Skepthical review: *Brain Microstructural Pattern Age Acceleration (BMPAA) in Long-Lived Bats: Disentangling Age-Related, Sex-Related, and Origin-Specific Signatures*

Summary

The manuscript proposes Brain Microstructural/Mean Diffusivity Pattern Age Acceleration (BMPAA): a PCA-based representation of brain-wide mean diffusivity (MD) covariance across 24 atlas ROIs in Egyptian fruit bats, followed by regression of PC scores on DNAmAge, sex, and colony of origin, with regression residuals interpreted as covariate-adjusted "age-acceleration"-like deviations (Secs. 2.3–2.4, 3.3). In an analytical sample of $\sim \mathbf{30}$ bats, six PCs explain $\sim \mathbf{87\%}$ of MD variance, with PC3 associated with DNAmAge, PC1 with sex, and PC2 with colony (Sec. 3.3.2). However, a systematic parsing error prevents extraction of behavioral/cognitive metrics, so the planned brain–behavior tests of "cognitive resilience" are not performed (Secs. 2.2, 2.5, 3.2.1, 3.4, 4.2–4.4). As a result, the current version is best supported as a methods/feasibility report demonstrating a pipeline for deriving multivariate MD pattern scores and separating age-, sex-, and origin-related covariance patterns, but it does not yet validate BMPAA against cognition and remains limited by under-specified imaging preprocessing/atlas details, inconsistencies in cohort reporting, and insufficient robustness and interpretability analyses for PCA and downstream regressions.

Strengths

- Strong motivation for moving beyond univariate ROI MD to multivariate/system-level covariance patterns for studying aging (Sec. 1).
- Clear, conceptually simple pipeline (z-scoring ROIs → PCA → regress PC scores on covariates → residuals as adjusted deviations) that is easy to implement and potentially transferable to other species/modalities (Secs. 2.4, 3.3).
- Use of DNAmAge (rather than only chronological age) and a long-lived bat model provides a distinctive aging context (Secs. 1, 2.1.3).
- The results suggest separable sources of variance (age vs sex vs colony) across PCs (PC3/PC1/PC2), supporting the premise that covariance patterns can disentangle biological/structural influences (Sec. 3.3.2).
- Commendable transparency about the behavioral pipeline failure and its implications for the central hypothesis (Sec. 3.2.1; Discussion/Conclusions).

Major issues

1. **The central scientific objective (linking BMPAA to cognitive resilience) is not tested because behavioral feature extraction failed due to a systematic parsing error; consequently, none of the planned brain–behavior models described in Sec. 2.5 are executed (Secs. 2.2, 2.5, 3.2.1, 3.4, 4.2–4.4).** As written, several sections still frame the work as if resilience were evaluated, which overstates what the current results can support.

   Recommendation: Choose one of two paths and revise the manuscript consistently: (i) Fix the parsing/extraction pipeline, regenerate behavioral variables, and run the pre-specified models in Sec. 2.5 with full reporting (N per model, effect sizes, uncertainty, diagnostics), or (ii) if recovery is not feasible, explicitly reframe as a methods/feasibility paper. For option (ii), revise the Abstract, Introduction, Sec. 2.5 (tense and positioning), Results (Sec. 3.2.1), and Conclusions (Secs. 4.2–4.4) to clearly state that cognition/resilience analyses are future work and remove/down-weight claims implying demonstrated cognitive relevance.

2. **Imaging preprocessing, tensor fitting, ROI extraction, and atlas definition are under-specified, limiting reproducibility and interpretability of MD (Secs. 2.3.1–2.3.2).** Key missing elements include: motion/eddy-current and distortion correction, tensor fitting method, registration/normalization to atlas/template space, smoothing (if any), QC/exclusion criteria, handling of partial-volume/CSF contamination, and the provenance/validation of the 24-ROI bat atlas (including the full ROI list). The presence of truncated/zero-filled path-like strings further obscures the pipeline.

