Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Prioritizing Geroscience Interventions for Human Translation: A Multi-Criteria Decision Analysis of 30 Candidates

P. Bouvet

b00108017@aus.edu

American University of Sharjah https://orcid.org/0009-0000-8758-2514

Method Article

Keywords: aging, geroscience, geriatric medicine, translational prioritization, multi-criteria decision analysis, healthspan, longevity interventions, evidence synthesis

Posted Date: March 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9110255/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Title: Prioritizing Geroscience Interventions for Human Translation: A Multi-Criteria Decision Analysis of 30 Candidates

Author: P. Bouvet

Abstract

Targeting fundamental mechanisms of aging could delay multiple chronic diseases simultaneously, but prioritizing among dozens of heterogeneous candidate interventions lacks a standardized evaluative framework. This study develops a multi-criteria decision analysis (MCDA) to quantitatively compare 30 geroscience interventions spanning repurposed drugs, nutraceuticals, regenerative strategies, and frontier modalities. Each intervention was scored across six prespecified translational domains (preclinical lifespan, healthspan, mechanistic conservation, human trial evidence, safety, accessibility) using predefined rubrics and aggregated with a weighted-sum model. Robustness was validated by Monte Carlo simulation (10,000 iterations) and sensitivity analysis under three stakeholder weighting schemes. The baseline ranking identified a stable top tier led by spermidine (3.7, 95% CI 3.44-3.93), rapamycin (3.6), and SGLT2 inhibitors (3.6). Sensitivity analysis revealed scenario-dependent leadership: metformin ranked first under regulator- and patient-focused weights, while stem cell therapy rose to third under investor-focused weights. No intervention occupied the high-readiness/high-impact quadrant, quantifying a persistent translational gap. An open-access interactive web tool accompanies the framework to support real-time exploration of ranking sensitivity.

Keywords: aging; geroscience; geriatric medicine; translational prioritization; multi-criteria decision analysis; healthspan; longevity interventions; evidence synthesis

Significance

Dozens of interventions claim to slow biological aging, yet clinicians, funders, and regulators lack a standardized way to compare translational readiness across heterogeneous evidence types. This study introduces a transparent, uncertainty-aware decision framework that shifts prioritization from maximal lifespan extension toward integrated human relevance. By explicitly quantifying trade-offs between efficacy, safety, and accessibility, the approach enables rational selection of candidates most suitable for near-term clinical testing and health-system deployment.

Introduction

Modern medicine has historically operated under a paradigm of treating individual diseases as they arise. Conditions such as heart disease, cancer, and neurodegeneration are addressed with distinct therapeutic strategies, a model that has yielded significant successes but faces mounting pressure from the demographic reality of aging populations. A transformative new discipline, geroscience, proposes a fundamental shift in this approach. It is built upon the hypothesis that the biological process of aging is the single greatest risk factor for the vast majority of chronic diseases and functional decline. Consequently, by targeting the fundamental mechanisms of aging itself, it may be possible to delay, prevent, or ameliorate multiple age-related diseases simultaneously, thereby extending "healthspan"—the period of life spent in good health, free from chronic disease and disability.

This hypothesis is grounded in decades of research that have identified a set of conserved biological processes that drive aging across species. These "hallmarks of aging," first codified in 2013 and expanded in 2023, include processes such as genomic instability, telomere attrition, epigenetic alterations, mitochondrial dysfunction, and cellular senescence, providing a mechanistic roadmap for therapeutic intervention [1,2]. The potential impact of this paradigm shift is profound. By intervening in the root causes of age-related decline, the goal is to compress morbidity into a shorter period at the end of life, allowing individuals to live not only longer but also healthier, more functional lives.

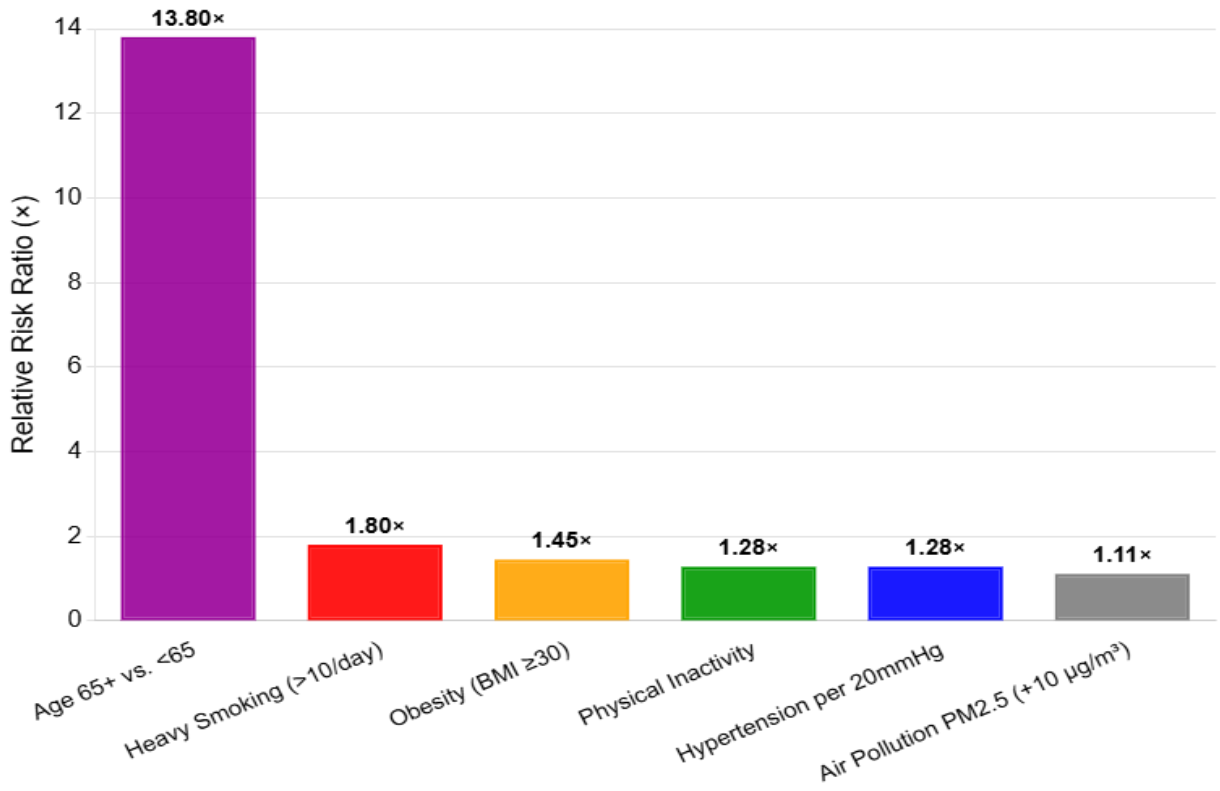
While the geroscience hypothesis provides a powerful conceptual framework, its translation into clinical reality is fraught with challenges. The field is rich with promising therapeutic targets and candidate interventions, from repurposed drugs to pig organ transplants, yet researchers, funders, and biopharmaceutical companies face the critical task of allocating finite resources to strategies with the highest probability of success in humans. This decision-making process is currently hampered by systemic hurdles, including the lack of formal recognition of aging as a treatable condition by major regulatory bodies and the difficulty of designing clinical trials with practical endpoints. Lifespan, while definitive in model organisms, is impractical for human studies, necessitating the development and validation of surrogate biomarkers that can reliably reflect biological age and predict healthspan outcomes.

The proliferation of candidate interventions, from repurposed drugs to advanced gene therapies, has created a "Tower of Babel" scenario, where progress in one area is difficult to weigh against advances in another. The absence of a common language or standardized metric for evaluating and comparing the translational potential of these diverse interventions is a fundamental, field-wide strategic bottleneck. This study aims to address this gap by implementing a systematic, multi-domain framework for the quantitative comparison of leading geroscience interventions. By establishing a common set of criteria and a transparent scoring methodology, this framework seeks to move the evaluation process from a qualitative art to a quantitative science, enabling a more rational and data-driven approach to prioritizing research and development efforts. This approach employs a Multi-Criteria Decision Analysis (MCDA), a methodology commonly used in health technology assessment (HTA) to structure and improve complex decision-making processes by explicitly evaluating multiple, often conflicting, criteria [3].

To contextualize the importance of this effort, we first benchmark the impact of biological aging against established risk factors. Data from meta-analyses (**Fig. 1**) reveal that advanced age (65+

versus <65) confers a 13.8-fold higher death ratio [4]—substantially higher than any modifiable lifestyle factor. In comparison, heavy smoking increases all-cause mortality risk by 1.8-fold [5], obesity (BMI≥30) by 1.45-fold [6], physical inactivity by 1.28-fold [7], hypertension by a similar magnitude per 20 mmHg elevation [8], and air pollution increases risk by 1.11-fold per 10 µg/m³ PM2.5 exposure [9]. While these lifestyle and environmental factors remain important for public health, the quantitative evidence underscores that chronological aging itself represents the single greatest risk factor. This epidemiological reality reinforces the geroscience hypothesis that targeting fundamental aging mechanisms offers the most promising pathway to compress morbidity and extend healthspan at the population level.

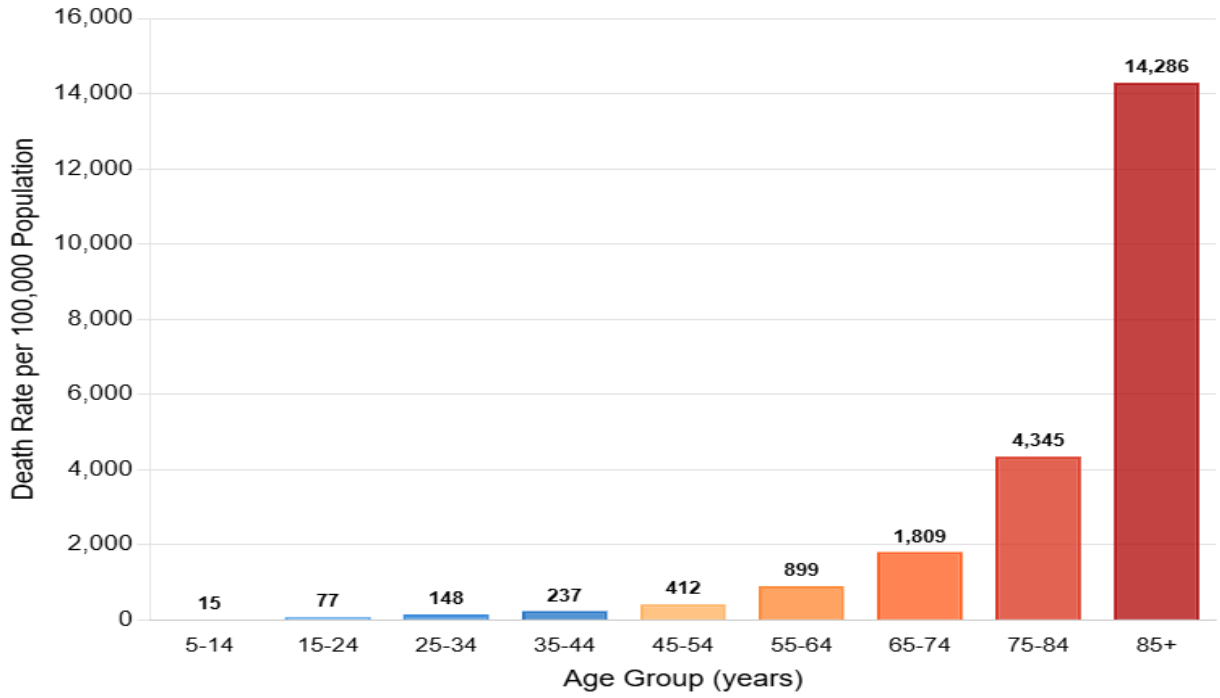
Fig. 1. Comparative Mortality Risk Factors



Relative risk ratios for all-cause mortality across major modifiable and non-modifiable risk factors. Advanced age (≥65 years vs. <65 years) confers a 13.8-fold increase in mortality risk (4), substantially exceeding all lifestyle and environmental factors.

