-defined proxies for healthspan within the XPRIZE Healthspan framework [16]. The competition requires improvement in all three domains and uses personalized thresholds indexed to expected 10-, 15- or 20-year declines in referent populations; the proportion of participants meeting these thresholds determines prize eligibility. Individual response classification is therefore the central analytical challenge. The protocol prespecified evaluation of the CRS-driven engine against Week 8 MIC-C responders, including the area under the receiver-operating-characteristic curve. This separates rapid response detection from clinical confirmation and makes the responder definition testable. The protocol specified a Bayesian multi-arm, multi-stage framework with potential adaptive reassignment. In this eight-week pilot, however, randomized arm allocation remained fixed throughout. CRS response states were computed and displayed to physicians for review and clinical oversight, but they did not trigger reassignment between arms.

### Clinical governance and auditability

For any response classification to be defensible, it must be traceable to verified source data and reviewed clinical decisions. Investigational-product accountability, adverse-event reporting, physician review, external monitoring and ethics oversight were therefore embedded in the method. The monitoring plan covered source-data verification, protocol-deviation handling, safety reporting, investigational-product management and database lock. The protocol specified that the trial would pause for data-safety review if any drug-related Grade 3 or any Grade 4 adverse event occurred, and defined individual discontinuation criteria including persistent non-compliance with digital data entry and a statistical harm rule of  $\Pr(\Delta\text{CRS} \leq 0) > 0.90$  at two consecutive evaluations. Safety oversight and study monitoring were provided under the approved protocol by Sheba's Sponsor's Office, with responsibility for source-data verification, protocol-deviation review, pharmacovigilance and data management. The ethics committee retained jurisdiction over amendments, deviations and safety events, and institutional pharmacovigilance procedures governed adverse-event capture, coding and escalation. No pause criterion was met. This governance layer linked source-backed longitudinal observations to conventional source, pharmacy and electronic case-report-form records summarized in Supplementary Table 4.

Adverse events were collected throughout the study through participant reporting and scheduled study contacts, with severity assessed using CTCAE version 5.0 and relatedness assessed by investigators.

## Statistical analysis

The operational recruitment target of 30 participants was selected on feasibility grounds. Safety and feasibility outcomes were summarized using counts, proportions, and descriptive statistics. Participant-level proportions are reported with Wilson 95% confidence intervals. Adherence summaries use means, sample standard deviations, medians and linearly interpolated interquartile ranges. Overall adherence and  $\geq 60\%$  threshold attainment are presented for completers, all randomized participants and the per-protocol population. Pooled scan-day and product-record summaries are descriptive because repeated observations are clustered within participants. Reporting drew on CONSORT 2025 and its explanation and elaboration, relevant principles of the pilot/feasibility extension, CONSORT-AI and SPIRIT-AI for algorithmic decision support, TIDieR for intervention reproducibility, and CONSORT Harms 2022 [17–23].

## Results

### Participant flow and intervention exposure

Of 282 individuals assessed for eligibility through the tracked SHARP registry and its pilot cohort [9], 22 enrolled: 20 of 261 screened from the registry database of 1,200 participants with at least one year of follow-up, and 2 of 21 screened from the pilot cohort. Of the 260 not enrolled, 104 did not meet the ELITE inclusion and exclusion criteria and 156 declined or could not be reached after direct telephone outreach; eligibility was not fully determined for this second group. Nine additional participants were enrolled through community and clinical-network outreach, for which a screening denominator was not tracked. Thirty-one participants were randomized between 17 November and 17 December 2025 (Arm A,  $n = 16$ ; Arm B,  $n = 15$ ); recruitment was closed once the enrollment target was reached, and the trial completed as planned. Follow-up continued to the last participant visit on 25 February 2026. Pharmacy dispensing and accountability records document treatment exposure in all 31 participants; close-out records document follow-up completion in 30 and no loss to follow-up. One Arm B participant discontinued because she could not independently complete home scanning despite training; the discontinuation was not adverse-event related. The safety population comprised all 31 randomized participants (Fig. 3; Table 1).

Table 1  
Baseline characteristics by randomized arm and overall

| Characteristic                                                                                                                                                                                   | Arm A (n = 16)       | Arm B (n = 15)       | Overall (n = 31)     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------|----------------------|----------------------|
| Age, years                                                                                                                                                                                       | 59.8 (6.7); n = 16   | 60.1 (5.8); n = 15   | 59.9 (6.2); n = 31   |
| Female sex                                                                                                                                                                                       | 5/16 (31.2%)         | 4/15 (26.7%)         | 9/31 (29.0%)         |
| Body-mass index, kg/m <sup>2</sup>                                                                                                                                                               | 29.1 (4.5); n = 16   | 28.7 (3.2); n = 15   | 28.9 (3.9); n = 31   |
| Systolic blood pressure, mm Hg                                                                                                                                                                   | 134.1 (15.7); n = 16 | 124.6 (13.0); n = 15 | 129.5 (15.0); n = 31 |
| Body fat, %                                                                                                                                                                                      | 31.9 (8.9); n = 16   | 32.3 (9.1); n = 15   | 32.1 (8.9); n = 31   |
| Fat-free mass index, kg/m <sup>2</sup>                                                                                                                                                           | 19.7 (2.8); n = 16   | 19.0 (1.7); n = 15   | 19.4 (2.3); n = 31   |
| Sarcopenia                                                                                                                                                                                       | 0/16 (0.0%)          | 0/15 (0.0%)          | 0/31 (0.0%)          |
| Frailty: not frail                                                                                                                                                                               | 13/16 (81.2%)        | 11/15 (73.3%)        | 24/31 (77.4%)        |
| Frailty: pre-frail                                                                                                                                                                               | 3/16 (18.8%)         | 3/15 (20.0%)         | 6/31 (19.4%)         |
| Frailty: frail                                                                                                                                                                                   | 0/16 (0.0%)          | 1/15 (6.7%)          | 1/31 (3.2%)          |
| 10-m gait speed, m/s                                                                                                                                                                             | 1.4 (0.1); n = 16    | 1.4 (0.2); n = 15    | 1.4 (0.2); n = 31    |
| Timed up-and-go, s                                                                                                                                                                               | 6.0 (0.8); n = 16    | 6.1 (1.2); n = 15    | 6.0 (1.0); n = 31    |
| Montreal Cognitive Assessment score                                                                                                                                                              | 27.6 (2.0); n = 16   | 27.6 (2.1); n = 15   | 27.6 (2.0); n = 31   |
| Any chronic medication reported                                                                                                                                                                  | 7/13 (53.8%)         | 10/14 (71.4%)        | 17/27 (63.0%)        |
| Values are mean (standard deviation); n unless otherwise indicated. Percentages use the non-missing denominator shown in each cell. Chronic medication status was available for 27 participants. |                      |                      |                      |

Prespecified secondary outcomes: safety and feasibility

## Study completion

Thirty of 31 randomized participants completed follow-up through TP3 (96.8%; Wilson 95% confidence interval, 83.8–99.4). Completion was 16/16 (100%) in Arm A and 14/15 (93.3%) in Arm B. Dropout was 3.2% (1/31), below the prespecified 20% benchmark. Twenty-five participants met per-protocol criteria (80.6%; 63.7–90.8), as defined in Methods (Table 2).