   Recommendation: Expand Sec. 2.3 into a stepwise, reproducible description: software and versions; preprocessing steps and parameters (e.g., eddy/motion, susceptibility correction if applicable); tensor fitting approach; template/atlas registration procedure; ROI summarization method (mean/median, voxel weighting); and QC metrics/thresholds. Add a table/appendix listing all 24 ROI labels with brief anatomical descriptions and atlas source/adaptation details for Rousettus. Replace internal absolute paths in Secs. 2.1.2/2.3.1/2.4.1 with human-readable placeholders and clarify whether any code/data will be shared.

3. **Cohort reporting is inconsistent ($N = 29$ vs $N = 30$; sex/origin counts differ across sections), and these discrepancies propagate into the stated MD matrix dimension (29×24 vs 30×24) used for PCA and regressions (Secs. 2.1.3, 3.1, 3.3.1; Table 1).** This undermines confidence in curation and in all downstream quantitative results.

   Recommendation: Reconcile the analytical cohort definitively: provide a single subject-flow summary (initial N, exclusions by reason, final N per modality/analysis). Update Sec. 2.1.3, Sec. 3.1, Table 1, and figure captions to match exactly (including

DNAmAge mean/SD). Explicitly state the final N used for (i) MD matrix construction, (ii) PCA fitting, and (iii) each regression.

4. **PCA component retention and stability are insufficiently justified given the small sample ($N \approx 30$) relative to 24 ROIs, and the reliance on Kaiser (> 1 eigenvalue) plus a variance threshold may over-retain components (Secs. 2.4.2, 3.3.1).** Without stability checks, it is unclear whether PC3's age association (and thus BMPAA on that component) is robust to sampling variation.

   Recommendation: Add robustness analyses for PCA (Secs. 2.4.2, 3.3.1): bootstrap or split-half resampling to quantify stability of eigenvalues and loadings (e.g., loading congruence/similarity), and include a parallel analysis (Horn) and/or scree-elbow justification. Report sensitivity of the key findings (PC3∼DNAmAge; PC1∼sex; PC2∼ origin) to retaining different numbers of PCs (e.g., 4–8). Frame conclusions explicitly as exploratory if stability is limited.

5. **Regression models linking PC scores to DNAmAge/sex/origin and defining BMPAA are under-specified and create avoidable inferential risk: the equation in Sec. 2.4.3 omits Origin_colony despite text claiming inclusion; coding for categorical predictors is not stated; standardized βs are not reproducible from the Methods; and multiple comparisons across PCs×predictors are not addressed (Secs. 2.4.3, 3.3.2).** Visual diagnostics are presented, but quantitative influence/leverage assessment is not reported despite small N.

   Recommendation: In Sec. 2.4.3, explicitly specify the full model used for each PC (e.g., $\mathbf{PC}_k \sim \mathbf{DNAmAge} + \mathbf{Sex} + \mathbf{Origin_colony}$), including intercept, error term, and the coding/reference levels for Sex and Origin. State exactly what was standardized (PC scores and/or DNAmAge) to produce "standardized β". In Sec. 3.3.2, report coefficient CIs and model fit (R^2; ideally partial R^2), and correct or control for multiplicity (e.g., FDR across 6 PCs for each predictor family, or a clearly stated exploratory stance). Add influence diagnostics (e.g., Cook's distance/leverage summary) and note whether any results depend on single points.

6. **Interpretability of PCs and of the term "age acceleration" is currently too thin.** PC3 is labeled "age-related" largely because it correlates with DNAmAge, but the manuscript does not show which ROIs drive PC1–PC3 nor how the sign/direction of the component maps to higher/lower MD in anatomically meaningful systems (Secs. 3.3.1–3.3.2, 3.4, 4.3–4.4). Additionally, interpreting residual sign as "more aged" is only justified for components with a defined aging direction and after resolving the arbitrary PC sign; it is not appropriate for PCs primarily associated with sex/origin (Sec. 2.4.3).

