Online content

Methods, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/xxxxxxxxxxx.

METHODS

Trial participants and proteome characterization

A total of 17,604, 529, 616, 3,533 and 9,650 participants were enrolled in the SELECT (NCT03574597)[1], STEP-HFpEF (heart failure with preserved ejection fraction; NCT04788511)[2], STEP-HFpEF DM (diabetes mellitus; NCT04916470)[3], FLOW (NCT03819153)[4] and SOUL (NCT03914326)[5] trials, respectively. A subset of these participants consented to inclusion in proximity extension assay-based proteomic analyses using the Olink Explore HT panel (Olink Proteomics, Uppsala, Sweden), which measures the relative abundance of 5,416 unique proteins using proximity extension assay technology[6].

After quality-control filtering, 1,641, 416, 469, 2,271 and 5,255 participants remained from SELECT, STEP-HFpEF, STEP-HFpEF DM, FLOW and SOUL, respectively. In SELECT, participants were selected for Olink Explore HT proteomic measurement based on a nested case-control design, in which individuals who experienced a major adverse cardiovascular event (MACE) or heart failure (HF) event during the trial were matched with an equal number of event-free participants. Analyses were conducted at week 20 for SELECT; weeks 20 and 52 for STEP-HFpEF and STEP-HFpEF DM; weeks 12, 52 and 156 for FLOW; and weeks 13 and 104 for SOUL.

Baseline tables and cross-sectional analyses included participants with available Olink proteomic measurements at baseline, whereas longitudinal analyses were restricted to participants with the required baseline and post-baseline proteomic measurements.

Application of organ-specific mortality clocks

Goeminne, L. J. E. *et al*[7] developed plasma protein-based prediction equations in UK Biobank to estimate chronological age and all-cause mortality risk/time to death[7]. In the present analysis, we used the mortality-trained equations and applied them to trial proteomics data without re-training or re-estimating model coefficients. In this manuscript, “clock” refers to the fixed published prediction model, whereas “biological-age score” refers to the participant-level output obtained after applying the model and converting the predicted mortality hazard to years.

For each trial participant and proteomic timepoint, we computed clock-derived scores for 18 organ systems using the published Goeminne, L. J. E. *et al*[7] mortality-clock equations: adipose,

adrenal, artery, brain, esophagus, heart, immune, intestine, kidney, liver, lung, muscle, pancreas, pituitary, salivary, skin, stomach and thyroid. For each organ-specific model, the participant's Olink Normalized Protein eXpression (NPX) value for each available protein was multiplied by the corresponding published coefficient and summed to generate a predicted relative log-mortality hazard score.

A protein could contribute to more than one organ-specific equation. When this occurred, the same participant-level NPX value was used separately in each relevant organ equation and multiplied by the coefficient specific to that equation. Thus, shared proteins contributed independently to each organ-specific score, but scores from different organs were not added together and proteins were not counted more than once within a given clock score.

Of the 2,916 proteins included in the published mortality-clock equations, excluding sex-specific clocks, 100 proteins were unavailable after quality-control filtering in SELECT, FLOW and SOUL, and 91 were unavailable in STEP-HFpEF and STEP-HFpEF DM. These proteins were omitted from downstream calculations.

In addition to the 18 organ-specific scores, we applied three composite mortality-clock equations defined by Goeminne, L. J. E. *et al*[7]: the multiorgan score, based on the combined protein set from the organ-specific models; the conventional score, based on the all-protein mortality model; and the organismal score, based on the conventional model after excluding proteins included in the multiorgan model. These composite scores were taken from Goeminne, L. J. E. *et al*[7] and were not newly derived in the present study. The conventional mortality clock was trained against all-cause mortality/time to death rather than an organ-specific endpoint. Predicted relative log-mortality hazard scores were then converted to biological age in years using the Gompertz-based transformation described by Goeminne, L. J. E. *et al*[7].

Multimodal computational analysis

All computational analyses were performed in R v4.4.1 (https://www.R-project.org/) unless otherwise stated.

Chronological-age calculation

Chronological age at the time of proteomics sample collection was calculated from participant birth dates and sample collection dates. Birth date information was available as complete dates (YYYY-MM-DD) for all participants in SOUL and FLOW and for most participants in SELECT, year-month (YYYY-MM) for most participants in STEP-HFpEF and STEP-HFpEF DM, and year-only (YYYY) for a subset of participants in SELECT. To standardize age calculations, birth dates with

year-only precision were assigned to the middle of the respective birth year, whereas birth dates with year-month precision were assigned to the 15th day of the specified month. Age was computed as the precise interval between the standardized birth date and the sample collection date, expressed in decimal years using the lubridate R package. This approach yielded age estimates with a maximum uncertainty of ±6 months for participants with year-only birth date records and ±15 days for participants with year-month birth date records.

Application of organ-specific mortality models to clinical trial proteomics data

Organ-specific mortality-clock scores were computed for each participant at each independent timepoint where proteomics was acquired, separately across all five trials. For each organ-specific, organismal, multiorgan and conventional mortality-based model, predicted relative log-mortality hazard scores were calculated by multiplying the published model coefficients by the corresponding Olink NPX values at the given timepoint and summing across all contributing proteins:

$$\hat{y}_i = \sum_p \beta_p \cdot x_{ip}$$

Where $\hat{y}_i$ is the predicted relative log(mortality hazard) for participant i, β_p is the published coefficient for protein p and x_{ip} is the baseline NPX value for protein p in participant i.