Table 2  
Feasibility outcomes and supporting operational measures

| Outcome or measure                               | Denominator                                         | Result                                                                                          | Interpretation                                                                                     |
|--------------------------------------------------|-----------------------------------------------------|-------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------|
| Enrolled/randomized                              | Operational recruitment target 30                   | 31 (103.3%)                                                                                     | Recruitment target met; recruitment closed once reached                                            |
| Follow-up completion through TP3                 | 31 randomized                                       | 30/31 (96.8%); 95% confidence interval 83.8–99.4. Arm A: 16/16 (100%); Arm B: 14/15 (93.3%)     | Completion measure; dropout 1/31 (3.2%); no loss to follow-up                                      |
| Overall protocol adherence rate                  | 30 completers; 31 randomized; PP: 25                | Completers: mean 91.5% (SD 7.4); $\geq 60\%$ : 30/30 (100%). All: 89.4% (13.8); PP: 93.3% (5.7) | Equal weights for the four assigned products and app scanning; Supplementary Table 6               |
| Per-protocol population                          | 31 randomized                                       | 25 (80.6%); 95% confidence interval 63.7–90.8                                                   | Excludes one study discontinuation, four component discontinuations and one late final visit       |
| Any treatment exposure                           | 31 randomized                                       | 31 (100%)                                                                                       | Pharmacy dispensing and accountability records document allocated treatment in all 31 participants |
| Required scan-days, observation window           | 1,978 expected                                      | 1,816 observed (91.8%)                                                                          | At least one source-linked scan document per expected participant-day                              |
| Participant scan adherence, TP3 completers       | 30 participants                                     | Median 97.7% (interquartile range 89.1–99.6)                                                    | Participant-level digital component                                                                |
| Participant scan adherence, all observed windows | 31 participants                                     | Median 97.1% (interquartile range 87.5–99.3)                                                    | Participant-level digital component; all randomized participants                                   |
| Participant scan adherence, per protocol         | 25 participants                                     | Median 98.3% (interquartile range 90.8–100.0)                                                   | Participant-level summary                                                                          |
| Four-component treatment adherence               | 123 product records / 31 participants; PP: 100 / 25 | Mean 89.2% overall; 93.5% per protocol; median 100% in both                                     | Equal weight per available participant-product record; Supplementary Table 5                       |
| Pipeline-status success                          | 2,251 documents                                     | 2,214 (98.4%)                                                                                   | Recorded processing completion                                                                     |
| Complete spectral arrays                         | 2,251 documents                                     | 2,121-2,162 (94.2–96.0%)                                                                        | Tongue, face and wrist lanes                                                                       |

| Outcome or measure              | Denominator       | Result                                | Interpretation                       |
|---------------------------------|-------------------|---------------------------------------|--------------------------------------|
| Optional post-TP3 participation | 30 TP3 completers | 30 with $\geq 1$ scan; median 14 days | Exploratory post-study participation |

## Protocol adherence

All 30 TP3 completers (100%; Wilson 95% confidence interval, 88.6–100.0) met the prespecified overall adherence threshold of at least 60%. Their mean overall adherence across the four treatment components and app scanning was 91.5% (standard deviation, 7.4), with a median of 92.4% (interquartile range, 86.2–97.2). Among completers, mean overall adherence was 90.2% (8.4;  $n = 16$ ) in Arm A and 92.9% (6.0;  $n = 14$ ) in Arm B. Mean overall adherence was 93.3% (5.7) among the 25 per-protocol participants and 89.4% (13.8) across all 31 randomized participants. All 25 per-protocol participants met the threshold. Table 2 and Supplementary Tables 5 and 6 present overall and component-specific results.

Across the four treatment components, 123 participant-product records from 31 participants yielded mean recorded adherence of 89.2% (standard deviation, 20.9) and a median of 100.0% (interquartile range, 89.6–100.0). Among 25 per-protocol participants contributing 100 records, the mean was 93.5% (12.1) and the median 100.0% (92.6–100.0). The sensitivity analysis excluding five stopped-component records gave a mean of 90.6% across 118 records. Mean product-specific adherence was 81.5% (20.2;  $n = 30$ ) for AB-01, 95.3% (17.3;  $n = 31$ ) for AB-02, 83.7% (25.4;  $n = 31$ ) for metformin, 99.2% (3.1;  $n = 16$ ) for semaglutide and 92.9% (22.8;  $n = 15$ ) for empagliflozin (Supplementary Table 5).

Digital capture occurred on 1,816 of 1,978 expected scan-days across all 31 participants (91.8%); the per-protocol result was 1,502 of 1,629 (92.2%). Median participant-level digital adherence was 97.7% (interquartile range, 89.1–99.6) among the 30 completers and 98.3% (90.8–100.0) among the 25 per-protocol participants. All 30 completers exceeded 60% digital adherence. The median inclusive observation window among completers was 65 days (range, 58–80), starting at baseline, which could precede randomization. TP3 was scheduled at Week 8; a final visit on treatment day 79 following intercurrent illness was recorded as a protocol deviation (Fig. 4; Table 2).

## Safety

In the safety population of 31 participants, 19 adverse events occurred in 17 participants (54.8%; Wilson 95% confidence interval, 37.8–70.8): 12 events in 10 of 16 Arm A participants (62.5%; 38.6–81.5) and 7 events in 7 of 15 Arm B participants (46.7%; 24.8–69.9). One participant reported three events. No serious adverse events, hospitalizations, deaths or adverse events of special interest were reported. Seventeen events were assessed as related to a study product, one as possibly related and one had relatedness recorded as not applicable. Reported events comprised gastrointestinal symptoms (10), oral-solution palatability complaints (7), fatigue (1) and concern about potential medication harm (1).

The clinical team confirmed that all 19 events were mild, consistent with CTCAE v5.0 Grade 1. A treatment component was withdrawn in three participants and interrupted in one (4/31; 12.9%; 5.1–28.9); all four continued study visits. No protocol-driven dose reductions were recorded (Fig. 4D; Table 3).