   Recommendation: Provide loading tables/plots for at least PC1–PC3 (top positive/negative ROI contributors) in Sec. 3.3.1–3.3.2 (main text or supplement), and interpret these patterns anatomically with relevant diffusion-aging context. Explicitly define the aging direction for the age-associated component (e.g., fix PC sign such that higher PC3 corresponds to higher predicted DNAmAge-associated MD pattern expression). Restrict "age acceleration" language to the age-associated component(s) and describe residuals for other PCs as "covariate-adjusted component expression," not aging.

7. **Sex and colony associations (PC1/PC2) may reflect technical/batch confounds (scanner/session differences, scan date, motion/SNR differences, processing batches) rather than biology, but technical covariates and potential confounding with age are not examined (Sec. 3.3.2; cohort description in Sec. 3.1).**

   Recommendation: Report available technical covariates (scan date/session, protocol changes, motion metrics, SNR/QC summaries) and test whether they correlate with sex/colony and/or with PCs. At minimum, add DNAmAge distributions by sex and by colony (and consider including chronological age if available) to assess confounding. If technical covariates exist, include them in sensitivity regressions or discuss explicitly as alternative explanations in Sec. 4.4.

Minor issues

1. The Abstract is truncated/incomplete and does not clearly communicate the key limitation that behavioral analyses were not performed, which can mislead readers about scope (Abstract).

   Recommendation: Complete and revise the Abstract to (i) finish all sentences, (ii) summarize the main quantitative findings (variance explained; which PCs associate with DNAmAge/sex/origin), and (iii) clearly state that behavioral/cognitive outcomes could not be analyzed due to a parsing error, positioning the manuscript as methodological/feasibility work in its current form.

2. DNAmAge is not described sufficiently for readers to interpret it (clock provenance, tissue/source implied by variable name, expected error, validation in this species), and the text sometimes conflates DNAmAge with chronological age (Secs. 1, 2.1.3, 2.4.3, 3.1, 4.4).

   Recommendation: Add a short description of the epigenetic clock used (training/validation reference, tissue, accuracy) in Methods (Sec. 2.1 or 2.4). Where possible, report DNAmAge vs chronological age correlation in this cohort (Sec. 3.1). Ensure terminology is consistent: BMPAA is adjusted for DNAmAge unless chronological age is also included.

3. Terminology and claims about residuals are imprecise: the manuscript says BMPAA is "independent" of covariates and sometimes of "chronological age," but residualization guarantees only in-sample orthogonality to included regressors under OLS with intercept, not statistical independence (Secs. 2.4.3, 3.3.2).

Recommendation: Replace "independent" with "linearly adjusted for" or "orthogonal to (in-sample)" and ensure claims match the actual model (DNAmAge vs chronological age). State explicitly that regressions included an intercept and specify categorical coding.

4. Figures 1–6 need readability and interpretability improvements: small fonts, unclear axes/units, and suboptimal encodings (e.g., pie charts). Several figure captions do not fully specify what is plotted (e.g., standardization status, N per panel), and Figure 2 does not provide actionable diagnostics about the behavioral parsing failure (Secs. Figures 1–6).

   Recommendation: Increase font/figure size and ensure axes include units and clear labels. Replace pie charts in Fig. 1 with bar charts (counts + percentages) and clearly distinguish DNAmAge vs chronological age. For Fig. 2, add a concise data-missingness/coverage summary (how many files/subjects affected) and, if appropriate, a minimal diagnostic (schema mismatch example) to make the failure mode reproducible. Add N and model formulas to relevant figure captions (Figs. 4–6).

5. Reproducibility is limited: software versions, key parameters, and a clear code/data availability statement are missing; internal absolute paths are not appropriate for a manuscript (Secs. 2.1–2.4).

   Recommendation: Add a Data/Code Availability section stating what will be shared (analysis scripts, ROI MD matrix, derived PC/BMPAA scores) and list software/packages and versions. Replace absolute internal paths with relative paths or descriptive placeholders.