Fig. 2 illustrates the exponential rise in death rates for each successive age group, particularly after age 55, with mortality reaching 3,900 per 100,000 at age 65+ versus 283 per 100,000 under 65—a 13.8-fold difference [4].

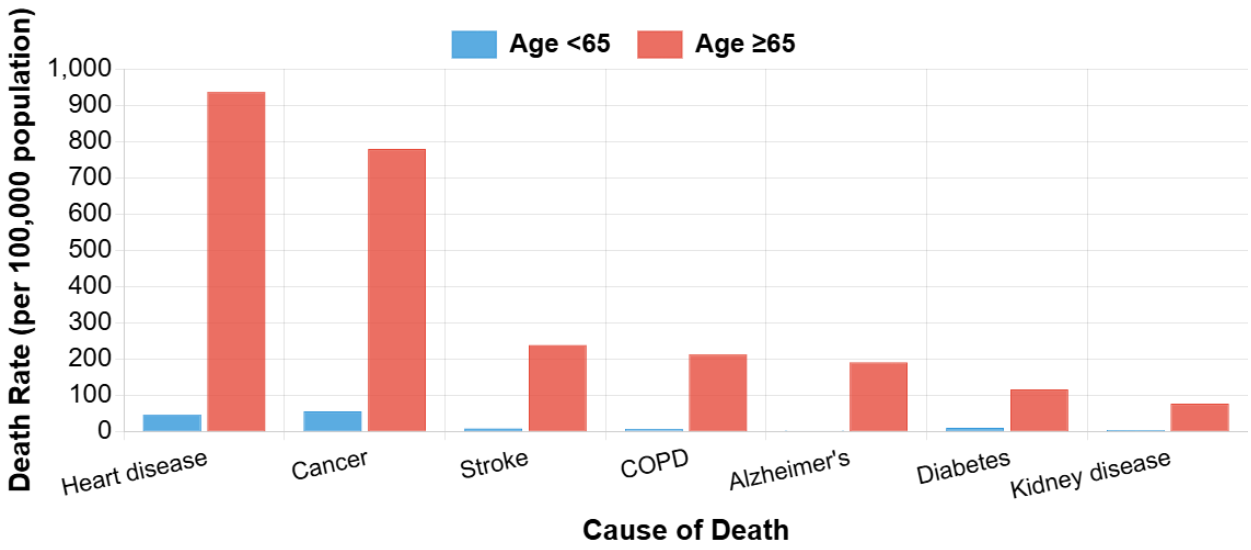
Fig. 2. US Mortality Rates by Age Group (2023, per 100,000)



Exponential increase in mortality rates across age groups in the United States (2023). Death rates per 100,000 population show significant acceleration after age 55, rising from 15 per 100,000 in ages 5-14 to 14,286 per 100,000 in ages 85+. The mortality rate for individuals aged 65+ (3,900 per 100,000) is 13.8 times higher than for those under 65 (283 per 100,000 [4].

Stratified mortality rates (**Fig. 3**) indicate that age is the dominant multiplier of mortality risk across disparate diseases. For every leading cause of death—heart disease, cancer, stroke, chronic lung disease, diabetes, kidney disease, and Alzheimer's—the risk in individuals ≥65 years is amplified at least 12-fold. The pattern is consistent despite different etiologies and organ systems, indicating a common upstream driver. The consistent risk jump with advancing age suggests these conditions are, in large part, manifestations of the same underlying process.

Fig. 3. Age-Specific Mortality Rates by Leading Causes of Death



Cause of Death	Rate <65 years	Rate ≥65 years	Age Ratio (≥65/<65)
Heart disease	45.9	935.7	**20.4×**
Cancer	55.1	778.7	**14.1×**
Stroke	7.9	237.7	**30.0×**
COPD	7.2	212.0	**29.6×**
Alzheimer's	0.5	190.0	**352.5×**
Diabetes	9.7	115.7	**12.0×**
Kidney disease	3.6	76.3	**20.9×**

Note: Death rates are expressed per 100,000 population. Age ratio represents the fold-difference in mortality rates between adults aged 65 and older compared to those under 65. COPD = Chronic Obstructive Pulmonary Disease (includes chronic lower respiratory diseases). Data excludes accidents and COVID-19. Source: CDC National Vital Statistics System, 2023.

This epidemiological dominance of age-related condition reveals a profound therapeutic mismatch. If one uses age-amplified noncommunicable disease (NCD) mortality as a proxy for aging-associated deaths, roughly three-quarters of global deaths fall into this category[10]. Yet, on the therapeutic side, essentially none of the FDA's approved indications explicitly target "aging" itself; instead, approved products overwhelmingly map to downstream single-disease endpoints—with oncology alone accumulating 573 FDA indication approvals (2000-2022)[11]. This asymmetry incentivizes siloed trials, complicates cross-program comparison, and helps explain the field's reliance on composite endpoints (e.g., TAME)[12]. It also underscores the need for a common, quantitative framework, such as the MCDA proposed here, to prioritize interventions under real-world regulatory constraints.

Intervention Selection and Evidence Synthesis

Thirty geroscience interventions were selected to represent the breadth of the current translational pipeline, spanning repurposed pharmacologics, nutraceutical and metabolic interventions, senolytics, cellular and regenerative strategies, and frontier modalities. For each intervention, the evidence base was assembled through a structured scoping-review approach designed to capture both preclinical longevity data and human translational signals, drawing on PubMed-indexed primary literature and reviews, clinical trial registries, regulatory summaries where relevant, and forward and backward citation tracing. The resulting domain scores are summarized in the scoring matrix, while their underlying sources and scoring rationales are documented in the evidence map provided in Supplementary Table S1. Detailed lifespan extraction data are provided in Supplementary Table S2, together providing the audit trail for the comparative ranking presented below.

Results

Lifespan Scoring Results

The structured evidence synthesis described above was first translated into preclinical lifespan scores, which provided the initial quantitative input to the broader multi-criteria decision analysis framework. **Table 1** presents the resulting net lifespan scores for the 30 geroscience interventions based on available preclinical data, ranked in descending order. Entries marked with (prog) indicate progeroid models, and those marked with (trans) indicate transgenic models.

Table 1. Lifespan score decomposition and net lifespan scores by intervention and model organism

Intervention	Model Organism	Average Lifespan Extension	Change Score	Organism Score	Reprod. Score	Net Lifespan Score
1. Rapamycin/mTOR Inhibitors	Mouse	13%	4.0	1.0	1.0	4.0
21. Stem Cell Therapy (FOXO3)	Rhesus macaques	11%	4.0	2.0	0.5	4.0
12. Autophagy Inducers (Spermidine)	Mouse	10%	4.0	1.0	1.0	4.0
16. Steroids (17α-estradiol)	Mouse	10%	4.0	1.0	1.0	4.0
21. Stem Cell Therapy (bone-marrow)	Mouse	28%	7.0	1.0	0.5	3.5
21. Stem Cell Therapy (CASIN-treated aged HSCs)	Mouse	25%	6.0	1.0	0.5	3.0
28. Anti-Inflammatory Therapy	Mouse	24%	6.0	1.0	0.5	3.0
27. Gut Microbiome (Probiotics)	C. elegans	24%	6.0	0.5	1.0	3.0

Intervention	Model Organism	Average Lifespan Extension	Change Score	Organism Score	Reprod. Score	Net Lifespan Score
25. Gene Therapy (Klotho)	Mouse	20%	6.0	1.0	0.5	3.0
3. Acarbose	Mouse	9%	3.0	1.0	1.0	3.0
13. Chloroquine	Mouse/Rat	9%	3.0	1.0	1.0	3.0
15. L-deprenyl	Mouse/Rat	9%	3.0	1.0	1.0	3.0
5. SGLT2 inhibitors	Mouse	6%	3.0	1.0	1.0	3.0
14. Glutathione Precursors	Mouse	5%	3.0	1.0	1.0	3.0
10. Mitochondria (Urolithin A)	Mouse	19%	5.0	1.0	0.5	2.5
18. Epigenetic Reprogramming (OSK/OSKM)	Mouse	19%	5.0	0.5	1.0	2.5
24. Exosome Therapy (hESCs)	Mouse	19%	5.0	1.0	0.5	2.5
26. Telomere Extension	Mouse	19%	5.0	1.0	0.5	2.5
19. Chemical Reprogramming	C. elegans	42%	9.0	0.5	0.5	2.3
25. Gene Therapy (FGF21)	Mouse	36% (trans)	9.0	0.5	0.5	2.3
21. Stem Cell (young HSCs)	Mouse	12%	4.0	1.0	0.5	2.0
24. Exosome Therapy (young plasma)	Mouse	12%	4.0	1.0	0.5	2.0
2. Metformin	Mouse/Rat	4%	2.0	1.0	1.0	2.0
6. Alpha-ketoglutarate (AKG)	Mouse	2%	2.0	1.0	1.0	2.0
8. Natural Senolytics (Fisetin)	Mouse	1%	2.0	1.0	1.0	2.0
17. Proteostasis & Nucleolar Function	C. elegans	25%	7.0	0.5	0.5	1.8
20. Immunotherapy (INK-ATTAC)	Mouse	24% (trans)	6.0	0.5	0.5	1.5
7. Senolytics (Dasatinib + Quercetin)	Mouse	6%	3.0	1.0	0.5	1.5
4. GLP-1 Receptor Agonists	C. elegans	10%	4.0	0.5	0.5	1.0
27. Gut Microbiome (FMT)	Mouse	14% (prog)	4.0	0.5	0.5	1.0
9. NAD+ Restoration (NMN)	Mouse	10% (prog)	4.0	0.5	0.5	1.0
9. NAD+ Restoration (NR)	Mouse	-2%	1.0	1.0	1.0	1.0

Table 1 presents the ranked geroscience interventions based on aggregated scores across three preclinical domains: magnitude of lifespan extension (Change Score), cross-species/strain reproducibility (Reprod. Score), and robustness of the underlying model organism evidence (Organism Score). The resulting Net Lifespan Score highlights a clear divide in preclinical performance. Rapamycin leads with consistent, reproducible effects in mammals, closely followed by advanced cellular therapies and autophagy inducers. In contrast, many interventions with human trial momentum—such as metformin, NAD+ precursors, and GLP-1 agonists—show only modest or inconsistent lifespan extension in model organisms.