Table 3  
Safety outcomes

| Safety outcome                                                                                                                                                                                                                                                                                                                                              | Arm A (n = 16)    | Arm B (n = 15)   | Total (n = 31)    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|------------------|-------------------|
| Participants with at least one adverse event                                                                                                                                                                                                                                                                                                                | 10 (62.5%)        | 7 (46.7%)        | 17 (54.8%)        |
| Adverse events reported                                                                                                                                                                                                                                                                                                                                     | 12                | 7                | 19                |
| Mild events (CTCAE Grade 1)                                                                                                                                                                                                                                                                                                                                 | 12/12 (100%)      | 7/7 (100%)       | 19/19 (100%)      |
| Events above CTCAE Grade 1                                                                                                                                                                                                                                                                                                                                  | 0/12              | 0/7              | 0/19              |
| Serious adverse events                                                                                                                                                                                                                                                                                                                                      | 0                 | 0                | 0                 |
| Hospitalizations / deaths / adverse events of special interest                                                                                                                                                                                                                                                                                              | 0 / 0 / 0         | 0 / 0 / 0        | 0 / 0 / 0         |
| Participants with an intervention component withdrawn for an adverse event                                                                                                                                                                                                                                                                                  | 1                 | 2                | 3                 |
| Participants with an intervention component interrupted for an adverse event                                                                                                                                                                                                                                                                                | 1                 | 0                | 1                 |
| Protocol-driven dose reductions                                                                                                                                                                                                                                                                                                                             | 0                 | 0                | 0                 |
| AE relatedness: related / possibly related / not applicable                                                                                                                                                                                                                                                                                                 | 11 / 0 / 1 events | 6 / 1 / 0 events | 17 / 1 / 1 events |
| Gastrointestinal symptoms (events)                                                                                                                                                                                                                                                                                                                          | 7                 | 3                | 10                |
| Oral-solution palatability complaints (events)                                                                                                                                                                                                                                                                                                              | 3                 | 4                | 7                 |
| Fatigue (events)                                                                                                                                                                                                                                                                                                                                            | 1                 | 0                | 1                 |
| Concern about potential medication harm (events)                                                                                                                                                                                                                                                                                                            | 1                 | 0                | 1                 |
| The clinical team confirmed that all reported side effects were mild, consistent with CTCAE v5.0 Grade 1. Severity and relatedness summaries use reported events as the denominator. Event types summarize the reported descriptions; gastrointestinal events include those accompanied by fatigue. Each event is counted once in the event-type breakdown. |                   |                  |                   |

Additional operational measures: data completeness

The canonical platform audit contained 2,251 unique scan documents, 2,221 participant-date pairs and records from all 31 participants (Fig. 5; Supplementary Table 2). Pipeline status was recorded as successful for 2,214 of 2,251 documents (98.4%). Complete multispectral arrays were present in 2,121

tongue (94.2%), 2,149 face (95.5%) and 2,162 wrist (96.0%) documents. The realized biological reference layer included completed metabolomic, proteomic, glycomic, metagenomic, epigenomic (WGBS) and transcriptomic (RNA-seq) datasets, with assay-specific quality-control denominators retained for integration with longitudinal digital and clinical-functional data (Supplementary Table 2). Together, the source-backed digital corpus, clinical-functional assessments and completed molecular datasets formed the realized analytical stack for subsequent participant-level response analyses.

### Responder-focused decision architecture and clinical integration

The decision-support system operated throughout the study, linking participant-specific intervention exposure to timestamped digital inputs and computed CRS trajectories and posterior probabilities. KEEP, AUGMENT, SWITCH and STOPP states were generated and displayed to physicians for review; however, randomized arm allocation remained fixed, and these states did not trigger treatment reassignment or CRS-guided treatment adaptation during the eight-week pilot. The resulting workflow generated the source-backed longitudinal series, multidomain clinical anchors, safety records and audit trail needed for later individual-response analyses while retaining physician authority throughout the fixed-window study.

### Physician oversight and monitored close-out

Sheba's Sponsor's Office monitored the study. Pharmacy accountability closed without open items, and site close-out was completed in April 2026.

### Exploratory post-trial participation

In the audited source snapshot, all 30 participants completing TP3 contributed at least one optional post-TP3 scan (Wilson 95% confidence interval, 88.6–100.0), with a median of 14 additional scan-days (interquartile range, 10.25-18; range, 1–24). This exploratory result indicates broad willingness to continue digital participation after the scheduled study window (Fig. 4C).

## Discussion

### Translational contribution

The prespecified secondary outcomes establish that the combined treatment and measurement requirements of ELITE could be sustained over the eight-week intervention. Their scientific significance lies in the integrated trial system: documented intervention exposure, dense within-person observation, periodic clinical and biological anchors, safety oversight and treatment accountability were connected within one clinical workflow (Table 4). High completion and adherence are particularly relevant to a regimen requiring repeated administration throughout the day alongside daily digital participation. The resulting observations retain the treatment context and source records needed for subsequent analyses of individual response. Continued follow-up after treatment-component withdrawal or interruption shows that tolerability-related changes could be managed while preserving retention. Pharmacy accountability and monitored close-out completed this pathway from intervention delivery to an auditable longitudinal

record. ELITE's methodological distinction lies in combined coverage rather than any single assay: molecular, clinical-functional, endocrine, psychological and high-frequency digital observations were linked to documented intervention exposure within the same governed longitudinal record. This cross-scale linkage preserves both temporal density and biological depth, allowing subsequent analyses to relate evolving digital trajectories to measured functional and molecular change.

Table 4  
Translational responder framework and supporting evidence in ELITE

| Translational layer                                                                                                                                                                  | ELITE implementation                                                                                                              | Responder contribution                                                                     | Evidence in this study                                                                                    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------|
| Algorithmic polytherapy                                                                                                                                                              | Arm-specific metabolic regimen plus shared AB-01/AB-02, longitudinal digital monitoring and physician-supervised decision support | Defines the integrated treatment-and-monitoring context for individual response assessment | 31 randomized; all 31 received allocated treatment; monitoring architecture operated throughout           |
| High-frequency trajectory                                                                                                                                                            | Daily source-backed multimodal capture with CRS updates                                                                           | Supports within-person temporal response detection                                         | 1,816/1,978 required scan-days; median digital adherence 97.7% among 30 completers                        |
| Clinical and molecular anchor                                                                                                                                                        | Week 8 MIC-C plus periodic clinical-functional and molecular assessments                                                          | Provides an independent multidomain criterion for responder validation                     | Clinical-functional anchors and completed molecular layers collected                                      |
| Bayesian response state and physician review                                                                                                                                         | $\Pr(\Delta\text{CRS} \geq 0)$ mapped to KEEP, AUGMENT, SWITCH and STOPP states; states reviewed by physicians                    | Translates uncertainty into accountable clinical interpretation                            | Thresholds prespecified; states generated in the fixed-window pilot without algorithm-guided reassignment |
| Audit and safety                                                                                                                                                                     | Source provenance, adverse-event capture, external monitoring, pharmacy accountability and close-out                              | Makes response classification traceable and safety-aware                                   | 19 events in 17 participants; no serious events; monitored close-out completed                            |
| The framework links defined intervention exposure to high-frequency response detection, multidomain clinical confirmation, physician interpretation and an auditable evidence trail. |                                                                                                                                   |                                                                                            |                                                                                                           |