6. Ethics/animal welfare approvals and key husbandry/anesthesia/handling details are not reported, which is required for animal imaging studies (Methods, Sec. 2.1).

   Recommendation: Add an Ethics/Animal Welfare subsection (e.g., Sec. 2.1.4) including oversight body, protocol numbers, housing conditions, anesthesia/handling procedures for MRI/DTI, and steps to minimize suffering, consistent with journal standards.

7. BMPAA acronym expansion is inconsistent (MD-specific vs microstructural-general), creating confusion about scope (Abstract, Secs. 1, 2.4, 4).

   Recommendation: Choose one expansion and apply consistently throughout. If intended as a general framework, explicitly state MD is the present instantiation and note other possible microstructural measures.

8. Internal cross-references and presentation of behavioral methods are inconsistent: Sec. 2.5.1 refers to behavioral metrics as being in Sec. 2.3, but Sec. 2.3 concerns DTI extraction; some behavioral analyses are described in past tense although not performed (Secs. 2.5.1, 3.2.1).

Recommendation: Correct section references (behavior: Sec. 2.2; imaging: Sec. 2.3; BMPAA: Sec. 2.4) and rewrite behavioral modeling text in future/conditional tense if not executed. Consider moving extensive planned behavioral model details to an Appendix/future-work subsection.

Very minor issues

1. Scattered typographic/OCR artifacts, truncated words, stray Markdown markers in headings, and visually distracting file-path artifacts reduce readability (e.g., Sec. 3.2.1 wording; stray "#" in headings; truncated terms in Sec. 2.4).

   Recommendation: Proofread and clean formatting: fix grammar, remove stray Markdown markers, repair truncated/split words, and replace non-informative path artifacts with concise descriptors.

2. Some numerical reporting appears inconsistent or overly precise (e.g., DNAmAge mean/SD differences across sections; borderline $p = 0.050$ described ambiguously), and repeated β/p values across sections may contain transcription errors (Secs. 2.1.3, 3.1, 3.3.2, 4.3).

   Recommendation: Standardize all cohort summary statistics to the final analytic sample, harmonize rounding, and verify that repeated statistics match the primary results tables/outputs. State the significance rule ($p < 0.05$ vs $p \leq 0.05$) and prefer effect sizes/CIs over thresholded language for borderline findings.

Key statements and references

- • **Principal Component Analysis of the standardized 30×24 Mean Diffusivity matrix identified six principal components with eigenvalues greater than 1 that collectively explained 87.33% of the total variance in brain-wide MD patterns, with individual variance contributions of 34.55% (PC1), 21.12% (PC2), 14.12% (PC3), 8.04% (PC4), 5.22% (PC5), and 4.28% (PC6).**
- *Reference(s):* (none)
- • **Regression of PC3 scores on DNAmAge, sex, and origin colony showed that PC3 scores were significantly and positively associated with DNAmAge (standardized regression coefficient $\beta = 0.56$, $p = 0.011$), while showing no significant association with sex or origin, indicating that PC3 captures a canonical age-related pattern of brain-wide MD change independent of these other covariates.**
- *Reference(s):* (none)
- • **Regression of PC1 scores on DNAmAge, sex, and origin colony revealed a significant association with sex, with male bats showing a negative association ($\beta = -2.94$, $p = 0.012$), indicating that PC1 represents a prominent**

pattern of MD covariance that differentiates male and female brains independently of age and origin.

- *Reference(s):* (none)
- • **Regression of PC2 scores on DNAmAge, sex, and origin colony demonstrated a significant association with origin colony, with bats from Herzeliya exhibiting a positive association ($\beta = 2.12$, $p = 0.005$), indicating that PC2 captures a large-scale pattern of brain microstructure that differs between colonies independently of age and sex.**
- *Reference(s):* (none)

Mathematical consistency audit

This section audits **symbolic/analytic** mathematical consistency (algebra, derivations, dimensional/unit checks, definition consistency).