Notably, interventions producing large lifespan gains in single studies or lower-order organisms—such as chemical reprogramming, proteostasis modulation, or microbiome manipulation in C. elegans—achieve lower net scores due to penalties for organismal relevance

and limited replication. Similarly, gene therapies showing large effects in progeroid or transgenic mice are down-weighted, as these models may not faithfully represent normative aging. This scoring logic intentionally prioritizes evolutionary conservation and wild-type mammalian evidence, filtering out impressive but potentially less translatable findings. The derived net lifespan score serves as one critical input domain in the full Multi-Criteria Decision Analysis, where human evidence, safety, and accessibility substantially reshape the overall prioritization landscape.

The Scoring and Ranking Matrix

This section translates the qualitative evidence synthesis across thirty interventions into a standardized quantitative framework using the prespecified 1–5 scoring rubric (Methods; **Tables 2A–2B**). **Table 3** summarizes side-by-side domain scores across the six translational domains and the resulting weighted composite score, from which the overall ranking is derived. A traceable evidence map supporting each domain score—including key studies and scoring inputs—is provided in **Supplementary Table S1**. Because minor decimal-level differences in composite scores can reshuffle within-tier ordering, rankings should be interpreted probabilistically, with inference focused on tier membership rather than small within-tier differences.

Table 3. Scoring and Ranking Matrix for Top 10 Geroscience Interventions (1-5 Scale)

Rank	Intervention	Lifespan (30%)	Healthspan (10%)	Conservation (10%)	Human Trials (20%)	Safety (20%)	Access (10%)	Weighted score
1	12. Autophagy Inducers (Spermidine)	4	4	5	2	4	4	**3.7**
2	1. Rapamycin/mTOR Inhibitors	4	4	5	3	3	3	**3.6**
2	5. SGLT2 inhibitors (Canagliflozin)	3	4	4	4	4	3	**3.6**
4	2. Metformin	2	4	4	3	5	5	**3.5**
5	3. Acarbose	3	4	4	3	3	5	**3.4**
5	14. Glutathione Precursors	3	3	4	3	4	4	**3.4**
5	27. Gut Microbiome Modulation	3	3	4	3	4	4	**3.4**
8	21. Stem Cell Therapy	4	4	4	3	3	1	**3.3**
8	15. L-deprenyl	3	3	4	3	4	3	**3.3**
8	10. Mitochondria (Urolithin A)	3	4	4	3	4	3	**3.3**
11	28. Anti-Inflammatory Therapy	3	3	4	3	3	4	**3.2**
12	30. Plasma dilution/apheresis	3	3	4	3	3	2	**3.0**

Rank	Intervention	Lifespan (30%)	Healthspan (10%)	Conservation (10%)	Human Trials (20%)	Safety (20%)	Access (10%)	Weighted score
12	16. Steroids (17α-estradiol)	4	3	4	1	3	3	**3.0**
14	6. Alpha-ketoglutarate (AKG)	2	3	4	2	4	4	**2.9**
14	8. Natural Senolytics (Fisetin)	2	3	4	2	4	4	**2.9**
16	9. NAD+ Restoration (NMN/NR)	1	3	5	3	4	3	**2.8**
17	4. GLP-1 Receptor Agonists	1	4	4	4	3	2	**2.7**
17	13. Chloroquine	3	2	4	2	2	4	**2.7**
17	7. Senolytics (D+Q)	2	4	4	3	3	2	**2.7**
20	24. Exosome Therapy	3	2	3	2	3	2	**2.5**
21	11. Mitochondria (Elamipretide/SS-31)	2	3	4	2	3	1	**2.4**
21	29. Young Blood Plasma and Parabiosis	3	3	3	2	2	1	**2.4**
21	26. Telomere Extension Therapy	3	3	4	2	2	1	**2.4**
24	25. Gene Therapy	3	3	4	1	2	1	**2.3**
25	18. Epigenetic Reprogramming	3	3	4	1	2	1	**2.2**
26	17. Proteostasis & Nucleolar Function	2	3	5	1	2	2	**2.1**
26	19. Chemical Reprogramming	2	2	4	1	2	2	**2.1**
28	22. Synthetic Tissues/Organs	2	2	3	2	2	1	**2.0**
29	20. Immunotherapy for Senescent Cells	2	2	3	2	1	1	**1.8**
29	23. Xenotransplantation	2	2	4	1	2	1	**1.8**

Table 4 summarizes the final rankings from the Multi-Criteria Decision Analysis (MCDA). For each intervention, it presents the rank, weighted score, and a concise overview of primary strengths and key limitations, providing qualitative context for the quantitative results.

Table 4. MCDA Ranking Results

Rank	Intervention	Score	Key Strengths	Key Limitations
1	12. Autophagy Inducers (Spermidine)	3.7	Natural, epidemiological support, lifespan impact	No ITP study
2	1. Rapamycin/mTOR Inhibitors	3.6	Lifespan impact	Dosing uncertainty, side effects

Rank	Intervention	Score	Key Strengths	Key Limitations
2	5. SGLT2 inhibitors	3.6	Multiple approved indications	Sex-specific effects
4	2. Metformin	3.5	TAME trial, decades of safety	Minimal preclinical efficacy
5	3. Acarbose	3.4	FDA approved, safety data	Sex-specific, GI effects
5	14. Glutathione Precursors	3.4	Positive human data, favorable safety	Limited evidence, ITP found minimal effects
5	27. Gut Microbiome Modulation	3.4	Personalized potential	Standardization challenges
8	21. Stem Cell Therapy	3.3	Proven regenerative capacity	Cost, temporary effects, delivery challenges
8	17. L-deprenyl	3.3	Some preclinical success	No human longevity data, mixed results
8	10. Mitochondria (Urolithin A)	3.3	Commercial availability, muscle benefits	Modest effects, mixed trial endpoints
11	28. Anti-Inflammatory Therapy	3.2	Multiple approaches available, targets inflammaging	Immunosuppression risk, lack of specific agent
12	30. Plasma dilution/apheresis	3.0	Novel mechanism, human proof-of-concept	Infrastructure requirements
12	16. Steroids (17α-estradiol)	3.0	Male-specific benefits	Safety concerns in primates, sex-specific effects
14	6. Alpha-ketoglutarate (AKG)	2.9	GRAS status, multi-pathway	Limited human data
14	8. Natural Senolytics (Fisetin)	2.9	Natural senolytic	Bioavailability issues, poor ITP results
16	9. NAD+ Restoration (NMN/NR)	2.8	Good safety, biomarker improvements	FDA regulatory issues, unproven long-term effects
17	4. GLP-1 Receptor Agonists	2.7	FDA approved, mortality benefits	Limited aging studies, gastrointestinal side effects
17	13. Chloroquine	2.7	Established safety	Paradoxical mechanism (Inhibits autophagy)
17	7. Senolytics (D+Q)	2.7	Addresses fundamental mechanism	Patient stratification needed
20	24. Exosome Therapy	2.5	Good safety profile, cell-free approach	Regulatory uncertainty, unproven systemic effects
21	11. Mitochondria (Elamipretide/SS-31)	2.4	Mitochondrial targeting	Injection required, mixed Phase 3 results
21	29. Young Blood Plasma and Parabiosis	2.4	Initial promise	Failed trials, FDA warnings, ethics
21	26. Telomere Extension Therapy	2.4	Strong mechanistic rationale	Cancer risk, lack of lifespan data
24	25. Gene Therapy	2.3	Targets fundamental mechanisms	Safety concerns, cost, delivery challenges
25	18. Epigenetic Reprogramming	2.2	Transformative preclinical potential	Early stage, safety risks (e.g., cancer)

Rank	Intervention	Score	Key Strengths	Key Limitations
26	17. Proteostasis & Nucleolar Function	2.1	Fundamental aging mechanism	Translation challenges, safety unproven
26	19. Chemical Reprogramming	2.1	Safer than viral delivery	Early stage
28	22. Synthetic Tissues/Organs	2.0	Transforming potential	Farther timeline, major technical hurdles
29	20. Immunotherapy for Senescent Cells	1.8	Powerful targeting	High cost, risk of side effects
29	23. Xenotransplantation	1.8	Addresses organ shortage	Compatibility issue

Strategic Profiles and the Translational Landscape

The quantitative MCDA results reveal a stratified translational landscape with three distinct strategic profiles. Translation-Ready interventions (weighted score >3.4), led by Spermidine (3.7), Rapamycin (3.6), and SGLT2 inhibitors (3.6), are characterized by strong preclinical data, favorable safety, and, crucially, existing human use or advanced trials. This group, which includes Metformin (3.5) and Acarbose (3.4), offers a balanced profile of proven mechanisms and lower regulatory risk, suitable for near-term clinical prioritization despite typically modest to moderate preclinical effect sizes.

A second tier of High-Promise/High-Risk candidates (score 2.5-3.4), such as Stem Cell Therapy (3.3) and Senolytics (2.7), shows compelling preclinical efficacy but is constrained by gaps in human evidence, safety questions, or delivery challenges. Their potential is significant but requires targeted research to de-risk translation. A 2025 study using FOXO3-enhanced stem cells demonstrates particular promise with 3.4 years biological age reduction in rhesus macaques[13], uniquely bridging the rodent-primate gap.