To facilitate interpretation, predicted relative log-mortality hazard scores were converted to mortality age in years using the Gompertz-based transformation described by Goeminne, L. J. E. *et al*[7]. Briefly, each participant's predicted score was mapped to the chronological age at which an average individual in the reference population would have the same predicted mortality hazard. Specifically, for a given participant i with predicted relative log-mortality hazard $\hat{y}_i$, mortality age was calculated as:

$$\text{Biological age}_i = \frac{\hat{y}_i - \alpha\left(\alpha + \beta \cdot \bar{a}\right)}{\beta}$$

where α and β are the intercept and slope, respectively, from a linear regression of predicted relative log(mortality hazard) on chronological age in the reference population. Biological age was therefore expressed in years, enabling direct interpretation of longitudinal treatment effects.

Age deviation was calculated as the residual from a linear regression of biological age on chronological age for each clock. Biological age was used for treatment-effect analyses, whereas age deviations were used for baseline association analyses and treatment-interaction analyses.

Application of UK Biobank proteomic frailty score (PFS) to clinical trial data

Frailty was defined according to Jia, X. *et al*[8] as a UK Biobank frailty index based on 49 self-reported items across sensory, cranial, mental well-being, infirmity, cardiometabolic, respiratory, musculoskeletal, immunologic, cancer, pain and gastrointestinal conditions, with the PFS representing an Olink NPX-based least absolute shrinkage and selection operator (LASSO)-derived molecular proxy of that index. PFS were obtained by averaging the LASSO regression coefficients across all ten models published by Jia, X. *et al*[8] to create a single ensemble model, and applying the averaged coefficients to the corresponding Olink NPX values: $\hat{y}_{ij} = \Sigma_p \beta_{pj} \cdot x_{ip}$, where $\hat{y}_{ij}$ is the predicted frailty score for participant i from model j, β_{pj} is the mean coefficient for protein p in LASSO model j and x_{ip} is the NPX value for protein p in participant i. To remove age-related confounding, PFS values were age-adjusted by extracting residuals from a linear regression of predicted frailty scores on chronological age. Treatment effects on PFS change from baseline were estimated using linear regression models fitted separately for each trial-timepoint combination: $CHG_i = \beta_0 + \beta_1(TRT01P_i) + \beta_2(BASE_i) + \varepsilon_i$, where CHG_i represents the change from baseline in PFS for participant i, $TRT01P_i$ indicates treatment assignment (semaglutide versus placebo), $BASE_i$ is the baseline age-adjusted PFS and ε_i is the residual error term. The coefficient β_1 represents the treatment effect, with negative values indicating a reduction in the PFS. As a single frailty score was tested per trial-timepoint combination, no multiple-testing correction was required. Treatment effects (β_1 coefficients) with 95% confidence intervals were visualized using forest plots across all trial-timepoint combinations.

Application of UK Biobank proteomic disease-risk models to clinical trial data

Disease-specific proteomic risk scores were computed for each participant at each independent timepoint where proteomics were acquired, separately across all five trials. For each of the 15 selected disease conditions, predicted disease-risk scores were calculated by multiplying the published linear elastic net regression coefficients[9] by the corresponding Olink NPX values and summing across all contributing proteins: $\hat{y}_i = \Sigma_p \beta_p \cdot x_{ip}$, where $\hat{y}_i$ is the predicted disease-risk score for participant i, β_p is the published coefficient for protein p from the UK Biobank model and x_{ip} is the NPX value for protein p in participant i. From the original 27 validated disease models developed using 2,919 proteins from 53,030 UK Biobank participants, diseases with ≥5 non-zero protein coefficients were selected to ensure sufficient model complexity and stability, yielding 15 diseases. Selected diseases included cardiometabolic conditions (type 2 diabetes, coronary heart disease, HF, atrial fibrillation), respiratory diseases (chronic obstructive pulmonary disease, asthma, primary pulmonary hypertension), neuropsychiatric conditions (dementia,

depression), thrombotic diseases (pulmonary embolism, ischemic stroke), autoimmune conditions (rheumatoid arthritis, lupus) and metabolically linked cancers (breast, lung). Treatment effects on change from baseline in disease risk were estimated using linear regression models fitted separately for each disease-trial-timepoint combination: $CHG_i = \beta_0 + \beta_1(TRT01P_i) + \beta_2(BASE_i) + \varepsilon_i$, where CHG_i represents the change from baseline in predicted disease-risk score for participant i, $TRT01P_i$ indicates treatment assignment (semaglutide versus placebo), $BASE_i$ is the baseline disease-risk score and ε_i is the residual error term. The coefficient β_1 represents the treatment effect, with negative values indicating beneficial risk reduction and positive values indicating increased disease risk. Treatment effects were estimated separately at each timepoint to provide trial- and visit-specific estimates consistent with the organ-age and PFS analyses. Treatment effects (β_1 coefficients) with 95% confidence intervals were visualized using forest plots across all trial-timepoint combinations. Statistical significance was assessed using the Holm–Bonferroni procedure for multiple comparisons across the 15 diseases and all trial-timepoint combinations, with family-wise error rate controlled at $\alpha = 0.05$.

Application of biological-age models to SomaScan data

SomaScan data from STEP 1 and STEP 2 were $\log_{10}$-transformed, and measurements for each aptamer were standardized using the mean and standard deviation of baseline samples. To apply the protein biological-age clock models to these data, we developed a cross-platform scaling procedure using matched samples from the SELECT trial. Specifically, among the 1,676 SELECT participants with both SomaScan and Olink measurements at baseline, we computed the Pearson correlation between the two platforms for each protein and retained proteins with $|r| > 0.2$. For each retained protein, a scaling multiplier was defined as the product of the absolute correlation and the standard deviation of the Olink measurements across the 1,676 matched participants (multiplier = $|r| \times SD\_Olink$). SomaScan protein values in STEP 1 and STEP 2 were then rescaled by multiplying each protein measurement by its corresponding protein-level multiplier. The rescaled SomaScan values were then passed directly to the organ aging coefficient models (gen1 and gen2) to derive predicted organ-specific biological ages.