Multidomain intervention as a responder-discovery test bed

The randomized regimens provided mechanistically distinct pharmacological perturbations alongside the same prescribed AB-01 and AB-02 regimen. Because intervention-related changes may emerge on different timescales across metabolic, immune, muscular and cognitive domains, the longitudinal design was intended to characterize not only the magnitude of change, but also its onset, direction and persistence within each participant. Linking documented exposure to daily observations and periodic

clinical anchors makes these trajectories available for individual responder assessment. Such longitudinal data provide a basis for latent state-space modelling toward a world model of individual biology, in which treatment perturbations are represented as inputs and their effects on inferred biological states and trajectories can be examined.

#### Future validation and decentralized monitoring

The next validation stage will evaluate four extensions of the architecture: validation of CRS against Week 8 MIC-C response, including calibration, discrimination and missing-data performance; reconstruction of participant-specific latent state from heterogeneous, irregular observations with explicit uncertainty [8]; physician-supervised adaptive intervention using longitudinal response trajectories; and prospective shadow-mode evaluation of candidate computational models, including in silico simulations conducted concurrently with the trial, with model versions, inputs and predictions fixed before subsequent outcomes are observed. The MIC-C provides a prespecified clinical, functional and multi-omic criterion independent of the digital CRS inputs. The present study provides the source-backed observation density, chronology, biological anchors, response vocabulary and governance framework needed to test these questions prospectively.

Over the longer term, this approach could support decentralized monitoring that combines frequent digital measurements at home with periodic clinic-based biological assessments, including blood sampling. Paired digital and biological observations could be used to train and prospectively validate models that estimate selected biological trajectories between visits, with direct measurements providing calibration and reassessment points. This would extend personalized monitoring into everyday life and help determine when further testing or treatment adjustment is warranted.

#### Regulatory relevance and prospective implementation

Beyond participant adherence, ELITE documents the practical delivery of multidrug geroscience in adults without diabetes under institutional ethics, clinical-safety and investigational-product oversight. This provides an institutional example of how repurposed metabolic therapies can be studied for an explicit healthy-aging objective. The trial also combined Bayesian response-monitoring architecture, digital health technologies and decentralized data acquisition, connecting methods addressed in FDA materials [24–29] to delivery within a supervised clinical study. Importantly, the study architecture was designed, ethics approval obtained and enrollment completed before publication of the FDA's January 2026 draft guidance on Bayesian methodology [27]; site close-out likewise preceded the agency's April 2026 real-time clinical trials initiative [28, 29]. These features were therefore implemented prospectively in an active randomized study rather than retrospectively mapped onto emerging regulatory concepts. In ELITE, Bayesian response quantities were prespecified and translated into explicit KEEP, AUGMENT, SWITCH and STOPP states; participant-generated data retained source provenance; computational states remained traceable to their underlying observations; clinical interpretation remained physician-gated; and safety triggers, adverse-event capture and external monitoring were embedded in the workflow. Together, these elements provide a field-tested example of how Bayesian decision support,

high-frequency digital measurement and clinical governance can operate within a single regulated trial architecture.

Published geroscience research includes an investigation of metformin in older adults [30], the ANTHEM trial protocol evaluating metformin stratified by antecedent metabolic health [31], a pilot trial of metformin and influenza vaccine response [32], a semaglutide trial with epigenetic-age endpoints [33], and a rapamycin trial in normative-aging adults [34]. Metformin/SGLT2-inhibitor combinations have also been studied for diabetes treatment [35], while healthy-volunteer co-administration studies have addressed pharmacokinetic and drug-interaction questions [36, 37]. ELITE extends this context to randomized combination treatment within a geroscience cohort, integrating dense longitudinal monitoring with physician-supervised response interpretation. To our knowledge, ELITE is the first reported randomized geroscience trial to administer metformin in protocol-defined combination with either a GLP-1 receptor agonist or an SGLT2 inhibitor to adults without diabetes for an explicit geroscience rather than pharmacokinetic, drug-interaction or disease-treatment objective. A structured search of PubMed/MEDLINE and ClinicalTrials.gov on 4 August 2026 screened 396 published records and 30 registry records; targeted verification on 22 August 2026 and an updated search on 29 September 2026 identified additional single-agent geroscience studies, including METFORAGING [38], and healthy-volunteer co-administration studies, but no previously published randomized trial matching this combination-treatment context and geroscience objective. This was a structured, targeted literature search rather than a systematic review.

### Study limitations and next-stage design

Generalizability is limited by the small cohort ( $n = 31$ ), single center, approximately eight-week intervention and selection of participants able to use the required technology. The cohort was 71% male, and most participants came from SHARP; those recruited from SHARP had completed that study before starting ELITE, but prior SHARP intervention exposure was not characterized by ELITE arm. No minimum interval between SHARP completion and ELITE initiation was required. The open-label design permits performance and assessment bias, and the absence of an inactive comparator limits causal interpretation of within-arm changes. The planned definitive study extends the architecture to approximately 200 participants over 12 months across three countries, with a matched comparator, masked outcome adjudication and adaptive reassignment. The single discontinuation related to inability to complete home scanning independently also suggests that future digitally intensive trials should assess task-level capability during screening rather than relying only on device access or general willingness to participate.

Principal strengths of the study include protocolized intervention exposure, dense within-person sampling, explicit separation of evidence layers, a prespecified response-state vocabulary, physician gating and a traceable path from capture through close-out. Because the cohort was predominantly male, broader validation should consider potential sex-related differences in response to metabolic interventions. Within its prespecified scope, recruitment reached target, retention and adherence

exceeded protocol benchmarks, dense multimodal measurement was sustained, and pharmacy, monitoring, safety and database close-out were completed.

## Conclusion

ELITE establishes the operational foundation for response-guided geroscience by connecting multicomponent treatment, dense longitudinal observation, periodic biological and functional anchoring, and clinical governance in one trial system. The prespecified secondary outcomes demonstrate that participants could sustain this protocol, with high completion and adherence and no serious adverse events. These findings support larger, longer trials of individual response validation and physician-supervised treatment adaptation. In this architecture, probabilistic response computation, physician review, safety oversight and monitored close-out remain linked within a provenance-preserving clinical system. The field-tested substrate therefore supports the next stage of the program: prospective validation of individual response classification, reconstruction of changing participant state between clinic visits, and physician-authorized adaptive intervention. More broadly, ELITE provides a practical template for geroscience trials seeking to move from sparse population-average measurements toward attributable longitudinal within-person response under clinical and regulatory governance.