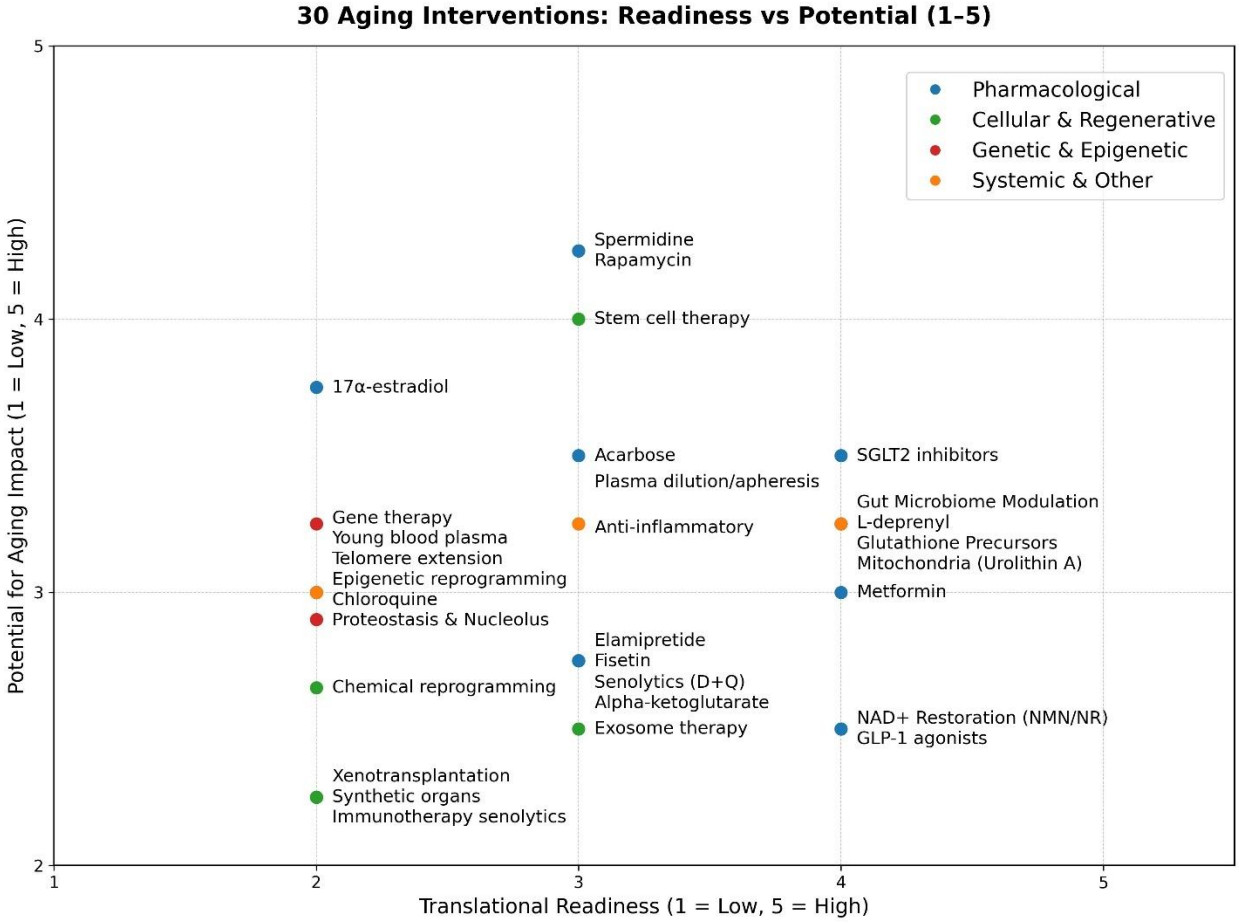
Finally, certain frontier modalities (score <2.5), including Epigenetic Reprogramming (2.2) and Gene Therapy (2.3), may possess transformative mechanistic potential. However, their low rankings directly reflect profound gaps in human safety data, delivery, and cost. Their position highlights the field's core translational challenge: the potentially most powerful interventions remain the farthest from clinical application, underscoring the critical need for continued fundamental research alongside the development of the leading pragmatic candidates.

Mapping Potential Impact vs. Translational Readiness

A central finding of this study is the distinct inverse relationship between "Potential Impact" and "Translational Readiness" for top ranked interventions as illustrated in **Fig. 4**. The analysis reveals that interventions with the highest preclinical impact—led by Spermidine, Rapamycin, and Stem Cell Therapy—cluster in the moderate-readiness quadrant (Impact ≥4.0, Readiness 3.0), constrained by safety risks or lack of large-scale human aging trials. Conversely, the most translationally mature interventions, such as SGLT2 inhibitors and Metformin, occupy the high-readiness but moderate-impact sector (Readiness ≥4.0, Impact 3.0–3.5).

Crucially, the "ideal" quadrant (High Impact/High Readiness) remains vacant, visually quantifying the field's translational gap. For instance, Metformin achieves a high readiness score (4.0) due to decades of clinical use, yet its impact potential is capped at 3.0, reflecting modest lifespan extension in preclinical studies[14]. In contrast, 17α-estradiol demonstrates higher impact potential (3.8) but lags significantly in readiness (2.0) due to safety concerns[15]. This distribution underscores a strategic trade-off: immediate clinical opportunities currently offer incremental gains, while the most revolutionary interventions remain farther from implementation.

Fig. 4: Translational Readiness vs Potential Impact

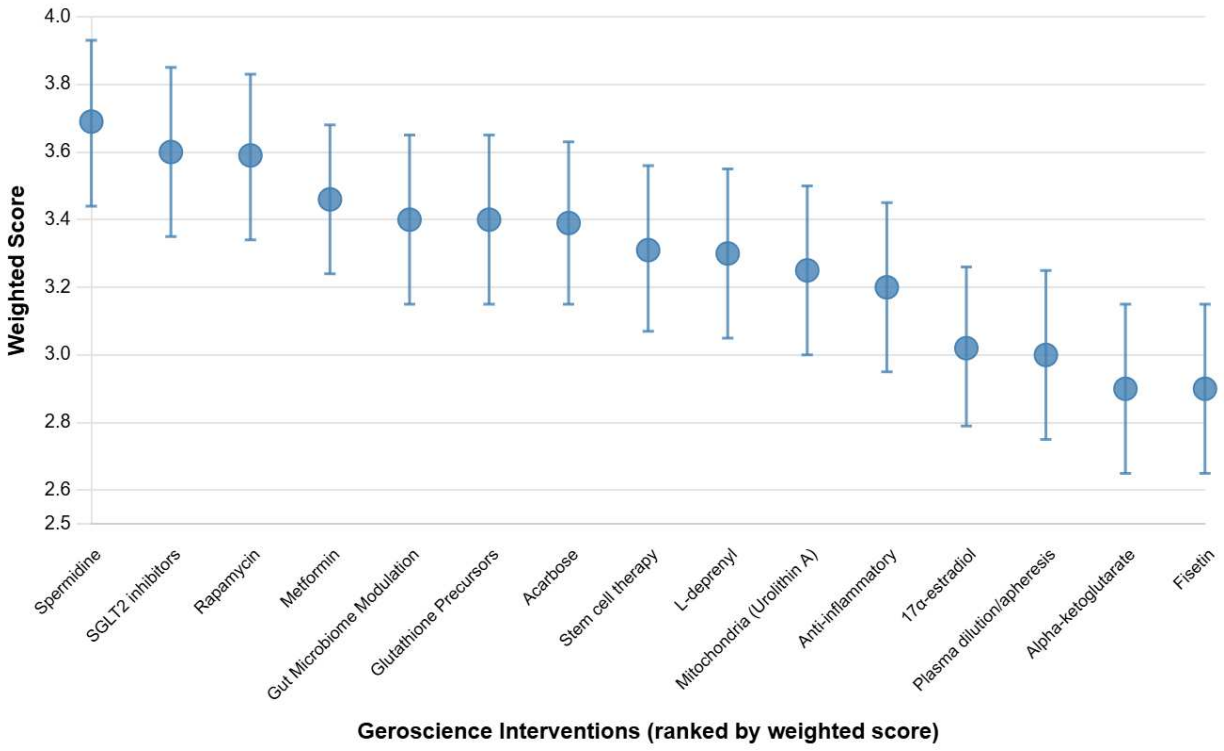


Plot mapping translational readiness against potential longevity impact across intervention categories. Each point represents one of 30 geroscience interventions, positioned according to current translational readiness (composite of human evidence and safety) and potential aging impact (weighted composite of preclinical lifespan, healthspan and mechanism efficacy). Interventions with highest preclinical impact (show lower translational readiness due to safety concerns and lack of human data, while most translationally ready interventions (e.g., Metformin) demonstrate modest preclinical efficacy. No intervention currently occupies the high-readiness/high-impact quadrant, highlighting the translational challenge.

Quantifying Uncertainty and Enhancing Quantitative Validation

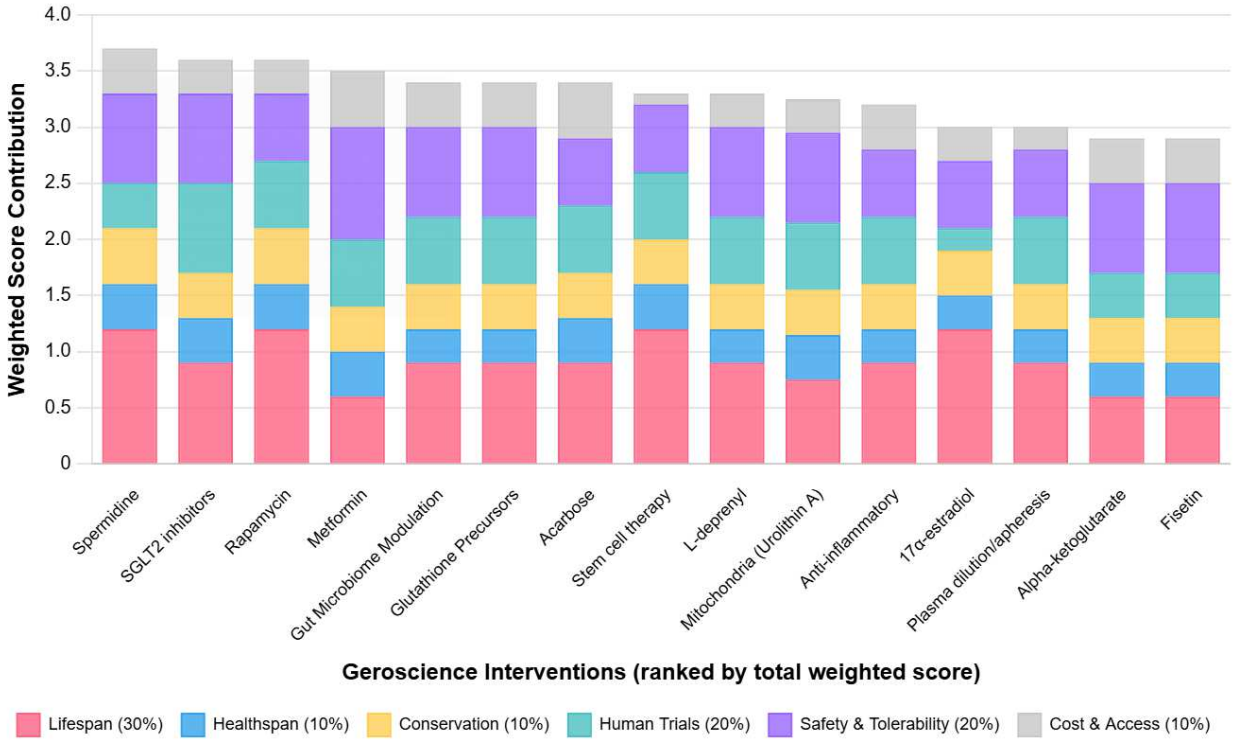
The Monte-Carlo uncertainty analysis revealed that the top tier remained tightly clustered and clearly separated from lower-ranked candidates, with substantial overlap in credible intervals reflecting similar composite evidence profiles. In the baseline MCDA, Spermidine ranked first (weighted score 3.7) and led the simulation with a narrow 95% credible interval of 3.44–3.93, followed by SGLT2 inhibitors (3.6) and Rapamycin (3.6) as the next most competitive interventions. Consistent with these results, weight-perturbation testing (±5% per domain) preserved the overall structure of the ranking with only minor positional changes among the leading candidates, indicating that the top tier is robust to shifts in stakeholder priorities. **Fig. 5** summarizes rank stability under Monte-Carlo sampling, and **Fig. 6** decomposes the weighted score into domain-wise contribution profiles to show which evidence pillars drive each intervention's composite performance.