Association of baseline biological age with medications

Associations between concomitant medications ongoing at randomization and predicted biological ages for each organ-specific clock were tested by defining medication use as a binary variable (use versus non-use at randomization). Medications included were selected based on the standardized categorization and mapping process detailed in Goeminne, L. J. E. *et al*[7] and only those medications with a usage frequency greater than 2.5% in each cohort were

included, to ensure robust statistical power. Linear regression was performed for each medication-clock pair, adjusting for chronological age, sex, their interaction, smoking status and race. Multiple-testing correction was applied using Holm-Bonferroni multiple test correction.

Association of baseline biological age with biomarkers

Associations between baseline clinical biomarkers and predicted biological ages for each organ-specific clock were tested using linear regression for each biomarker–clock pair, adjusting for chronological age, sex, their interaction, smoking status and race, except for the smoking model where smoking was not included as a covariate. To facilitate clinical interpretation across trial populations, continuous biomarkers were modeled as binary variables using clinically established thresholds; continuous-variable sensitivity analyses were not performed. Chronic kidney disease was defined as an estimated glomerular filtration rate below 60 ml/min/1.73 m^2, microalbuminuria as a urine albumin-to-creatinine ratio exceeding 30 mg/g, prediabetes as glycated hemoglobin (HbA_{1c}) level between 5.7% and 6.4%, low levels of high-density lipoprotein cholesterol (below 40 mg/dl), elevated total cholesterol as values exceeding 200 mg/dl, elevated systolic blood pressure as values exceeding 120 mmHg and elevated diastolic blood pressure as values exceeding 80 mmHg. HF with reduced ejection fraction and HFpEF were defined based on trial-adjudicated diagnoses. Smoking was defined as reported current smoking status at randomization. Multiple-testing correction was applied using the Holm–Bonferroni procedure.

Association of baseline biological age with self-reported questionnaire

For each trial, patient-reported outcome questionnaires were assessed using instruments available at baseline: the EuroQol 5-dimension 5-level (EQ-5D-5L)[10] supplemented by the Weight-Related Symptom Scale (WRSS)[11] and the Kansas City Cardiomyopathy Questionnaire (KCCQ)[12] in STEP-HFpEF and STEP-HFpEF DM. The EQ-5D-5L captures general health-related quality of life across five domains (mobility, self-care, usual activities, pain or discomfort, and anxiety or depression), whereas the WRSS assesses the presence and bother associated with weight-related symptoms and the KCCQ captures HF-specific symptom burden and functional limitations. Each measure was harmonized by dichotomizing participants into the worst-performing quartile (worst 25%) versus all others, using rank-based assignment to handle ties. Because the instruments and their score ranges differed across trials, the within-trial worst-quartile indicator was used to provide a common definition of poorer health status; continuous-score sensitivity analyses were not performed. Quality-of-life measures reported by fewer than 10% of participants within a trial were excluded. Associations between the binary worst-quartile indicator and predicted organ biological age were tested using linear regression for each

questionnaire-clock pair, adjusting for chronological age, sex, their interaction, smoking status and race. Multiple-testing correction was applied using the Holm–Bonferroni procedure.

Prognostication of baseline biological age to clinical endpoints

Cox proportional hazards models were fitted in participants receiving placebo from SELECT for 22 non-COVID–related time-to-event endpoints and from SOUL for 30 time-to-event endpoints. Baseline organ-specific predictive biological ages were included as predictors, with chronological age included as a linear covariate; alternative non-linear specifications were not evaluated. These analyses were intended as age-adjusted prognostic validation rather than assessment of incremental predictive value beyond established baseline clinical risk factors; endpoint-specific baseline disease covariates were therefore not included. Multiple-testing correction was applied using the Holm–Bonferroni procedure.

Estimating the effect of semaglutide on biological age

Linear regression was used to assess the effect of semaglutide versus placebo on changes from baseline in organ-specific biological age (in years). The model included the treatment indicator and the baseline biological-age value as a covariate (CHG ~ TRT01P + BASE). Active treatment arms were semaglutide 2.4 mg for SELECT, STEP-HFpEF and STEP-HFpEF DM; semaglutide 1.0 mg for FLOW; and oral semaglutide 14 mg for SOUL. To account for multiple testing across organ-specific biological-age models, *P* values were adjusted for multiple testing using the Holm–Bonferroni procedure ($\alpha = 0.05$). All analyses were conducted separately by trial.

As a sensitivity analysis to account for the case-enriched design of SELECT, treatment effects were re-estimated with inverse probability-style sampling weights. All participants with a MACE or HF event were included in the Olink subsample ($w = 1$), whereas non-event participants were sampled separately within each treatment arm (~402 per arm) and assigned weights equal to the number of non-event participants in that arm divided by the number sampled (placebo: $w \approx 12.6$; semaglutide: $w \approx 12.6$). Treatment effects were then estimated using weighted linear regression via svyglm from the R survey package[13], which incorporates these weights to produce estimates.