## Abbreviations

AE  
adverse event  
CRS  
Composite Response Score  
eCRF  
electronic case-report form  
GLP-1  
glucagon-like peptide-1  
IQR  
interquartile range  
MIC-C  
Muscular-Immune-Cognitive Composite  
SAE  
serious adverse event  
SD  
standard deviation  
SGLT2  
sodium-glucose cotransporter-2  
TP  
time point.

# Declarations

## Ethics approval and consent to participate

The study was approved by the Sheba Helsinki Committee / institutional review board (2361-25-SMC) on 2 September 2025. All participants provided informed consent.

## Declaration of interests

B.B. and N.P. are affiliated with ANI BIOME PBC, which developed the platform, response-state architecture and investigational AgeBiotics products described in this work. B.B. and N.P. are named inventors on patents and patent applications held by or assigned to ANI BIOME PBC relating to multispectral measurement technology, the digital clinical-trial platform and associated computational methods, and compositions related to the AB investigational products. M.P.S. and E.B. serve as scientific advisers to ANI BIOME PBC. V.N.G. served on advisory panels for Retro Biosciences, Dior, LongeVC and Thorne. The Regents of the University of California are the sole owner of patents and patent applications directed at epigenetic biomarkers, including GrimAge and PhenoAge, for which S.H. is a named inventor. S.H. is a founder of the non-profit Epigenetic Clock Development Foundation, which licenses these patents. S.H. was a Principal Investigator at Altos Labs until March 2026 and holds equity in Altos Labs, Inc. The remaining authors declare no competing interests.

## Authors' contributions

B.B. and N.P. developed the platform and methodology, analyzed the feasibility data and drafted the manuscript. T.S. served as principal investigator, with E.B. and A.G. as co-principal investigators. E.B. led clinical execution, and A.G. and T.S. contributed to regulatory and clinical operations and interpretation. C.C.H. contributed to clinical assessment and participant care. M.P.S., V.N.G., S.H. and B.K.K. contributed scientific interpretation and critical revision. All authors read and approved the final manuscript.

## Acknowledgments

The authors thank the ELITE participants; the Sheba Medical Center clinical, pharmacy, monitoring and data-management teams; YonaLink and trial-operations partners involved in electronic case-report-form closure; and assay and analysis collaborators. The study was funded by ANI BIOME PBC, with additional financial support from an award made by the XPRIZE Foundation upon ANI BIOME PBC's selection as an XPRIZE Healthspan Semifinalist. ANI BIOME PBC also provided the platform and investigational products and, through its affiliated authors, led the study design, data analysis and manuscript preparation. The Sheba Fund for Health Services and Research served as regulatory sponsor of record and provided clinical and regulatory oversight.

## Supplementary information

Supplementary Table 1 maps protocol objectives and outcomes to their reporting status. Supplementary Table 2 summarizes participant measurements, biospecimen types and collection time points. Supplementary Table 3 presents the prespecified CRS response-state thresholds. Supplementary Table 4 summarizes regulatory and monitoring evidence. Supplementary Table 5 summarizes aggregate treatment and component-specific adherence; Supplementary Table 6 reports overall participant-level adherence.

## Declaration of generative AI and AI-assisted technologies in the writing process

During preparation of this manuscript, generative AI tools were used to assist with language editing and document preparation. The authors reviewed all outputs and take full responsibility for the final content.

# Availability of data and materials

De-identified participant-level data and analytical code are not publicly available because access is subject to ethics, sponsor, privacy, participant-consent and data-use requirements. Requests may be considered subject to applicable approvals.

## References

1. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G (2023) Hallmarks of aging: an expanding universe. *Cell* 186:243–278. <https://doi.org/10.1016/j.cell.2022.11.001>
2. Kennedy BK, Berger SL, Brunet A et al (2014) Geroscience: linking aging to chronic disease. *Cell* 159:709–713. <https://doi.org/10.1016/j.cell.2014.10.039>
3. Rolland Y, Sierra F, Ferrucci L et al (2023) Challenges in developing geroscience trials. *Nat Commun* 14:5038. <https://doi.org/10.1038/s41467-023-39786-7>
4. Ahadi S, Zhou W, Schüssler-Fiorenza Rose SM et al (2020) Personal aging markers and ageotypes revealed by deep longitudinal profiling. *Nat Med* 26:83–90. <https://doi.org/10.1038/s41591-019-0719-5>
5. Izmailova ES, Wagner JA, Perakslis ED (2018) Wearable devices in clinical trials: hype and hypothesis. *Clin Pharmacol Ther* 104:42–52. <https://doi.org/10.1002/cpt.966>
6. Berry DA (2006) Bayesian clinical trials. *Nat Rev Drug Discov* 5:27–36. <https://doi.org/10.1038/nrd1927>
7. Duan N, Kravitz RL, Schmid CH (2013) Single-patient (n-of-1) trials: a pragmatic clinical decision methodology for patient-centered comparative effectiveness research. *J Clin Epidemiol* 66:S21–S28. <https://doi.org/10.1016/j.jclinepi.2013.04.006>
8. Katsoulakis E, Wang Q, Wu H et al (2024) Digital twins for health: a scoping review. *npj Digit Med* 7:77. <https://doi.org/10.1038/s41746-024-01073-0>