Fig. 5. Rank Stability Analysis: Monte Carlo Simulation of Weighted Scores with 95% Credible Intervals



Monte Carlo analysis showing rank stability for top-tier geroscience interventions. Error bars represent 95% credible intervals from 10,000 simulations where domain scores were sampled from uniform distributions (±0.5 around assigned values). Spermidine leads with narrow confidence intervals (3.4-3.9), followed by SGLT2 inhibitors and Rapamycin. Top-tier interventions show tight clustering (weighted scores 3.4-3.7) with overlapping confidence intervals, indicating robust evidence profiles. Novel modalities in lower tiers show wider intervals reflecting greater uncertainty. Weight perturbation analysis (±5%) preserved overall ranking structure with minimal position changes among top performers.

Fig. 6. Domain-wise Contribution Profiles: Weighted Score Decomposition



Domain-wise contributions to final weighted scores for top 15 interventions. Each segment represents weighted contribution from six evaluation domains: Preclinical Efficacy - Lifespan (30%), Preclinical Efficacy - Healthspan (10%), Mechanism Conservation (10%), Human Trial Evidence (20%), Safety & Tolerability (20%), and Cost & Accessibility (10%). Top performers (Spermidine, SGLT2 inhibitors, Rapamycin) derive balanced strength from multiple domains, with particularly strong lifespan component. Metformin achieves high ranking primarily through exceptional safety despite modest lifespan evidence. Stem cell therapy shows strong preclinical promise but faces cost/accessibility challenges.

Sensitivity Analysis

To test robustness, interventions were re-scored under three alternative stakeholder weighting schemes (**Table 5**). The results show that the identity of the leaders is scenario-dependent, but the upper tier as a group is fairly stable: most baseline top performers remain within roughly the top 10–15 across schemes. Under regulator-focused weights (Safety 40%, Human Evidence 30%), rankings shift toward the most clinically de-risked programs, with Metformin rising to rank 1 and SGLT2 inhibitors to rank 2, while more efficacy-driven candidates (e.g., Rapamycin, Spermidine) fall into mid–upper positions (ranks 9 and 7, respectively). Under patient-focused weights, the same pattern holds, with Metformin (rank 1) and SGLT2 inhibitors (rank 2) leading due to combined tolerability, real-world familiarity, and accessibility.

In contrast, the investor-focused model (Preclinical Lifespan 40%, Healthspan 20%) systematically elevates preclinical-heavy candidates. Here, Stem Cell Therapy surges from rank 8

to rank 3 and 17α-estradiol jumps from rank 13 to rank 6, while Rapamycin maintains its strong position (rank 2), quantitatively illustrating the trade-off between preclinical upside that attracts capital and clinical readiness that determines near-term deployability.

Table 5. Sensitivity Analysis of Intervention Rankings with Alternative Weighting Schemes

Intervention	Baseline Weighted (Score, Rank)	Regulator-Focused Weights* (Score, Rank)	Investor-Focused Weights** (Score, Rank)	Patient-Focused Weights*** (Score, Rank)
Spermidine	3.7 (1)	3.4 (7)	3.9 (1)	3.5 (3)
Rapamycin	3.6 (2)	3.2 (9)	3.8 (2)	3.3 (8)
SGLT2 inhibitors	3.6 (3)	3.8 (2)	3.5 (4)	3.7 (2)
Metformin	3.5 (4)	4 (1)	3.3 (6)	3.9 (1)
Acarbose	3.4 (5)	3.3 (8)	3.5 (4)	3.5 (4)
Glutathione Precursors	3.4 (5)	3.5 (3)	3.3 (6)	3.5 (4)
Gut Microbiome Modulation	3.4 (5)	3.5 (3)	3.3 (6)	3.5 (4)
Stem cell therapy	3.3 (8)	3 (15)	3.5 (3)	3 (14)
L-deprenyl	3.3 (9)	3.4 (6)	3.2 (10)	3.3 (8)
Mitochondria (Urolithin A)	3.3 (10)	3.5 (5)	3.2 (10)	3.4 (7)
Anti-inflammatory	3.2 (11)	3.1 (12)	3.2 (12)	3.2 (10)
Plasma dilution/apheresis	3 (12)	2.9 (16)	3 (13)	2.9 (16)
17α-estradiol	3 (13)	2.5 (18)	3.3 (6)	2.7 (18)
Alpha-ketoglutarate	2.9 (14)	3.1 (12)	2.8 (14)	3.1 (11)
Fisetin	2.9 (14)	3.1 (12)	2.8 (14)	3.1 (11)
NAD+ Restoration (NMN/NR)	2.8 (16)	3.2 (9)	2.5 (21)	3 (13)
GLP-1 agonists	2.7 (17)	3.1 (11)	2.5 (21)	3 (15)
Chloroquine	2.7 (18)	2.3 (21)	2.8 (14)	2.5 (19)
Senolytics (D+Q)	2.7 (19)	2.9 (17)	2.6 (18)	2.8 (17)
Exosome therapy	2.5 (20)	2.5 (19)	2.4 (24)	2.4 (20)
Elamipretide	2.4 (21)	2.4 (20)	2.4 (23)	2.3 (21)
Young blood plasma	2.4 (22)	2.1 (22)	2.6 (17)	2.2 (22)
Telomere extension	2.4 (23)	2.1 (23)	2.5 (20)	2.1 (23)
Gene therapy	2.3 (24)	1.8 (25)	2.6 (18)	1.9 (24)
Epigenetic reprogramming	2.2 (25)	1.8 (27)	2.4 (24)	1.8 (27)
Proteostasis & Nucleolus	2.1 (26)	1.8 (26)	2.3 (26)	1.9 (25)
Chemical reprogramming	2.1 (27)	1.7 (28)	2.2 (27)	1.8 (28)
Synthetic organs	2 (28)	1.9 (24)	2 (28)	1.9 (26)
Xenotransplantation	1.8 (29)	1.5 (30)	1.9 (29)	1.6 (29)

Intervention	Baseline Weighted (Score, Rank)	Regulator-Focused Weights* (Score, Rank)	Investor-Focused Weights** (Score, Rank)	Patient-Focused Weights*** (Score, Rank)
Immunotherapy senolytics	1.8 (30)	1.6 (29)	1.8 (30)	1.5 (30)

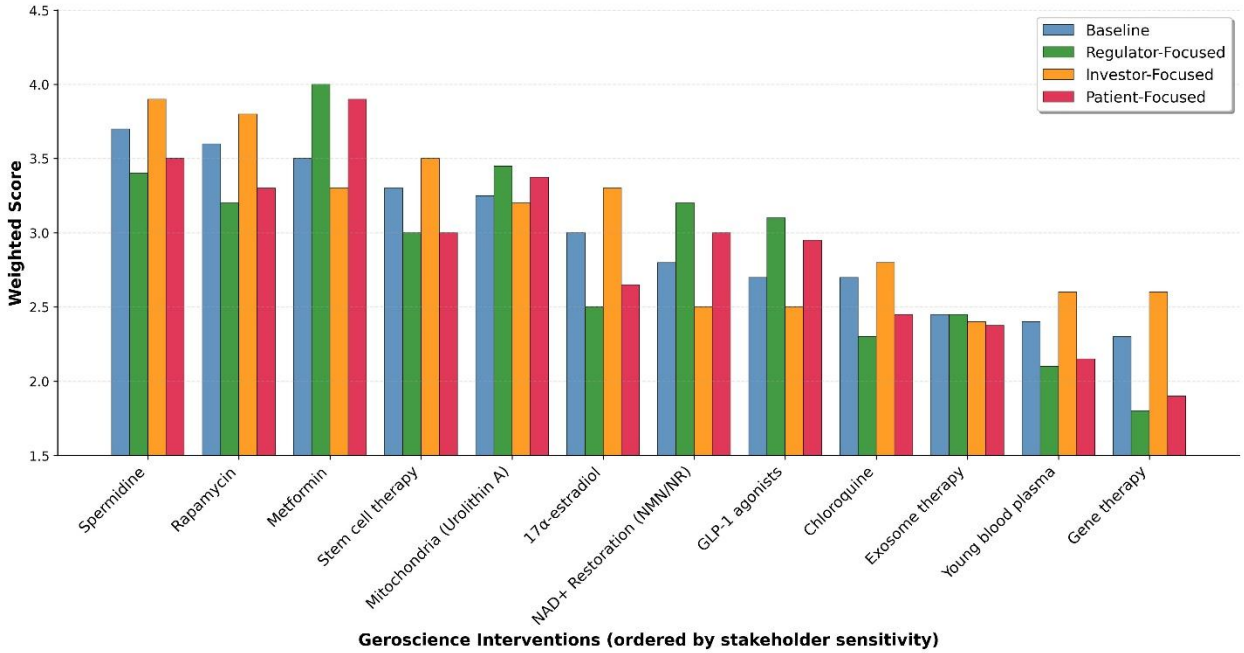
*Regulator-focused (Safety 40%, Human Evidence 30%, Preclinical Lifespan 10%, Preclinical Healthspan 10%, Accessibility 10%, Mechanism 0%)

**Investor-focused (Preclinical Lifespan 40%, Preclinical Healthspan 20%, Mechanism 10%, Human Evidence 10%, Safety 10%, Accessibility 10%)

***Patient-focused (Safety 30%, Human Evidence 25%, Preclinical Lifespan 15%, Preclinical Healthspan 15%, Accessibility 15%, Mechanism 0%)

Patient-tilted weights increase the relative standing of low-cost generics; investor-tilted weights elevate mechanistically novel candidates; regulator-tilted weights favor programs with clear IND alignment and safety record. **Fig. 7** presents tornado plots by stakeholder scenario.