Maths relevance: light

The paper's mathematics centers on (i) z-score standardization of region-wise MD values, (ii) PCA on the standardized subject×region matrix with component selection by eigenvalue > 1 and cumulative variance, and (iii) defining BMPAA as residuals from linear regressions of PC scores on DNAmAge, sex, and origin. There are few explicit equations and no step-by-step derivations; most checks are about definition consistency and whether interpretations logically follow from the stated linear-algebra/regression operations.

Checked items

1. ✖ **Cohort size vs MD matrix dimensions** (Sec. 2.1.3 (p.3) vs Sec. 3.1 (p.5–6) vs Sec. 3.2.2 & 3.3.1 (p.7))
    - **Claim:** The final analytical cohort size and the subject×region MD matrix used for PCA are consistently defined.
    - **Checks:** definition consistency, dimensional consistency
    - **Verdict:** FAIL; confidence: high; impact: critical
    - **Assumptions/inputs:** The PCA input matrix rows correspond exactly to the final cohort subjects., No additional exclusions occur after cohort definition.
    - **Notes:** Methods state final cohort $N = 29$, while Results state $N = 30$ and explicitly use a 30×24 matrix for PCA. This inconsistency leaves the PCA/regression residualization analytically ambiguous (sample size affects PCs, scores, and residuals).
2. ✔ **Region-wise z-scoring before PCA** (Sec. 2.4.1 (p.4) and Sec. 3.3.1 (p.7))

- **Claim:** Each region/column is z-scored across subjects so that PCA captures covariance/correlation structure rather than absolute MD scale differences.
- **Checks:** notation/definition consistency, sanity check
- **Verdict:** PASS; confidence: high; impact: minor
- **Assumptions/inputs:** Z-scoring means subtract column mean and divide by column standard deviation.
- **Notes:** The stated motivation and procedure (column-wise standardization) are internally coherent for PCA on multi-region measures.

3. ✔ **Kaiser criterion compatibility with standardization** (Sec. 2.4.2 (p.5) and Sec. 3.3.1 (p.7))

    - **Claim:** Retaining PCs with eigenvalues > 1 is an appropriate selection rule here.
    - **Checks:** assumption check, definition consistency
    - **Verdict:** PASS; confidence: medium; impact: minor
    - **Assumptions/inputs:** Eigenvalues referenced are for PCA on standardized variables (unit variance per region).
    - **Notes:** Because the paper explicitly z-scores each region, eigenvalues are on a unit-variance scale where the > 1 heuristic is at least internally consistent with the described preprocessing.

4. ⚠ **Variance explained by selected PCs** (Sec. 3.3.1 (p.7))

    - **Claim:** Six PCs collectively explain 87.33% of the variance, with listed per-PC percentages.
    - **Checks:** algebra consistency (sum check without recomputation), missing detail check
    - **Verdict:** UNCERTAIN; confidence: low; impact: minor
    - **Assumptions/inputs:** Percent variance explained sums to the cumulative explained variance for the retained components.
    - **Notes:** The claim is plausible, but the paper does not provide eigenvalues or enough intermediate quantities to verify internal arithmetic from the PDF alone.

5. ⚠ **Regression model form for PC scores** (Sec. 2.4.3 (p.5))

    - **Claim:** For each retained PC, a linear regression predicts PC_score from DNAmAge, Sex, and Origin_colony.
    - **Checks:** model specification completeness, notation consistency
    - **Verdict:** UNCERTAIN; confidence: medium; impact: moderate
    - **Assumptions/inputs:** A linear model with an intercept is fit., Sex and Origin are coded via a specified contrast scheme (not provided).

- **Notes:** The formula is given only in shorthand (tilde notation) with no explicit intercept/error term and no categorical coding description, so the exact meaning of coefficients and residual properties cannot be fully verified.