To assess whether treatment-associated changes in biological-age scores were statistically explained by concurrent changes in body weight, HbA_{1c} or high-sensitivity C-reactive protein (hsCRP), we performed formal causal mediation analysis using the mediation package in R[14]. This analysis separates the total semaglutide effect into two parts: an indirect effect, representing the portion statistically explained by change in the mediator, and a remaining estimated effect conditional on the measured mediator. Mediation was conducted only for organ-specific biological-age models showing a statistically significant treatment effect (Holm-adjusted *P* <

0.05). We examined whether the treatment effect on changes in biological age was mediated by changes in body weight, HbA_{1c} and hsCRP, where available. PFS and clinically measured frailty were not evaluated as candidate mediators; consequently, these analyses did not assess whether frailty-associated weight loss contributed to changes in the adipose biological-age score. For each organ clock and mediator, two models were fit: a mediator model predicting the change in the mediator from treatment assignment and the corresponding baseline values (e.g. body weight change ~ treatment + baseline weight), and an outcome model predicting the change in biological age from treatment assignment, the change in the mediator, baseline mediator value and baseline biological age. For SELECT, models were fit using svyglm from the survey package to incorporate inverse probability-style sampling weights accounting for the case-enriched design; for all other trials, standard linear regression was used. The average causal mediation effect and average direct effect were estimated using the counterfactual framework, with confidence intervals obtained via nonparametric bootstrap (100 simulations)[15].

Mediation analysis of the contribution of organ-specific biological aging to the effect of semaglutide on outcomes

To investigate whether the effect of semaglutide on mortality outcomes was mediated through changes in organ-specific biological aging, we performed counterfactual mediation analyses for time-to-event outcomes. Analyses were conducted in SELECT (week 20) and SOUL (week 13) participants with complete measurements. For each organ-specific clock, changes from baseline in mortality-clock scores served as the mediator. The total treatment effect on time-to-event outcomes was decomposed into a mediated (clock-dependent) and non-mediated (clock-independent) component. The outcome was modeled using a Cox proportional hazards model (survival package), and the mediator was modeled using linear regression. Counterfactual survival probabilities were estimated using riskRegression::predictCox, and the mediated proportion was computed as the natural indirect effect divided by the total effect, averaged across individuals. To address confounding introduced by the post-randomization mediator measurement, a pre-mediator Cox model was fitted to estimate survival probabilities at the mediator timepoint, which served as inverse probability weights adjusting for selective dropout. Only participants surviving beyond week 20 (SELECT) or week 13 (SOUL) were included. All models were adjusted for sex, age, smoking status, baseline body weight, baseline clock score and sample storage time. As a sensitivity analysis, individual protein changes from baseline were evaluated as mediators using the product-of-coefficients approach, in which the indirect effect for each protein was estimated as the product of the treatment effect on the mediator (α) and the mediator–outcome association (β). Statistical significance of mediation hypotheses was

assessed using the joint-significance approach, with false-discovery-rate control using the high-dimensional mediation testing procedure described by Dai, J. Y. *et al*[16].

Methods references

1. Lincoff, A. M. *et al*. Semaglutide and Cardiovascular Outcomes in Obesity without Diabetes. *N. Engl. J. Med.* **389**, 2221–2232 (2023).
2. Kosiborod, M. N. et al. Semaglutide in Patients with Heart Failure with Preserved Ejection Fraction and Obesity. N. Engl. J. Med. 389, 1069–1084 (2023).
3. Kosiborod, M. N. et al. Semaglutide in Patients with Obesity-Related Heart Failure and Type 2 Diabetes. N. Engl. J. Med. 390, 1394–1407 (2024).
4. Perkovic, V. et al. Effects of Semaglutide on Chronic Kidney Disease in Patients with Type 2 Diabetes. N. Engl. J. Med. 391, 109–121 (2024).
5. McGuire, D. K. et al. Oral Semaglutide and Cardiovascular Outcomes in High-Risk Type 2 Diabetes. N. Engl. J. Med. 392, 2001–2012 (2025).
6. Assarsson, E. et al. Homogenous 96-Plex PEA Immunoassay Exhibiting High Sensitivity, Specificity, and Excellent Scalability. PLoS ONE 9, e95192 (2014).
7. Goeminne, L. J. E. et al. Plasma protein-based organ-specific aging and mortality models unveil diseases as accelerated aging of organismal systems. Cell Metab. 37, 205-222.e6 (2025).
8. Jia, X. et al. Plasma proteomic signature of frailty in 50,506 adults. Cell Metab. 38, 1012–1028.e6 (2026).
9. Smith, A., Elliott, P., Mayr, M., Dehghan, A. & Tzoulaki, I. Proteomic risk scores for predicting common diseases using linear and neural network models in the UK biobank. Sci. Rep. 15, 20520 (2025).
10. Herdman, M. et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual. Life Res. Int. J. Qual. Life Asp. Treat. Care Rehabil. 20, 1727–1736 (2011).
11. Brod, M. et al. Development of the Weight-Related Sign and Symptom Measure. J. Patient-Rep. Outcomes 2, 17 (2018).
12. Spertus, J. A. & Jones, P. G. Development and Validation of a Short Version of the Kansas City Cardiomyopathy Questionnaire. Circ. Cardiovasc. Qual. Outcomes 8, 469–476 (2015).
13. Lumley, T. Analysis of Complex Survey Samples. J. Stat. Softw. 9, (2004).
14. Tingley, D., Yamamoto, T., Hirose, K., Keele, L. & Imai, K. mediation : R Package for Causal Mediation Analysis. J. Stat. Softw. 59, (2014).

15. Imai, K., Keele, L. & Tingley, D. A general approach to causal mediation analysis. Psychol. Methods 15, 309–334 (2010).
16. Dai, J. Y., Stanford, J. L. & LeBlanc, M. A multiple-testing procedure for high-dimensional mediation hypotheses. J. Am. Stat. Assoc. 117, 198–213 (2022).

Supplementary information

Semaglutide reduces proteomic organ-specific biological age across randomized clinical trials

Dermit, M *et al*.