Fig. 7. Impact of Stakeholder Weighting Schemes on Intervention Rankings



Tornado plot showing weighted scores for 12 key geroscience interventions under four different stakeholder weighting schemes. Interventions are ordered by sensitivity to weighting changes (largest range first). Baseline

weighting balances preclinical lifespan (30%) and healthspan (10%) with safety (20%) and human evidence (20%); Regulator-focused emphasizes safety (40%) and human evidence (30%); Investor-focused prioritizes preclinical lifespan (40%) and healthspan (20%); Patient-focused balances safety (30%), human evidence (25%), and accessibility (15%). Horizontal bars show score ranges, with markers indicating specific stakeholder scores. Notable findings: (1) Spermidine and rapamycin remain consistently high across scenarios. (2) Metformin is most favored by regulator- and patient-weightings, consistent with extensive human safety data. (3) Stem cell therapy and NAD^+ restoration (NMN/NR) show the largest investor-driven uplift, (4) whereas gene therapy and young blood plasma are most penalized under regulator/patient models, reflecting safety, cost, and limited human evidence. Error bars represent 95% credible intervals from Monte Carlo analysis (n=10,000).

The Interactive Web-Based Framework

To maximize the framework's utility and accessibility, a public, interactive web-based tool has been developed. This tool is designed to be a dynamic decision-support system and serves as a platform for disseminating the framework's findings. Its core function is to allow stakeholders—researchers, investors, clinicians, and policymakers—to explore how different strategic priorities alter the ranking landscape. The key feature is adjustable weighting: users can modify the importance of each of the six domains (e.g., an investor might increase the weight of "Preclinical Lifespan," while a regulator might maximize "Safety & Tolerability"). The rankings recalculate in real-time, visually demonstrating the inherent trade-offs in the field. For transparency, each score in the ranking table is hyperlinked to its underlying evidence and rationale. While the tool is interactive, the core analysis presented in this report represents a current snapshot based on the baseline weighting. The web platform will serve as a living tool, incorporating new evidence through regular updates and a curated community submission portal, ensuring the framework evolves with the science.

Discussion

This multi-criteria decision analysis (MCDA) provides a standardized, uncertainty-aware framework for comparing heterogeneous geroscience interventions on translationally relevant criteria[3]. The framework is intended to support real-world prioritization by making assumptions explicit and updateable. It does not replace randomized evidence; rather, it structures heterogeneous signals to identify candidates most justified for near-term clinical testing under stated values and constraints.

Across weighting scenarios, a consistent pattern emerges. Clinically familiar, relatively low-cost repurposed agents (for example, metformin, SGLT2 inhibitors and spermidine) tend to perform well because they combine established human safety with substantial human evidence[16–18]. In contrast, mechanistically transformative modalities (including stem cell-based approaches and gene or epigenetic interventions) are typically penalized primarily by translational readiness rather than biological plausibility: delivery feasibility, safety uncertainty and accessibility remain the dominant constraints[13,19,20]. This framing helps distinguish near-term clinical viability from scientific promise, recognizing that some interventions may be compelling mechanistically but still poorly positioned for immediate deployment.

Beyond producing a ranking, the domain-level profiles highlight intervention-specific knowledge gaps that define practical next steps. For leading repurposed agents, prospective trials with aging-relevant primary endpoints remain essential for establishing clinical utility and for informing plausible regulatory strategies, particularly where the aim is to address multimorbidity or functional decline rather than a single downstream disease. The TAME trial is frequently cited as a paradigmatic example of a geroscience-oriented clinical strategy for repurposed agents, although broader endpoint harmonization and replication will be required[16]. For programs that rank more favorably under investor-oriented weighting despite limited human evidence, priorities include biomarker qualification and validation, dose/exposure optimization, and longer-term safety characterization in higher-order models—steps that de-risk translation and enable interpretable early-phase studies[12].

The analysis yields actionable implications for stakeholders. For funders, results highlight the value of studies that directly address translational bottlenecks—particularly aging-endpoint trials for repurposed agents and rigorous safety, delivery and biomarker programs for advanced modalities. More broadly for regulators, a persistent system-level barrier is that aging is not widely treated as a formal therapeutic indication, which can incentivize disease-specific development strategies rather than interventions targeting shared upstream biology[21]. A potential path forward is the development of regulatory guidance for geroscience-oriented indications, including standards for the qualification and acceptance of validated surrogate biomarkers like epigenetic clocks, that could enable feasible trial durations while maintaining clinical relevance[22].

Several limitations warrant consideration. First, ordinal scoring compresses complex biological and clinical evidence into discrete categories; accordingly, uncertainty and robustness analyses are emphasized to reduce overinterpretation of point ranks and small within-tier differences. Second, evidence availability remains uneven across interventions, particularly for emerging modalities where limited human data may reflect developmental stage rather than ineffectiveness; this framework therefore prioritizes verifiable evidence while avoiding the inference that absence of evidence constitutes evidence of absence. Third, although scoring follows prespecified rubrics, some domains necessarily involve judgment in mapping heterogeneous studies to ordinal levels. The lifespan domain is largely rule-based, relying on quantitative criteria (magnitude of extension, organismal model and replication), which reduces discretionary interpretation for this component. Finally, the framework represents a snapshot of the current evidence landscape and rankings may change as new trials and mechanistic studies emerge. The accompanying web tool supports iterative updating, but outputs should be interpreted as time-indexed rather than definitive.

Despite these limitations, this work provides a systematic and reproducible approach for comparing geroscience interventions on translational criteria, offering a structured starting point for evidence-based prioritization while acknowledging the irreducible uncertainty inherent in translating aging biology into human therapeutics.

Conclusion

The field of geroscience stands at a pivotal moment, possessing a wealth of promising interventions but lacking a unified method to strategically navigate their development. The path forward is unlikely to be a one-size-fits-all solution but rather a sophisticated, multi-pronged approach. The findings suggest a fundamental limitation in current pharmacological approaches. Even the leading compounds fall short of the longevity increases that capture public imagination, suggesting that breakthrough advances may require combination therapies or entirely novel approaches beyond traditional drug development.

This framework is not a static final judgment but a dynamic tool. As new data emerge, the scores and rankings within this matrix will undoubtedly change. By transparently and systematically integrating new evidence, this framework can serve as a living document, continuously guiding the geroscience community toward the interventions with the highest probability of success. By moving from qualitative debate to quantitative comparison, we can accelerate the translation of geroscience from a promising hypothesis into a clinical reality.

Methods

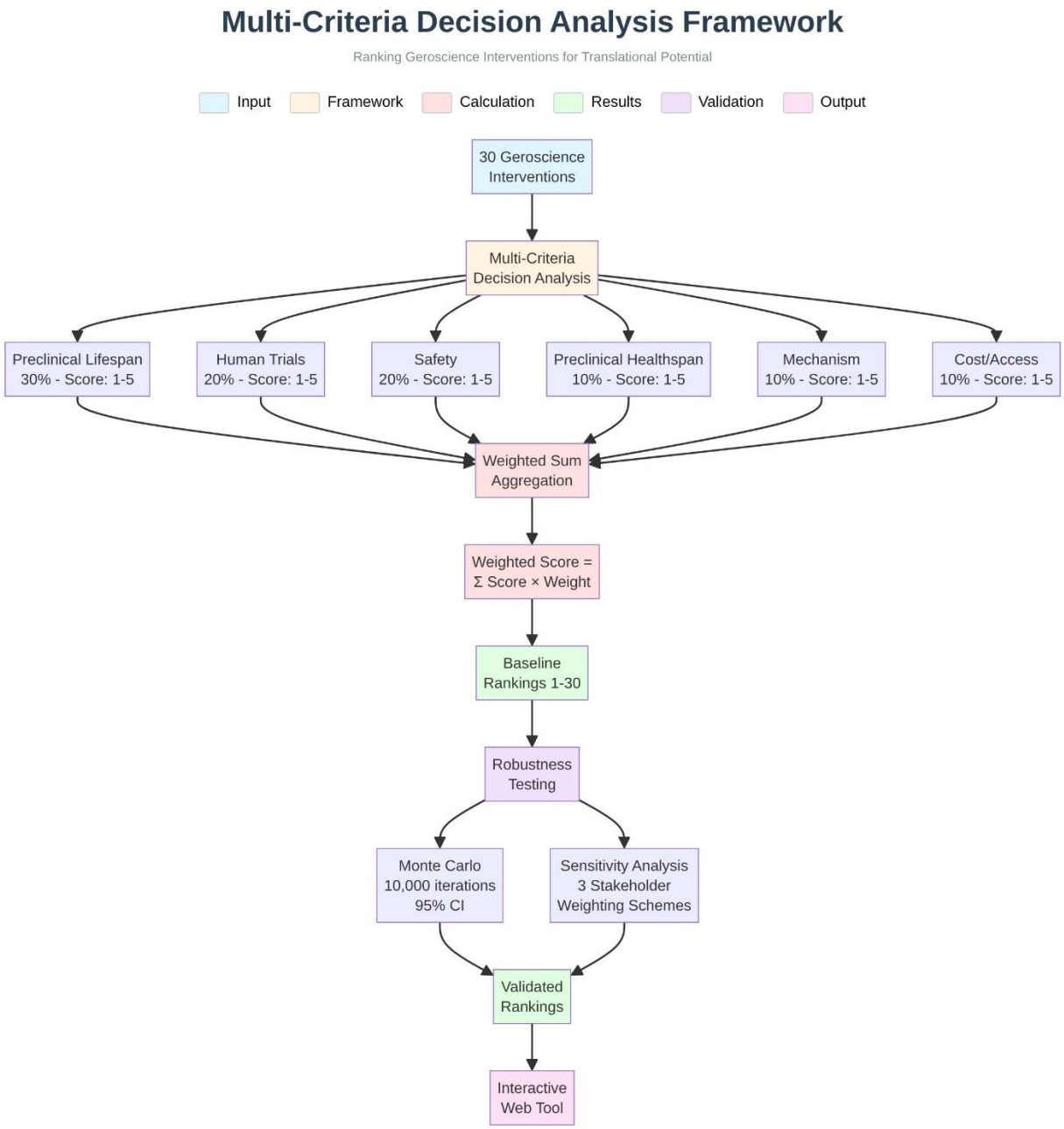
Study design and scope

A structured scoping review was conducted followed by a multi-criteria decision analysis (MCDA) to evaluate 30 geroscience interventions selected a priori to represent the breadth of the current pipeline, including repurposed pharmacologics, nutraceuticals, biologics, and frontier modalities. The objective was comparative prioritization for human translation.

Overview of the Multi-Criteria Decision Analysis (MCDA) Framework

The candidate interventions were evaluated across six domains: preclinical lifespan efficacy, preclinical healthspan, mechanistic conservation, human trial evidence, safety, and accessibility. Domain scores (1–5) were aggregated using a weighted-sum model, followed by robustness validation via Monte Carlo simulation and sensitivity analysis under alternative stakeholder weighting schemes.