9. Sheba Medical Center (2026) Sheba Healthspan Research Population (SHARP) trial - Sheba Longevity Center diagnostic and intervention protocol to lower biological age in older adults. ClinicalTrials.gov identifier NCT07596576. First posted 19 May 2026.  
<https://clinicaltrials.gov/study/NCT07596576>. Accessed 3 August
10. Barzilai N, Crandall JP, Kritchevsky SB, Espeland MA (2016) Metformin as a tool to target aging. *Cell Metab* 23:1060–1065. <https://doi.org/10.1016/j.cmet.2016.05.011>
11. Lincoff AM, Brown-Frandsen K, Colhoun HM et al (2023) Semaglutide and cardiovascular outcomes in obesity without diabetes. *N Engl J Med* 389:2221–2232.  
<https://doi.org/10.1056/NEJMoa2307563>
12. EMPA-KIDNEY Collaborative Group (2023) Empagliflozin in patients with chronic kidney disease. *N Engl J Med* 388:117–127. <https://doi.org/10.1056/NEJMoa2204233>
13. Smith RL, Soeters MR, Wüst RCI, Houtkooper RH (2018) Metabolic flexibility as an adaptation to energy resources and requirements in health and disease. *Endocr Rev* 39:489–517.  
<https://doi.org/10.1210/er.2017-00211>
14. Bartsch RP, Liu KKL, Bashan A, Ivanov PC (2015) Network physiology: how organ systems dynamically interact. *PLoS ONE* 10:e0142143. <https://doi.org/10.1371/journal.pone.0142143>
15. Picard M, Trumpff C, Burelle Y (2019) Mitochondrial psychobiology: foundations and applications. *Curr Opin Behav Sci* 28:142–151. <https://doi.org/10.1016/j.cobeha.2019.04.015>
16. XPRIZE Foundation. XPRIZE Healthspan Competition Guidelines, Version 2.2 (2025) December  
<https://xprize.org/prizes/healthspan/guidelines>
17. Hopewell S, Chan AW, Collins GS et al (2025) CONSORT 2025 statement: updated guideline for reporting randomised trials. *BMJ* 389:e081123. <https://doi.org/10.1136/bmj-2024-081123>
18. Hopewell S, Chan AW, Collins GS et al (2025) CONSORT 2025 explanation and elaboration: updated guideline for reporting randomised trials. *BMJ* 389:e081124. <https://doi.org/10.1136/bmj-2024-081124>
19. Eldridge SM, Chan CL, Campbell MJ et al (2016) CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *BMJ* 355:i5239. <https://doi.org/10.1136/bmj.i5239>
20. Liu X, Cruz Rivera S, Moher D et al (2020) Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nat Med* 26:1364–1374.  
<https://doi.org/10.1038/s41591-020-1034-x>
21. Cruz Rivera S, Liu X, Chan AW et al (2020) Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Nat Med* 26:1351–1363.  
<https://doi.org/10.1038/s41591-020-1037-7>
22. Hoffmann TC, Glasziou PP, Boutron I et al (2014) Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 348:g1687.  
<https://doi.org/10.1136/bmj.g1687>
23. Junqueira DR, Zorzela L, Golder S et al (2023) CONSORT Harms 2022 statement, explanation, and elaboration: updated guideline for the reporting of harms in randomised trials. *BMJ* 381:e073725.

<https://doi.org/10.1136/bmj-2022-073725>

24. U.S. Food and Drug Administration. Guidance for the use of Bayesian statistics in medical device clinical trials. Silver Spring: FDA; (2010)
25. U.S. Food and Drug Administration. Conducting clinical trials with decentralized elements: guidance for industry. Silver Spring: FDA; (2024)
26. U.S. Food and Drug Administration. Digital health technologies for remote data acquisition in clinical investigations: guidance for industry. Silver Spring: FDA; (2023)
27. U.S. Food and Drug Administration. Use of Bayesian methodology in clinical trials of drug and biological products: draft guidance for industry. January (2026) <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-bayesian-methodology-clinical-trials-drug-and-biological-products>. Accessed 20 July 2026
28. Food US, Administration D FDA announces major steps to implement real-time clinical trials. 28 April 2026. <https://www.fda.gov/news-events/press-announcements/fda-announces-major-steps-implement-real-time-clinical-trials>. Accessed 20 July 2026
29. Food US, Administration D (2026) AI-enabled optimization of early-phase clinical trials pilot program; request for information. Fed Regist 91:23100
30. Kulkarni AS, Brutsaert EF, Anghel V et al (2018) Metformin regulates metabolic and nonmetabolic pathways in skeletal muscle and subcutaneous adipose tissues of older adults. Aging Cell 17:e12723. <https://doi.org/10.1111/accel.12723>
31. Kumari S, Bubak MT, Schoenberg HM et al (2022) Antecedent Metabolic Health and Metformin (ANTHEM) Aging Study: rationale and study design for a randomized controlled trial. J Gerontol Biol Sci Med Sci 77:2373–2377. <https://doi.org/10.1093/gerona/ghab358>
32. Martin DE, Cadar AN, Panier H, Torrance BL, Kuchel GA, Bartley JM (2023) The effect of metformin on influenza vaccine responses in nondiabetic older adults: a pilot trial. Immun Ageing 20:18. <https://doi.org/10.1186/s12979-023-00343-x>
33. Corley MJ, Dwaraka VB, Pang AP et al (2026) Semaglutide slows epigenetic aging in a randomized trial of HIV-associated lipohypertrophy. Nat Commun 17:6606. <https://doi.org/10.1038/s41467-026-72861-3>
34. Moel M, Harinath G, Lee V et al (2025) Influence of rapamycin on safety and healthspan metrics after one year: PEARL trial results. Aging 17:908–936. <https://doi.org/10.18632/aging.206235>
35. Häring HU, Merker L, Seewaldt-Becker E et al (2014) Empagliflozin as add-on to metformin in patients with type 2 diabetes: a 24-week, randomized, double-blind, placebo-controlled trial. Diabetes Care 37:1650–1659. <https://doi.org/10.2337/dc13-2105>
36. Hausner H, Derving Karsbøl J, Holst AG, Jacobsen JB, Wagner FD, Golor G, Anderson TW (2017) Effect of semaglutide on the pharmacokinetics of metformin, warfarin, atorvastatin and digoxin in healthy subjects. Clin Pharmacokinet 56(11):1391–1401. <https://doi.org/10.1007/s40262-017-0532-6>

37. Macha S, Dieterich S, Mattheus M, Seman LJ, Broedl UC, Woerle HJ (2013) Pharmacokinetics of empagliflozin, a sodium glucose cotransporter-2 (SGLT2) inhibitor, and metformin following co-administration in healthy volunteers. *Int J Clin Pharmacol Ther* 51(2):132–140. <https://doi.org/10.5414/CP201794>

38. Marcelo-Calvo C, Esteban-Cantos A, Jurado F et al (2026) Metformin and epigenetic age in non-diabetic older people with HIV in Madrid (METFORAGING): a double-blind, randomised, placebo-controlled, pilot trial. *EClinicalMedicine* 95:103874. <https://doi.org/10.1016/j.eclinm.2026.103874>