6. ✔ **BMPAA defined as regression residuals** (Sec. 2.4.3 (p.5) and Sec. 3.3.2 (p.8))
    - **Claim:** BMPAA_PCk equals the residual from regressing PCk scores on DNAmAge, sex, and origin.
    - **Checks:** definition consistency, sanity check
    - **Verdict:** PASS; confidence: high; impact: moderate
    - **Assumptions/inputs:** Residuals are computed from the fitted linear model per PC.
    - **Notes:** This definition is consistent across Methods and Results.

7. ✖ **Interpretation of BMPAA sign as "more aged"** (Sec. 2.4.3 (p.5); also narrative in Sec. 3.3.2 (p.8))
    - **Claim:** A positive BMPAA score indicates a brain-wide MD pattern that is more "aged" than expected given age/sex/origin; negative indicates "younger".
    - **Checks:** logical implication check, definition adequacy
    - **Verdict:** FAIL; confidence: high; impact: critical
    - **Assumptions/inputs:** Higher PC score corresponds to a more aged microstructural pattern., PC sign is fixed in a biologically meaningful direction.
    - **Notes:** Residual sign does not generally equate to "more aged" unless (i) the specific PC is an age-related axis with known direction, and (ii) the arbitrary PCA sign has been fixed accordingly. The paper later notes PC1/PC2 are associated with sex/origin; for those PCs, calling residuals "aged/younger" is not mathematically justified from the stated construction.

8. ✖ **"Independence" of BMPAA from covariates** (Sec. 2.4.3 (p.5); Sec. 3.3.2 (p.8); Fig. 5 caption (p.8))
    - **Claim:** BMPAA scores are independent of DNAmAge, sex, and origin after residualization.
    - **Checks:** statistical property check (orthogonality vs independence)
    - **Verdict:** FAIL; confidence: medium; impact: moderate
    - **Assumptions/inputs:** OLS residuals are used and the model includes an intercept.
    - **Notes:** Linear regression residuals are orthogonal (zero in-sample covariance) to included regressors under standard OLS with intercept, but this is not the same as statistical independence. The paper's language overclaims what the math guarantees.

9. ✖ **Use of DNAmAge vs "chronological age" in claims** (Sec. 2.4.3 and Sec. 2.5.2 (p.5); Conclusions Sec. 4.2–4.3 (p.9))
    - **Claim:** BMPAA is adjusted for and interpreted relative to chronological age.
    - **Checks:** symbol/definition consistency
    - **Verdict:** FAIL; confidence: high; impact: moderate
    - **Assumptions/inputs:** Chronological age equals DNAmAge or both are included/defined.
    - **Notes:** The regression uses DNAmAge as the age term, but multiple passages describe adjustment beyond "chronological age." Without defining chronological age or including it in the model, the math supports adjustment for DNAmAge only.

10. ✔ **Downstream behavioral model includes DNAmAge plus BMPAA** (Sec. 2.5.2 (p.5))
    - **Claim:** Behavioral_Metric ∼ BMPAA_PC1 + BMPAA_PC2 + ... + DNAmAge + Sex + Origin tests BMPAA effects beyond age/sex/origin.
    - **Checks:** algebraic redundancy/sanity check
    - **Verdict:** PASS; confidence: medium; impact: minor
    - **Assumptions/inputs:** BMPAA terms are residuals from regressions using DNAmAge/Sex/Origin with the same coding and intercept treatment.
    - **Notes:** Including DNAmAge/Sex/Origin alongside BMPAA is not inconsistent; BMPAA is (in-sample) orthogonal to those covariates if computed via OLS residuals with intercept. The inclusion may be redundant but does not violate linear-model algebra.

11. ⚠ **"Standardized regression coefficient" reporting** (Sec. 3.3.2 (p.8))
    - **Claim:** Reported β values are standardized regression coefficients.
    - **Checks:** definition completeness
    - **Verdict:** UNCERTAIN; confidence: medium; impact: minor
    - **Assumptions/inputs:** Standardization refers to scaling predictors/response to unit variance (exact convention unspecified).
    - **Notes:** The paper does not state the standardization convention for β in the regression stage; PCA input z-scoring alone does not imply regression coefficients are standardized.