Contents

Supplementary Table 1 | Baseline characteristics by trial and treatment (full analysis set)

	SELECT		STEP-HFpEF		STEP-HFpEF DM		FLOW		SOUL	
Characteristic	**Placebo (*N* = 8,801)**	**Semaglutide 2.4 mg† SC (*N* = 8,803)**	**Placebo (*N* = 266)**	**Semaglutide 2.4 mg† SC (*N* = 263)**	**Placebo (*N* = 306)**	**Semaglutide 2.4 mg† SC (*N* = 310)**	**Placebo (*N* = 1,766)**	**Semaglutide 1.0 mg† SC (*N* = 1,767)**	**Placebo (*N* = 4,825)**	**Oral semaglutide 14 mg† (*N* = 4,825)**
Age, years	61.6±8.8	61.6±8.9	67.5±10.4	68.5±9.4	68.8±9.1	67.9±8.7	66.7±9.0	66.6±9.0	66.1±7.5	66.1±7.6
Female sex	2424 (28)	2448 (28)	148 (56)	149 (57)	145 (47)	128 (41)	550 (31)	519 (29)	1414 (29)	1376 (29)
White race	7404 (84)	7387 (84)	252 (95)	255 (97)	268 (88)	251 (81)	1168 (66)	1155 (65)	3321 (69)	3327 (69)
Body weight, kg	96.8±17.8	96.5±17.5	108.4±23	108.3±222	105.2±21.4	106.4±20.9	89.8±21.2	89.5±19.8	88.3±19.6	87.5±19.1
BMI, kg/m^2	33.4±5	33.3±5	38.5±6.9	38.4±6.5	37.9±6.2	37.9±5.5	32±6.5	31.9±6.1	31.2±5.9	31 ±5.7
Current smoker	1464 (17)	1486 (17)	27 (10)	24 (9)	32 (10)	28 (9)	206 (12)	223 (13)	584 (12)	545 (11)
SBP, mmHg	130.9±15.3	131±15.6	131±15.4	133.3±15.3	134.8±14.1	132.8±13.9	138.4±15.4	138.9±16.1	134.7±16.4	134.6±16.3
DBP, mmHg	79.2±9.9	79.4±10.0	78.7±10.1	77.9±10.5	76.4±10.2	78.1±9.3	76.1±10.0	76.8±10.0	76.7±10.1	76.6±10.1
T2D	0 (0)	0 (0)	1 (0)	1 (0)	306 (100)	310 (100)	1,766 (100)	1,766 (100)	4,825 (100)	4,824 (100)
eGFR, ml/min/1.73 m^2	82.5±17.3	82.4±17.5	Not performed	Not performed	67.2±23.0	69.9±23.1	47.1±14.7	46.9±15.6	73.6±22.6	74.0±22.6

HbA_{1c}, %	5.78±0.33	5.78±0.34	Not performed	Not performed	7.08±1.09	6.95±1.15	7.77±1.28	7.79±1.30	7.96±1.13	7.96±1.16
Median (IQR) hsCRP, mg/l	1.8 (0.9–4.1)	1.9 (0.9–4.2)	3.9 (2.0–8.4)	3.8 (1.9–7.0)	3.3 (1.6–8.4)	3.7 (1.8–8.4)	2.5 (1.1–6.1)	2.4 (1.1–5.6)	2.0 (0.9–4.5)	2.0 (0.9–4.3)

Data are presented as mean±s.d. or *n* (%) unless otherwise indicated.

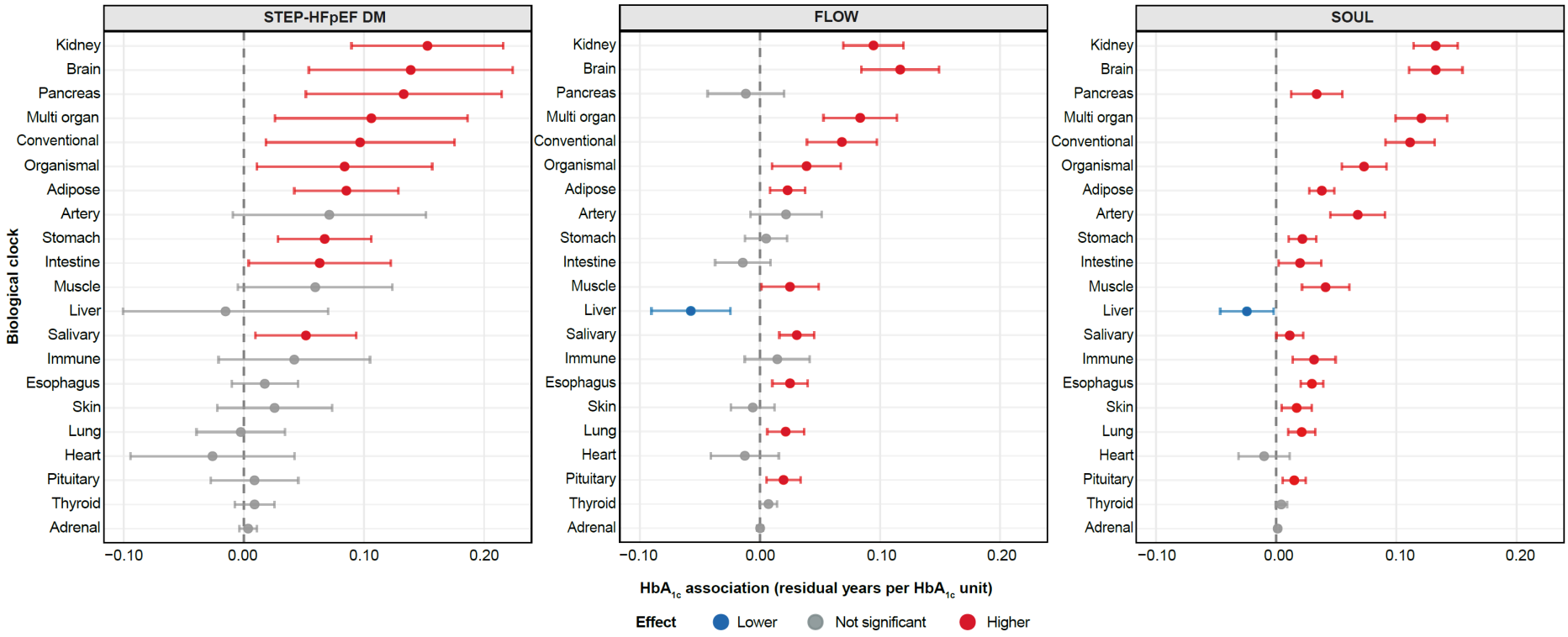
[†]Maximum dose.