Fig.8 Multi-criteria decision analysis (MCDA) framework for ranking geroscience interventions.



Evidence identification and study selection

The evidence base was assembled using a structured scoping-review approach designed to capture both preclinical longevity studies and human translational signals. Sources included PubMed-indexed primary literature and reviews, registered/posted clinical studies (e.g., ClinicalTrials.gov and equivalent registries), regulatory summaries (FDA/EMA where relevant), and landmark consortium resources in aging biology. Searches combined intervention names with aging and healthspan terms (e.g., lifespan, frailty, epigenetic age, senescence, mTOR, NAD,

mitochondrial function) and were complemented by forward/backward citation chasing. Interventions were eligible if they had at least one peer-reviewed study providing (i) lifespan and/or multi-domain healthspan outcomes in an established model organism, or (ii) human trial/real-world evidence relevant to aging-related functional or cardiometabolic endpoints. For each intervention, priority was given to the most direct evidence for the domain in question (e.g., randomized trials for Human Trial Evidence; replicated lifespan studies for Lifespan).

Data extraction, evidence mapping, and audit trail

From each eligible study, data were extracted covering study design, model organism/strain, primary endpoints, effect sizes (or best-available summary statistics), and key limitations (**Supplementary Table S2)**. Each domain score is accompanied by a written rationale and linked to one or more primary sources in an evidence map maintained alongside the scoring matrix (**Supplementary Table S1**). This audit trail enables readers to trace each score back to the specific evidence used and to update scores as new studies emerge.

Pre-specified exclusion criteria were applied to minimize bias and ensure comparability of data. Studies were excluded if (i) published prior to 1980, (ii) the primary evidence originated from a single research group with a direct commercial conflict (e.g., patent ownership or licensing arrangements) and lacked independent replication, or (iii) the intervention evidence relied on a single animal cohort. Excluded studies and their rationales are reported in **Supplementary Table S2**.

Defining the Six Critical Scoring Domains

Interventions were scored across six prespecified domains, each representing a key pillar of translational potential.

- **Preclinical Efficacy (Lifespan):** Assesses the foundational proof-of-concept for an intervention's effect on the most definitive aging outcome in model organisms: lifespan. A robust ability to extend lifespan in multiple species provides fundamental evidence that the targeted pathway is involved in the biology of aging.
- **Preclinical Efficacy (Healthspan):** Evaluates the breadth and consistency of an intervention's ability to improve health-related metrics in preclinical models, such as mobility, metabolic health, and cognitive function.
- **Mechanism Conservation:** Assesses the degree to which the intervention's molecular target and mechanism of action are conserved across evolution. A highly conserved pathway suggests that findings in mice are more likely to be translatable to human physiology.
- **Human Trial Evidence:** Evaluates the quantity, quality, and outcomes of existing human clinical trial data. It represents the direct evidence for an intervention's effects in the target species.

- **Safety & Tolerability:** Scrutinizes the known safety profile of the intervention in humans. An intervention with profound efficacy but an unacceptable safety profile has zero translational potential as a broad geroprotector.
- **Accessibility:** Evaluates the affordability and accessibility of the intervention relative to its expected benefit. This domain incorporates considerations such as mode of delivery (oral vs. injection), drug pricing and whether lower-cost generic or compounded formulations exist.

The Scoring Rubric and Weighting Methodology

Each of the six domains were scored on a 1-5 scale according to a detailed, predefined rubric (**Tables 2A and 2B**). To calculate a final translational potential score, the scores from each domain were weighted to reflect their relative importance in the drug development process.

The weighting scheme used is a formal additive weighted-sum Multi-Criteria Decision Analysis (MCDA), consistent with the good-practice recommendations of organizations like the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) (4). The standard weight for most domains is 10%, but three criteria receive a higher weight:

- **Preclinical Efficacy (Lifespan) (30%):** Lifespan extension in a mammal remains the foundational proof-of-concept for a geroprotective agent. This criterion's highest weight acknowledges that a compound must first demonstrate that it targets a fundamental aging pathway before it can be considered for human translation. The elevated weighting also acknowledges the field's nascent state and the paucity of clinical data in geroscience.
- **Human Trial Evidence (20%):** A positive human trial outcome is the single strongest predictor of clinical success, with less than 12% of interventions showing preclinical efficacy successfully translating to humans [23]. Therefore, this criterion is given a double weight to reflect its importance as the ultimate validation of an intervention's potential.
- **Safety & Tolerability (20%):** Safety failures account for a substantial percentage of late-stage drug development terminations, and this risk is particularly high for a preventive therapy intended for a broad, healthy population [24]. A spectacular effect on lifespan or healthspan is rendered irrelevant if the intervention has an unacceptable side effect profile.

The disproportionate weighting of lifespan over healthspan reflects both a pragmatic and a biological rationale. First, lifespan serves as a standardized, quantitative metric that enables direct comparison across mechanistically diverse interventions and is less sensitive to cross-study heterogeneity than multi-domain healthspan batteries with variable protocols. Biologically, significant lifespan extension in model organisms is intrinsically linked to improved healthspan; interventions that slow core aging biology often shift multiple functional domains in parallel. Thus, lifespan operates as an integrative readout that inherently captures sustained

functional benefit, while its quantitative and comparative nature provides a necessary anchor for cross-intervention evaluation.

This weighting reflects a pragmatic, development-focused perspective: an intervention with a perfect safety record and promising human data may have higher immediate translational potential than an intervention with spectacular preclinical results but significant safety concerns (e.g., Epigenetic Reprogramming), even if the latter holds greater long-term promise. This balanced approach prevents overreliance on either preclinical promise or limited human evidence alone.

Tables 2A and 2B present the detailed scoring rubric that operationalizes these six evaluation domains. To keep the rubric readable, domains are split into (A) evidence and biological plausibility (preclinical efficacy, healthspan, human trial evidence) and (B) feasibility and implementation (mechanism conservation, safety, and accessibility). Each domain is scored on a standardized 1-5 scale, with specific, predefined criteria that ensure consistency and reproducibility in the evaluation process. A score of 5 represents the highest level of evidence or most favorable profile within that domain (e.g., FDA-approved drugs with extensive safety data, or robust lifespan extension across multiple mammalian species), while a score of 1 indicates minimal evidence, unfavorable characteristics, or substantial concerns. This granular rubric allows for nuanced assessment of each intervention's strengths and limitations across the translational pipeline, from preclinical proof-of-concept through human safety and accessibility considerations.

Table 2A. Translational potential scoring rubric (Domains 1-3; 1-5 scale)

Score	Preclinical Efficacy (Lifespan)	Preclinical Efficacy (Healthspan)	Human Trial Evidence
5	Large lifespan extension (≥15%) in mammalian models with reproducible findings across multiple studies, OR exceptional extension (≥30%) in invertebrate models with reproducible findings.	Broad, robust improvements across multiple (≥3) healthspan domains demonstrated in multiple mammalian species.	High-certainty evidence from multiple, large, low-risk-of-bias RCTs showing a clinically meaningful benefit on an age-related indication; FDA-approved for a relevant chronic condition.
4	Significant (≥10%) median lifespan extension in multiple independent studies in mammalian species, OR Interventions showing high magnitude (≥20%) in single mammalian studies.	Broad, robust improvements across two healthspan domains demonstrated in at least one mammalian species.	Moderate-certainty evidence from at least one large, low-risk-of-bias RCT or multiple smaller RCTs showing a clinically meaningful benefit.

3	Moderate to robust lifespan extension (5-10%) in mammalian models with strong reproducibility, OR substantial extension (10-20%) in mammalian models from single landmark studies.	Consistent improvements demonstrated in one major healthspan domain in mammalian models.	Low-certainty evidence from small or moderate-risk-of-bias RCTs suggesting a benefit on a primary aging/healthspan endpoint.
2	Modest lifespan extension (0-5%) in mammalian models with reproducibility, OR moderate effects (5-10%) in single mammalian studies, OR substantial effects in invertebrate models (5-10%), OR effects observed only in transgenic/progeroid models.	Positive effects on multiple healthspan metrics reported, but results are not yet replicated, or broad improvements only in invertebrate models.	Very low-certainty evidence from high-risk-of-bias RCTs or observational studies for surrogate endpoints.
1	Inconsistent, or evidence suggests negative/null effect.	Inconsistent or modest healthspan improvements in mammals, or evidence suggests negative/null effect. No direct evidence.	No human trials completed or planned. Only preclinical or observational data exists.

Table 2B. Translational potential scoring rubric (Domains 4-6; 1-5 scale)

Score	Mechanism Conservation	Safety & Tolerability	Accessibility
5	Core pathway components and function are ubiquitously conserved across all eukaryotes, with clear orthologs and functional equivalence.	FDA-approved for a chronic condition with decades of widespread use (>10M patients). Side effect profile is well-characterized, predominantly mild, and non-life-threatening. No black box warnings.	Very low cost (<$1/day). Widely available over-the-counter or as a generic drug. Simple oral administration.
4	Pathway is highly conserved across all vertebrates with clear functional equivalence in key regulatory nodes.	Long-term use for a chronic condition in large populations (>1M patients). Side effects are well-known,	Low cost (<$5/day). Available as a generic prescription drug. Simple oral administration.

		mostly mild-to-moderate, and reversible.	
3	Pathway is conserved across metazoans, though some components may have diverged.	Approved for a chronic condition with a well-understood safety profile. Common side effects are moderate but manageable. Low incidence of serious adverse events (SAEs).	Moderate cost (<$20/day). Available as a branded prescription drug with good insurance coverage.
2	Pathway is conserved in mammals, but its role in invertebrates is unclear or different.	Used clinically for an acute or intermittent indication. Side effects can be significant but are predictable and manageable in a clinical setting.	High cost (<$100,000/year). Requires specialist prescription and monitoring. May require injection or infusion.
1	Pathway is specific to a particular lineage or its role in aging is highly speculative.	Significant safety concerns based on mechanism of action. High potential for severe, irreversible side effects. Known toxicity in humans.	Prohibitively high cost (>$100,000/year). Complex administration (e.g., gene therapy). Highly restricted access.

Lifespan Scoring Methodology

To standardize preclinical lifespan evidence, a composite scoring system was developed to account for three critical dimensions: the magnitude of lifespan extension, the translational relevance of the model organism, and the reproducibility of findings. A hierarchical approach prioritizes higher-quality evidence: when an Interventions Testing Program (ITP) study is available, the Lifespan Change Score is derived exclusively from the ITP study; when both rodent and invertebrate studies are available, the score is derived exclusively from rodent studies. For each intervention-model pair, the proportional change in average lifespan relative to matched controls was extracted across eligible studies, and an average effect size was computed when multiple studies were available.