12. ✖ **Section reference for behavioral metric definitions** (Sec. 2.5.1 (p.5))
    - **Claim:** Behavioral metrics are defined in Section 2.3.
    - **Checks:** cross-reference consistency
    - **Verdict:** FAIL; confidence: high; impact: minor
    - **Assumptions/inputs:** Section numbering is correct.

- **Notes:** Behavioral extraction is in Sec. 2.2, while Sec. 2.3 is DTI extraction. This is a documentation inconsistency that can confuse which definitions underpin later models.

Limitations

- The PDF provides no explicit equations for PCA (e.g., SVD/eigendecomposition), z-score formula, or regression/residual definitions; verification is limited to consistency of stated procedures and logical implications.
- Eigenvalues, loading matrices, and exact regression design matrices (intercept inclusion, categorical coding, handling of missingness) are not given, preventing full analytic verification of properties like residual centering and coefficient standardization from the PDF alone.
- This audit does not assess numerical correctness of reported percentages, p-values, or plots; it only checks symbolic/analytic consistency.

Numerical results audit

This section audits **numerical/empirical** consistency: reported metrics, experimental design, baseline comparisons, statistical evidence, leakage risks, and reproducibility.

Performed internal arithmetic and cross-text consistency checks on reported sample sizes, subgroup percentages, DNAmAge summaries, matrix dimensions, PCA variance totals, and repeated regression statistics. Several arithmetic/rounding checks passed, but there are major inconsistencies in cohort N across sections and in repeated reporting of regression statistics (PC1–PC3).

Checked items

1. ✖ **C01** (Page 1 Abstract: '...behavioral data from 30 ...')
    - **Claim:** The abstract states the study used behavioral data from 30 bats.
    - **Checks:** cross_section_sample_size_consistency
    - **Verdict:** FAIL
    - **Notes:** Inconsistent sample size across sections: methods (29) differs from abstract/results (30/30).
2. ✔ **C02** (Page 3 Table (EDA of final cohort): Sex counts and percentages)
    - **Claim:** Sex: Male 17 (58.6%), Female 12 (41.4%) for $N = 29$.
    - **Checks:** percent_from_counts
    - **Verdict:** PASS
    - **Notes:** Counts sum to N and percents match within rounding tolerance.
3. ✔ **C03** (Page 3 Table (EDA of final cohort): Origin counts and percentages)
    - **Claim:** Origin Colony: Aseret 15 (51.7%), Herzeliya 14 (48.3%) for $N = 29$.
    - **Checks:** percent_from_counts

- **Verdict:** PASS
- **Notes:** Counts sum to N and percents match within rounding tolerance.

4. ✔ **C04** (Page 3 Table (EDA of final cohort): DNAmAge mean (SD) and range)
   - **Claim:** DNAmAge (years): Mean (SD) 9.53 (1.67), Range $6.62 - 13.84$.
   - **Checks:** range_order_and_span
   - **Verdict:** PASS
   - **Notes:** Logical consistency holds (min < mean < max; positive span).

5. ✔ **C05** (Page 5 Results 3.1: cohort filtering numbers)
   - **Claim:** A total of 41 unique subjects were initially considered; final cohort 30; 11 excluded.
   - **Checks:** parts_to_total
   - **Verdict:** PASS
   - **Notes:** 41 equals 30 + 11 exactly.

6. ✔ **C06** (Page 6 Results 3.1 and Figure 1 caption: Sex distribution for $N = 30$)
   - **Claim:** Sex: 20 males (66.7%) and 10 females (33.3%) for cohort $N = 30$.
   - **Checks:** percent_from_counts
   - **Verdict:** PASS
   - **Notes:** Counts sum to N and percents match within rounding tolerance.