Figures

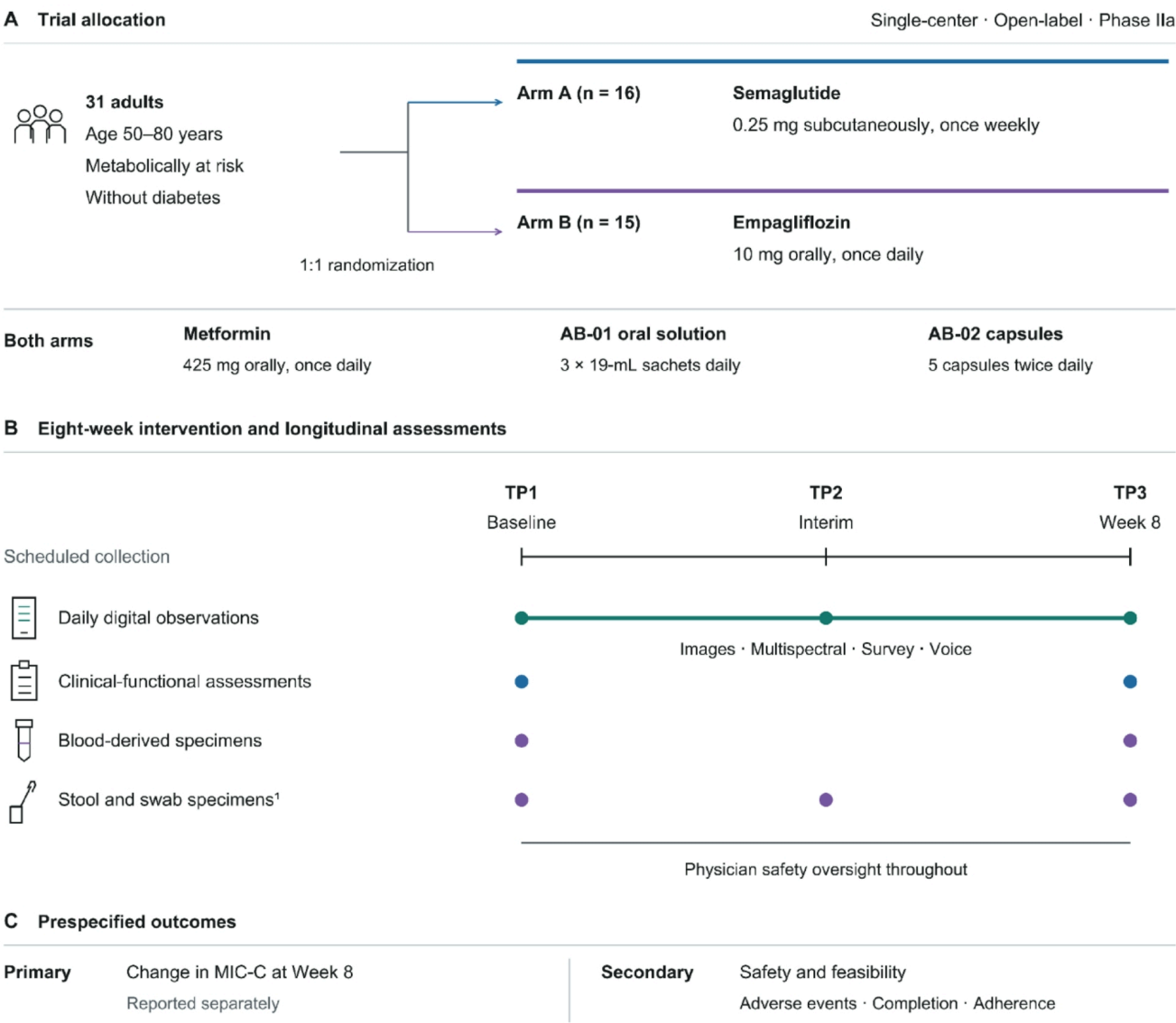


Figure 1

Overall design of the ELITE trial. A, Thirty-one metabolically at-risk adults aged 50–80 years without diabetes were randomized 1:1 in a single-center, open-label phase IIa pilot at Sheba Medical Center. Arm A received semaglutide and Arm B received empagliflozin; both arms received metformin and the investigational AgeBiotics formulations AB-01 and AB-02 at the schedules shown. B, The approximately eight-week intervention combined daily digital observations with clinical-functional assessments and blood-derived specimen collection at baseline (TP1) and Week 8 (TP3). Stool and swab specimens were collected at TP1, the intermediate time point (TP2), and TP3. The TP2 position is schematic and does not specify an exact week. Physician safety oversight continued throughout the intervention. C, The primary endpoint was change in the Muscular-Immune-Cognitive Composite (MIC-C) at Week 8. This report addresses the prespecified secondary safety and feasibility outcomes. <sup>1</sup>Swabs comprised tongue swabs and vaginal swabs in female participants. Symbols denote scheduled collection, not individual observations. TP denotes time point.