BMI, body mass index; DBP, diastolic blood pressure. DM, diabetes mellitus. eGFR, estimated glomerular filtration rate; HbA_{1c}, glycated hemoglobin; HFpEF, heart failure with preserved ejection fraction; hsCRP, high-sensitivity C-reactive protein; IQR: interquartile range; *N*, number of participants; SBP, systolic blood pressure. s.d., standard deviation; T2D, type 2 diabetes.

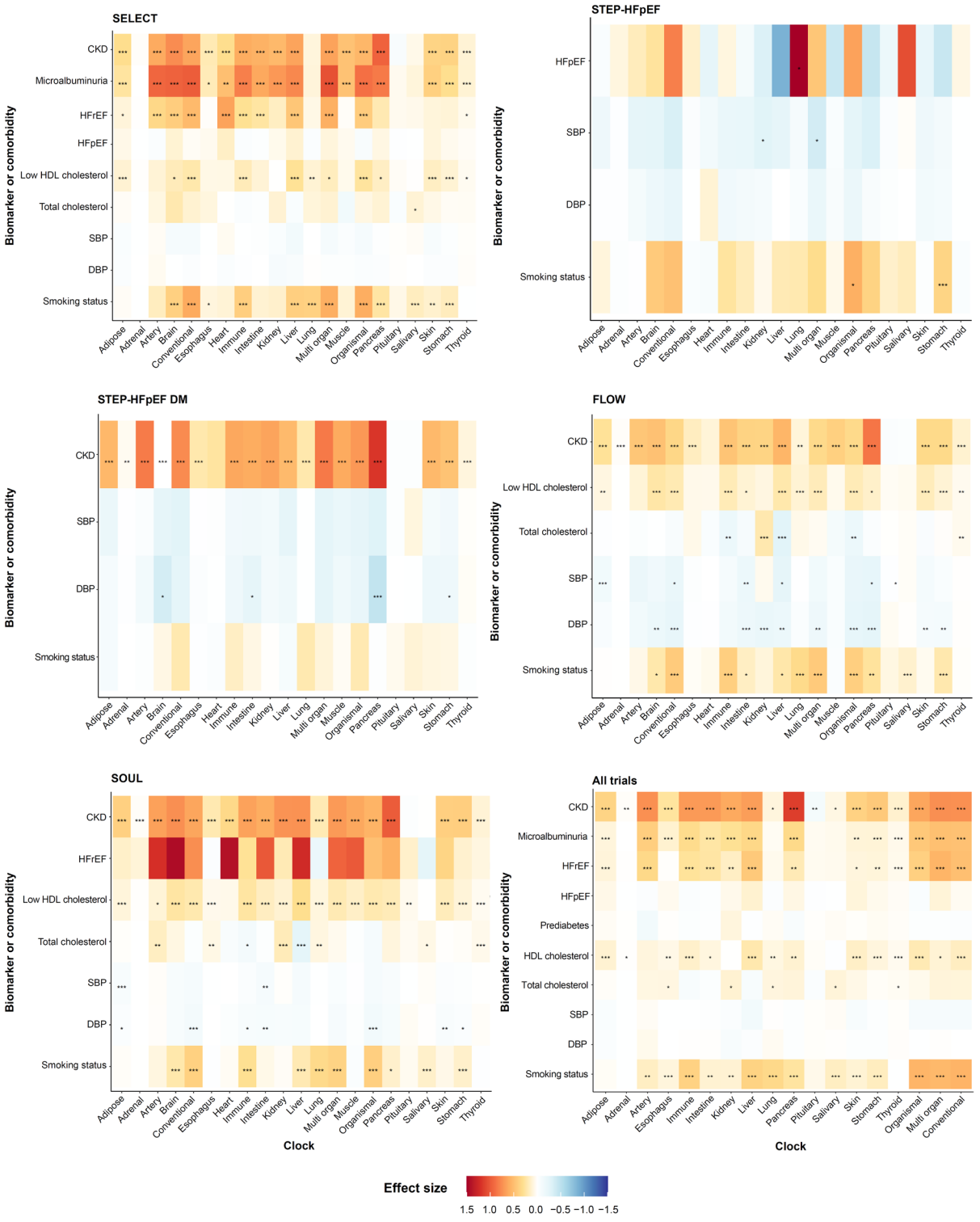
Supplementary Fig. 1 | Clinical associations and prognostic relevance of organ-specific proteomic biological-age clocks