7. ✔ **C07** (Page 6 Results 3.1 and Figure 1 caption: Origin distribution for $N = 30$)
   - **Claim:** Origin Colony: 16 (53.3%) Aseret and 14 (46.7%) Herzeliya for $N = 30$.
   - **Checks:** percent_from_counts
   - **Verdict:** PASS
   - **Notes:** Counts sum to N and percents match within rounding tolerance.

8. ✔ **C08** (Page 6 Results 3.1: DNAmAge summary statistics)
   - **Claim:** Mean DNAmAge 9.55 years, SD 1.63 years, range 6.62 to 13.84 years.
   - **Checks:** range_order_and_span
   - **Verdict:** PASS
   - **Notes:** Logical consistency holds; cross-section deltas vs Page 3 table: mean $+0.02$, SD -0.04.

9. ✔ **C09** (Page 7 Results 3.2.2: MD matrix dimensions)
   - **Claim:** Subject-by-region MD matrix has dimensions 30 subjects × 24 ROIs.
   - **Checks:** dimension_multiplication

- **Verdict:** PASS
- **Notes:** $30 \times 24 = 720$ total cells (sanity check).

10. ✔ **C10** (Page 7 Results 3.3.1: PCA variance explained totals)

    - **Claim:** Six PCs explain 87.33% total variance; individual: 34.55%, 21.12%, 14.12%, 8.04%, 5.22%, 4.28%.
    - **Checks:** sum_of_percentages
    - **Verdict:** PASS
    - **Notes:** Sum of the six reported percentages equals the reported total (87.33%).

11. ✔ **C11** (Page 7 Results 3.3.1 and Figure 4 caption: number of PCs selected by Kaiser criterion)

    - **Claim:** Kaiser criterion (eigenvalue > 1) selected six principal components.
    - **Checks:** cross_reference_numeric_repeat
    - **Verdict:** PASS
    - **Notes:** The number of PCs (6) is consistent between the two cited locations.

12. ✖ **C12** (Page 8 Results 3.3.2 and Page 9 Results 4.3: PC3 association statistics)

    - **Claim:** PC3: standardized regression coefficient $\beta = 0.56$, $p = 0.011$ (repeated in conclusions).
    - **Checks:** cross_reference_numeric_repeat
    - **Verdict:** FAIL
    - **Notes:** Automated cross-reference check flagged a mismatch in repeated numeric claims.

13. ✖ **C13** (Page 8 Results 3.3.2 and Page 9 Results 4.3: PC1 association statistics)

    - **Claim:** PC1 associated with sex: $\beta = -2.94$, $p = 0.012$ (repeated in conclusions).
    - **Checks:** cross_reference_numeric_repeat
    - **Verdict:** FAIL
    - **Notes:** Automated cross-reference check flagged a mismatch in repeated numeric claims.

14. ✖ **C14** (Page 8 Results 3.3.2 and Page 9 Results 4.3: PC2 association statistics)

    - **Claim:** PC2 associated with origin colony: $\beta = 2.12$, $p = 0.005$ (repeated in conclusions).
    - **Checks:** cross_reference_numeric_repeat
    - **Verdict:** FAIL
    - **Notes:** Automated cross-reference check flagged a mismatch in repeated numeric claims.

15. ✖ **C15** (Page 8 Results 3.3.2: borderline p-value)

    - **Claim:** PC4 showed a weak association with origin ($p = 0.050$).
    - **Checks:** p_value_threshold_classification
    - **Verdict:** FAIL
    - **Notes:** Flagged as borderline because p is exactly 0.05; interpretation depends on the chosen significance rule.

Limitations

- Only parsed text was available; tables/figures (e.g., Table 1, Table 2) were not provided as machine-readable numeric content, limiting verification to arithmetic and cross-text consistency checks.
- No underlying datasets (e.g., dti_md_matrix.csv, behavioral_metrics.csv) were included, so PCA/regression results cannot be recomputed—only internal consistency of reported numbers can be tested.
- Plot-based numeric extraction is excluded by instruction; therefore any claims supported only by figures cannot be quantitatively audited here.