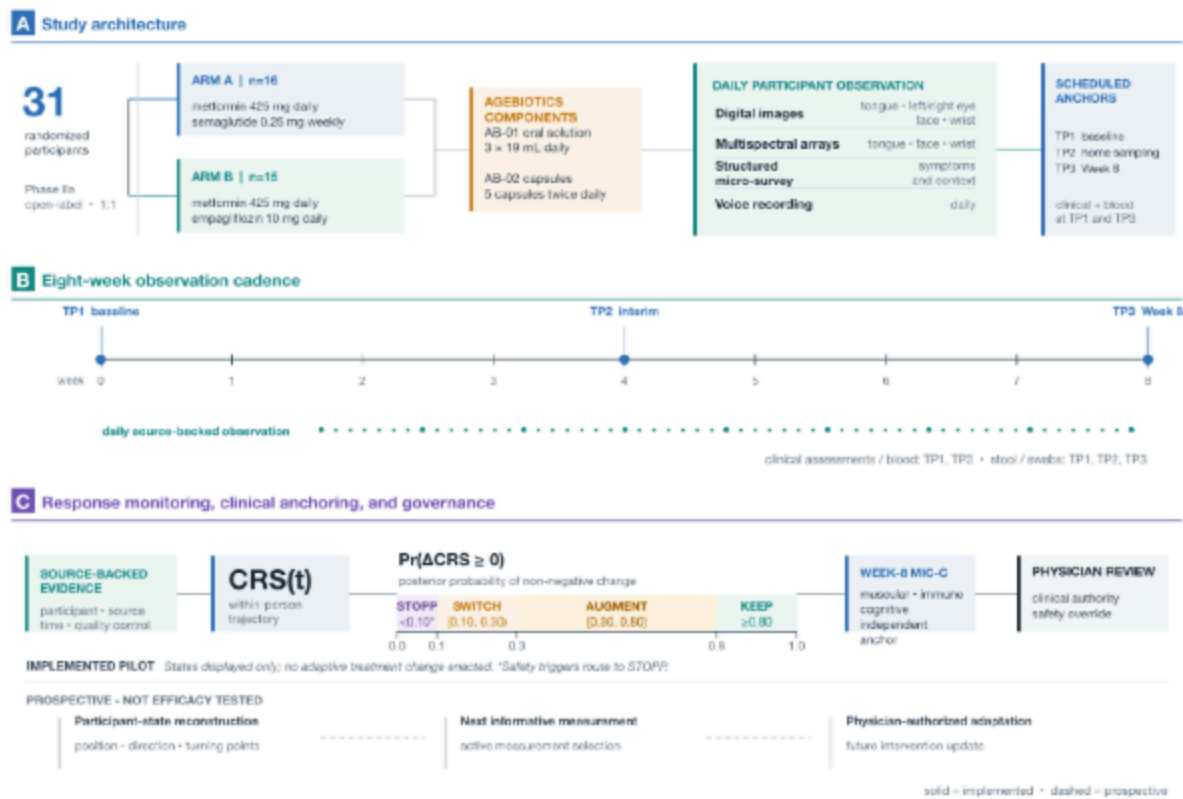


Figure 2

ELITE end-to-end measurement and response architecture. A, Thirty-one participants were randomized to Arm A, comprising metformin plus semaglutide, or Arm B, comprising metformin plus empagliflozin. Both arms received AB-01 oral solution and AB-02 capsules, together with daily digital observation and scheduled clinical-functional and molecular assessments. B, The eight-week window combined daily source-backed digital observation with clinical-functional assessments and blood collection at TP1 and TP3, alongside home stool and swab sampling at TP1, TP2 and TP3. C, Timestamped participant evidence was summarized through the CRS, CRS(t), and the posterior probability of non-negative

change, which was mapped to the prespecified KEEP, AUGMENT, SWITCH and STOPP response states. Response states were referenced against the prespecified Week 8 MIC-C clinical anchor and remained subject to physician interpretation. Solid arrows denote components implemented in the fixed-window pilot. Dashed elements denote prospective state-reconstruction and adaptive-action layers. CRS denotes Composite Response Score; MIC-C denotes Muscular-Immune-Cognitive Composite; TP denotes time point.



Figure 3

Participant flow, intervention allocation and operational timeline. A, Of 282 individuals assessed through the tracked SHARP registry and pilot cohort, 22 enrolled, and nine additional participants enrolled through community and clinical-network recruitment. Thirty-one participants were randomized to Arm A (n=16) or Arm B (n=15), both also assigned AB-01 and AB-02; pharmacy dispensing records document that all 31 received allocated treatment. Thirty participants completed follow-up: 16 of 16 in Arm A and 14 of 15 in Arm B. One Arm B participant discontinued after a few days of treatment because she could not independently complete home scanning; none were lost to follow-up in either arm. The 25/31 figure denotes the per-protocol population defined in Methods. B, Ethics approval, the enrollment and randomization period, the TP1-to-TP3 intervention period, and monitored site close-out and final-report milestones. TP denotes time point.

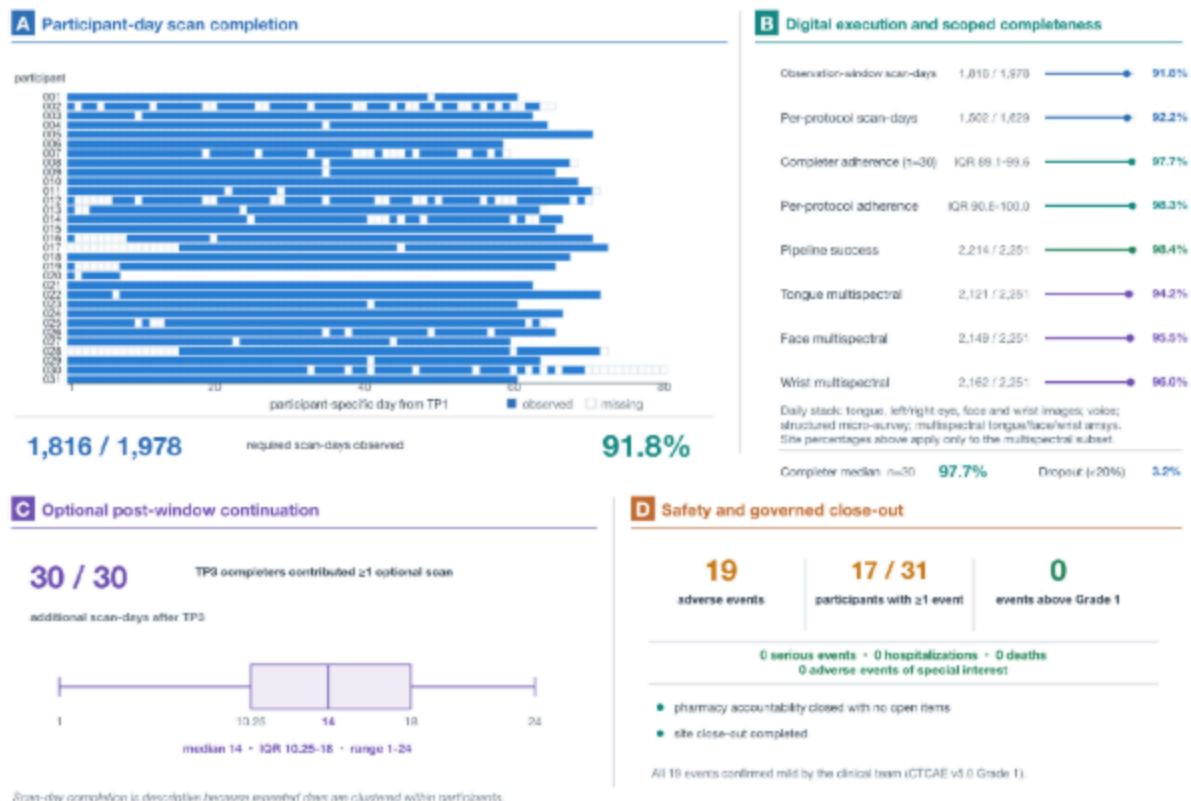


**Figure 4**

**Clinical and operational execution.** A, Participant-day heatmap of observed and missing scan-days within each participant's baseline-to-TP3 or withdrawal window; days beyond each window are unfilled. Overall, 1,816 of 1,978 expected scan-days were observed (91.8%). B, Digital capture and processing: 92.2% per-protocol scan-day completion, median participant-level digital adherence of 97.7% among 30 completers and 98.3% among 25 per-protocol participants, 98.4% pipeline success, and 94.2-96.0% spectral completeness across tongue, face and wrist arrays. C, All 30 completers contributed optional post-TP3 scans, with a median of 14 additional scan-days (interquartile range, 10.25-18; range, 1-24). D, Nineteen adverse events occurred in 17 of 31 participants, with no serious adverse events, hospitalizations, deaths or adverse events of special interest. Pharmacy accountability and monitored site close-out were completed. IQR denotes interquartile range; TP denotes time point.