a, Association of baseline HbA_{1c} with organ-specific biological-age residuals in trials enrolling participants with diabetes (STEP-HFpEF DM, FLOW and SOUL). Effect sizes represent residual years per 1 mmol/mol HbA_{1c} unit from linear regression models adjusted for age, sex, smoking status and ethnicity. Significant associations (Holm-adjusted $P < 0.05$) are highlighted in red.**b**,Relation between biomarkers and organ-specific mortality models in SELECT, STEP-HFpEF, STEP-HFpEF DM, FLOW, and SOUL. Heatmap shows the associations of biomarkers with organ log(mortality hazard) baseline values, after correcting for chronological age, sex, their interaction, smoking status and race, except for the smoking model, where smoking was not included as covariate. **c, d**, Heatmaps showing effect sizes for the association between baseline worst-quartile quality of life and organ-specific predicted biological age in SELECT and STEP-HFpEF DM after correcting for chronological age, sex, their interaction, current smoking status and race. Each row represents a quality-of-life measure and each column an organ-specific or composite clock. * Holm–Bonferroni-adjusted < 0.05; ** Holm–Bonferroni-adjusted < 0.01; *** Holm–Bonferroni-adjusted < 0.001. Quality-of-life measures reported by fewer than 10% of participants within a trial were excluded. EQ-5D-5L, EuroQol 5-dimension 5-level; WRSS, Weight-Related Symptom Scale; KCCQ, Kansas City Cardiomyopathy Questionnaire. **e, f,** Hierarchical clustering of organ-specific biological-age clock models predictive value for organ-relevant outcomes in **e,** SELECT, and **f,** SOUL. Hazard ratios per unit of standard deviation in the kidney and heart models for significant recorded time to event. All-cause death is shown in the heatmap above. Whiskers denote 95% confidence intervals. **g,** Prognostic specificity for kidney and heart clocks in SOUL. Hazard ratios per unit of standard deviation in the kidney model and heart models for significant recorded time to event. All- cause death is shown in the heatmap above. Whiskers denote 95% confidence intervals.

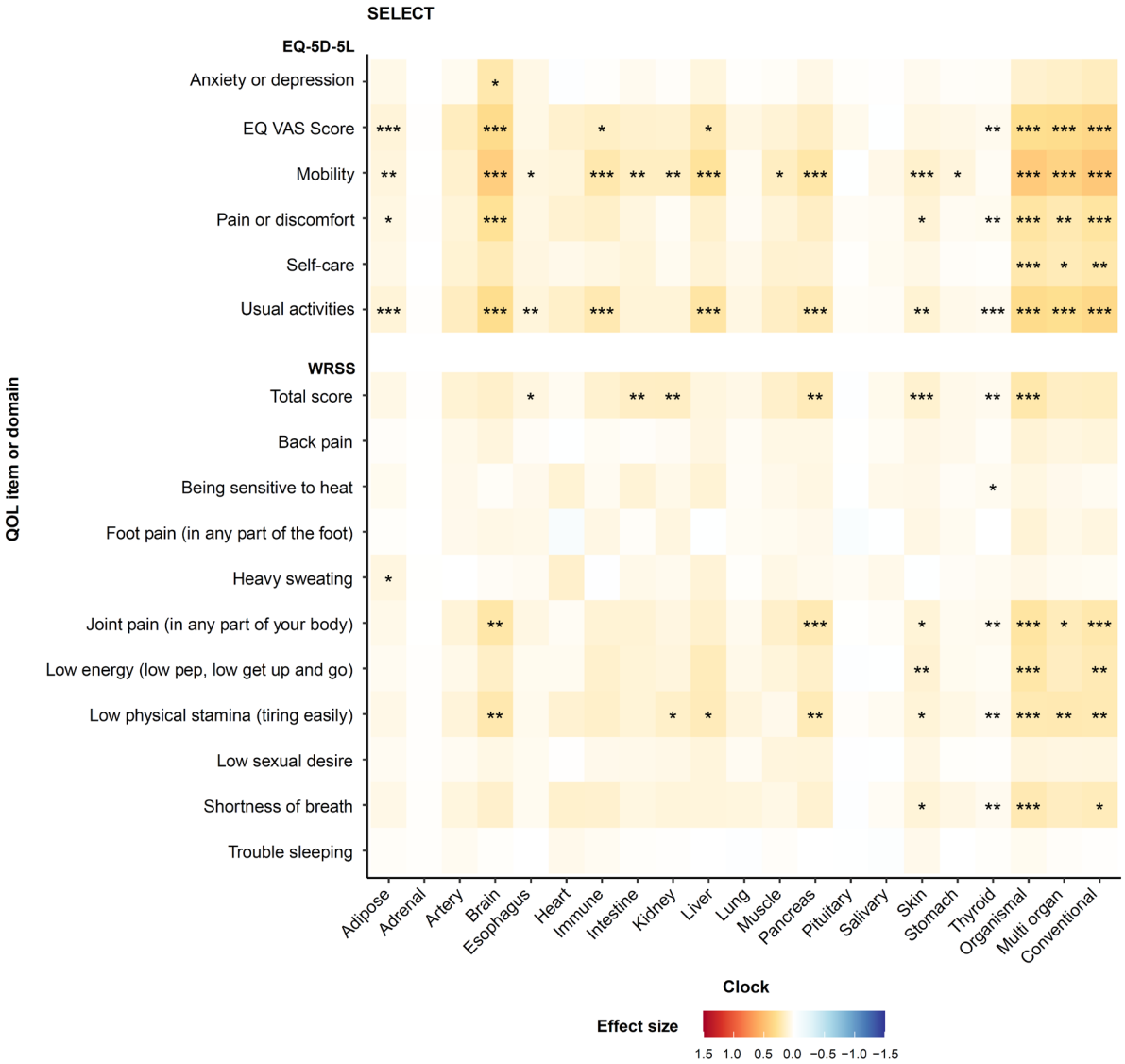
a


STEP-HFpEF DM
FLOW
SOUL
Biological clock
Kidney
Brain
Pancreas
Multi organ
Conventional
Organismal
Adipose
Artery
Stomach
Intestine
Muscle
Liver
Salivary
Immune
Esophagus
Skin
Lung
Heart
Pituitary
Thyroid
Adrenal
−0.10
0.00
0.10
0.20
HbA1c association (residual years per HbA1c unit)
Effect
Lower
Not significant
Higher