When several distinct interventions fell under the same broader category, (e.g., Stem Cell Therapy with five different approaches: bone-marrow transplant, FOXO3, CASIN-treated aged HSCs, young-donor HSCs, and Hypothalamic NSC), the highest Net Lifespan Score among these interventions was selected for the category's ranking. This approach ensured that each category is represented by its strongest preclinical evidence.

The Net Lifespan Score was calculated as:

Net Lifespan Score = Lifespan Change Score × Model Organism Score × Reproduced Study Score

- **Lifespan Change Score:**

  This component evaluates the magnitude of lifespan extension observed in preclinical studies. For studies in wild-type healthy organisms, scores are assigned based on the percentage increase in median or mean lifespan: 9 points (≥35%), 8 points (≥30%), 7 points (≥25%), 6 points (≥20%), 5 points (≥15%), 4 points (≥10%), 3 points (≥5%), 2 points (>0%), and 1 point (nil or <0%).

- **Model Organism Score:**

  This component adjusts for the translational relevance of the model organism: 2.0 points for primates (e.g., Rhesus macaques), 1.0 point for mammalian wild-type models (e.g., Mus musculus, Rattus norvegicus), 0.5 points for transgenic, progeroid or diseased mammalian models, and 0.5 points for other animal model organisms (e.g., *C. elegans*, Drosophila, fish).

- **Reproduced Study Score:**

  This component accounts for the reproducibility and robustness of findings: 1.0 point for multiple independent studies (≥2 studies) demonstrating lifespan extension, and 0.5 points for a single published study. Studies from the NIA Interventions Testing Program (ITP) also receive a score of 1.0, as the ITP design inherently incorporates reproducibility by testing interventions simultaneously across three independent sites[25].

The resulting Net Lifespan Score provided a normalized metric that allowed direct comparison across interventions with different model systems and study designs. Higher scores indicated interventions with robust preclinical evidence from translationally relevant models.

Scoring procedure and internal consistency checks

Domain scores were assigned using the predefined 1–5 rubric (**Tables 2A-2B**), with ties resolved by applying the rubric's explicit hierarchy. To reduce arbitrariness, a second-pass consistency audit was performed across the full matrix: (i) each score was required to have at least one mapped primary source; (ii) adjacent scores (e.g., 3 vs 4) were cross-checked against the rubric language; and (iii) outliers were stress-tested via uncertainty and sensitivity analyses (Monte-Carlo sampling and weight perturbation). When an intervention had data across multiple phylogenetic levels, only the highest-order model organism available was considered in the lifespan score to maximize translational relevance.

Uncertainty & sensitivity analyses

To quantify the uncertainty in the final scores, a Monte-Carlo simulation was conducted where each domain score was sampled 10,000 times from a uniform distribution of ±0.5 around its

assigned value (bounded by 1 and 5) to generate 95% credible intervals. Robustness to stakeholder preference shifts was evaluated through weight-perturbation analyses (±5% per domain), examining changes in rank ordering under alternative but reasonable weighting schemes.

Transparency and reproducibility.

All domain definitions, scoring rubrics, weights, and analytical code were predefined and are intended for open dissemination to allow independent replication, reweighting, and updating as new evidence emerges.

Competing interests.

The author declares no competing financial or non-financial interests. No payments or services were received from any third party that could be perceived to influence the submitted work.

Funding.

This study did not receive any external funding.

Ethics statement.

This study is a computational analysis of previously published preclinical and clinical data. No new human or animal subjects were involved, and no institutional review board (IRB) or ethics committee approval was required.

Data and Code Availability.

Data underlying the findings of this study are provided in the manuscript and Supplementary Materials. Additional curated input files and analytical materials are available from the corresponding author upon reasonable request. An interactive web tool accompanying this study is accessible at https://translational-geroscience-mcda-ranking.streamlit.app/

Supplementary Materials

- Table 1: Lifespan score decomposition and net lifespan scores by intervention and model organism
- Table 2A: Translational potential scoring rubric (Domains 1-3; 1-5 scale)
- Table 2B: Translational potential scoring rubric (Domains 4-6; 1-5 scale)

- Table 3: Scoring and Ranking Matrix for Geroscience Interventions (1-5 Scale)
- Table 4: MCDA Ranking Results
- Table 5. Extended sensitivity analysis results
- Supplementary Table S1: Evidence map for all 30 interventions across all six domains
- Supplementary Table S2: Detailed lifespan data
- Fig. 1: Comparative Mortality Risk Factors
- Fig. 2: US Mortality Rates by Age Group (2023, per 100,000)
- Fig. 3. Age-Specific Mortality Rates by Leading Causes of Death
- Fig. 4: Translational Readiness vs Potential Impact
- Fig. 5: Rank Stability Analysis: Monte Carlo Simulation of Weighted Scores
- Fig. 6. Domain-wise Contribution Profiles: Weighted Score Decomposition
- Fig. 7. Impact of Stakeholder Weighting Schemes on Intervention Rankings
- Fig. 8: Multi-criteria decision analysis (MCDA) framework
- Interactive web tool

References

1. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The Hallmarks of Aging. *Cell*. 2013;153(6):1194-1217. doi:10.1016/j.cell.2013.05.039

2. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. Hallmarks of aging: An expanding universe. *Cell*. 2023;186(2):243-278. doi:10.1016/j.cell.2022.11.001

3. Thokala P, Devlin N, Marsh K, et al. Multiple Criteria Decision Analysis for Health Care Decision Making–Emerging Good Practices: Report 2 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health*. 2016;19(1):1-13. doi:10.1016/j.jval.2015.12.003

4. Ahmad FB, Cisewski JA, Anderson RN. Mortality in the United States — Provisional Data, 2023. *MMWR Morb Mortal Wkly Rep*. 2024;73(31):677-681. doi:10.15585/mmwr.mm7331a1

5. Jacobs Jr DR, Adachi H, Mulder I, et al. Cigarette smoking and mortality risk: twenty-five-year follow-up of the Seven Countries Study. *Arch Intern Med*. 1999;159(7):733-740. doi:10.1001/archinte.159.7.733

6. Di Angelantonio E, Bhupathiraju SN, Wormser D, et al. Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. *The Lancet*. 2016;388(10046):776-786. doi:10.1016/S0140-6736(16)30175-1

7. Martinez-Gomez D, Luo M, Huang Y, et al. Physical Activity and All-Cause Mortality by Age in 4 Multinational Megacohorts. *JAMA Netw Open*. 2024;7(11):e2446802. doi:10.1001/jamanetworkopen.2024.46802

8. Pan H, Hibino M, Kobeissi E, Aune D. Blood pressure, hypertension and the risk of sudden cardiac death: a systematic review and meta-analysis of cohort studies. *Eur J Epidemiol*. 2020;35(5):443-454. doi:10.1007/s10654-019-00593-4

9. Oh J, Hevia-Ramos G, Ha E, Hong YC, Kim H, Lim YH. A Systematic Review and Meta-Analysis on Long-Term Exposure to Particulate Matter and All-Cause and Cause-Specific Mortality in the Asia-Pacific States. *J Korean Med Sci*. 2025;40(27):e156. doi:10.3346/jkms.2025.40.e156

10. WHO. *Noncommunicable Diseases*. 2025. https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases

11. Scott EC, Baines AC, Gong Y, et al. Trends in the approval of cancer therapies by the FDA in the twenty-first century. *Nat Rev Drug Discov*. 2023;22(8):625-640. doi:10.1038/s41573-023-00723-4

12. Justice JN, Ferrucci L, Newman AB, et al. A framework for selection of blood-based biomarkers for geroscience-guided clinical trials: report from the TAME Biomarkers Workgroup. *GeroScience*. 2018;40(5-6):419-436. doi:10.1007/s11357-018-0042-y

13. Lei J, Xin Z, Liu N, et al. Senescence-resistant human mesenchymal progenitor cells counter aging in primates. *Cell*. 2025;188(18):5039-5061.e35. doi:10.1016/j.cell.2025.05.021

14. Strong R, Miller RA, Antebi A, et al. Longer lifespan in male mice treated with a weakly estrogenic agonist, an antioxidant, an α-glucosidase inhibitor or a Nrf2-inducer. *Aging Cell*. 2016;15(5):872-884. doi:10.1111/acel.12496

15. Stout MB, Vaughan KL, Isola JVV, et al. Assessing tolerability and physiological responses to 17α-estradiol administration in male rhesus macaques. *GeroScience*. 2023;45(4):2337-2349. doi:10.1007/s11357-023-00767-9

16. Barzilai N, Crandall JP, Kritchevsky SB, Espeland MA. Metformin as a Tool to Target Aging. *Cell Metab*. 2016;23(6):1060-1065. doi:10.1016/j.cmet.2016.05.011

17. Miller RA, Harrison DE, Cortopassi GA, et al. Lifespan effects in male UM-HET3 mice treated with sodium thiosulfate, 16-hydroxyestriol, and late-start canagliflozin. *GeroScience*. 2024;46(5):4657-4670. doi:10.1007/s11357-024-01176-2

18. Eisenberg T, Abdellatif M, Schroeder S, et al. Cardioprotection and lifespan extension by the natural polyamine spermidine. *Nat Med*. 2016;22(12):1428-1438. doi:10.1038/nm.4222

19. Ocampo A, Reddy P, Martinez-Redondo P, et al. In Vivo Amelioration of Age-Associated Hallmarks by Partial Reprogramming. *Cell*. 2016;167(7):1719-1733.e12. doi:10.1016/j.cell.2016.11.052

20. Zinman B, Wanner C, Lachin JM, et al. Empagliflozin, Cardiovascular Outcomes, and Mortality in Type 2 Diabetes. *N Engl J Med*. 2015;373(22):2117-2128. doi:10.1056/NEJMoa1504720

21. Kennedy BK, Berger SL, Brunet A, et al. Geroscience: Linking Aging to Chronic Disease. *Cell*. 2014;159(4):709-713. doi:10.1016/j.cell.2014.10.039

22. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013;14(10):3156. doi:10.1186/gb-2013-14-10-r115

23. Hay M, Thomas DW, Craighead JL, Economides C, Rosenthal J. Clinical development success rates for investigational drugs. *Nat Biotechnol*. 2014;32(1):40-51. doi:10.1038/nbt.2786

24. Harrison RK. Phase II and phase III failures: 2013-2015. *Nat Rev Drug Discov*. 2016;15(12):817-818. doi:10.1038/nrd.2016.184

25. Nadon NL, Strong R, Miller RA, Harrison DE. NIA Interventions Testing Program: Investigating Putative Aging Intervention Agents in a Genetically Heterogeneous Mouse Model. *EBioMedicine*. 2017;21:3-4. doi:10.1016/j.ebiom.2016.11.038

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTableS1Evidencemapforall30interventionsacrossallsixdomains.xlsx
- SupplementaryTableS2Detailedlifespandata.xlsx