b


SELECT
STEP-HFpEF
STEP-HFpEF DM
FLOW
SOUL
All trials
Biomarker or comorbidity
Clock
CKD
Microalbuminuria
HFrEF
HFpEF
Prediabetes
Low HDL cholesterol
HDL cholesterol
Total cholesterol
SBP
DBP
Smoking status
Adipose
Adrenal
Artery
Brain
Conventional
Esophagus
Heart
Immune
Intestine
Kidney
Liver
Lung
Multi organ
Muscle
Organismal
Pancreas
Pituitary
Salivary
Skin
Stomach
Thyroid
Effect size
1.5 1.0 0.5 0.0 −0.5 −1.0 −1.5

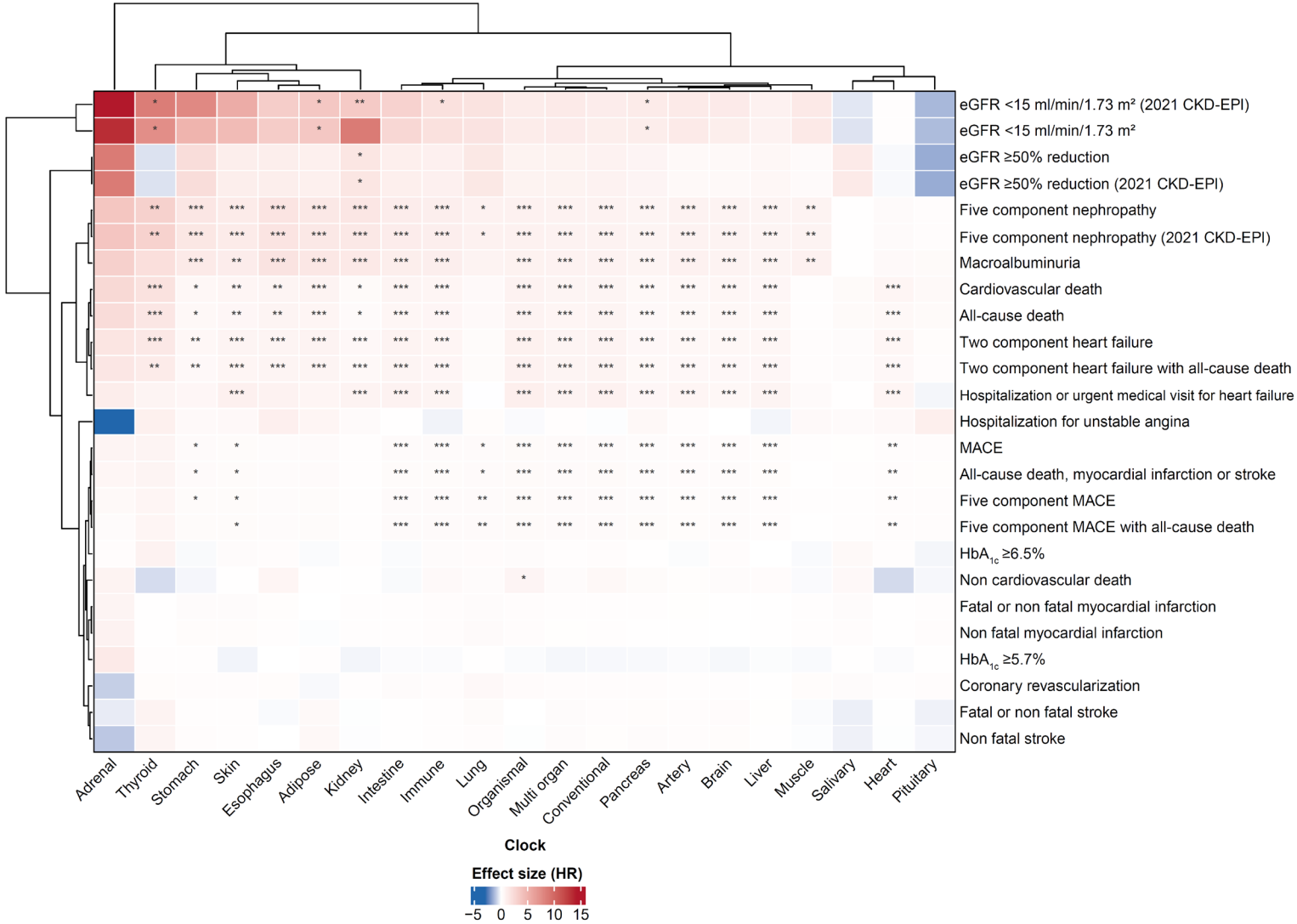
c


SELECT
EQ-5D-5L
Anxiety or depression
EQ VAS Score
Mobility
Pain or discomfort
Self-care
Usual activities
WRSS
Total score
Back pain
Being sensitive to heat
Foot pain (in any part of the foot)
Heavy sweating
Joint pain (in any part of your body)
Low energy (low pep, low get up and go)
Low physical stamina (tiring easily)
Low sexual desire
Shortness of breath
Trouble sleeping
QOL item or domain
Adipose
Adrenal
Artery
Brain
Esophagus
Heart
Immune
Intestine
Kidney
Liver
Lung
Muscle
Pancreas
Pituitary
Salivary
Skin
Stomach
Thyroid
Organismal
Multi organ
Conventional
Clock
Effect size
1.5 1.0 0.5 0.0 −0.5 −1.0 −1.5

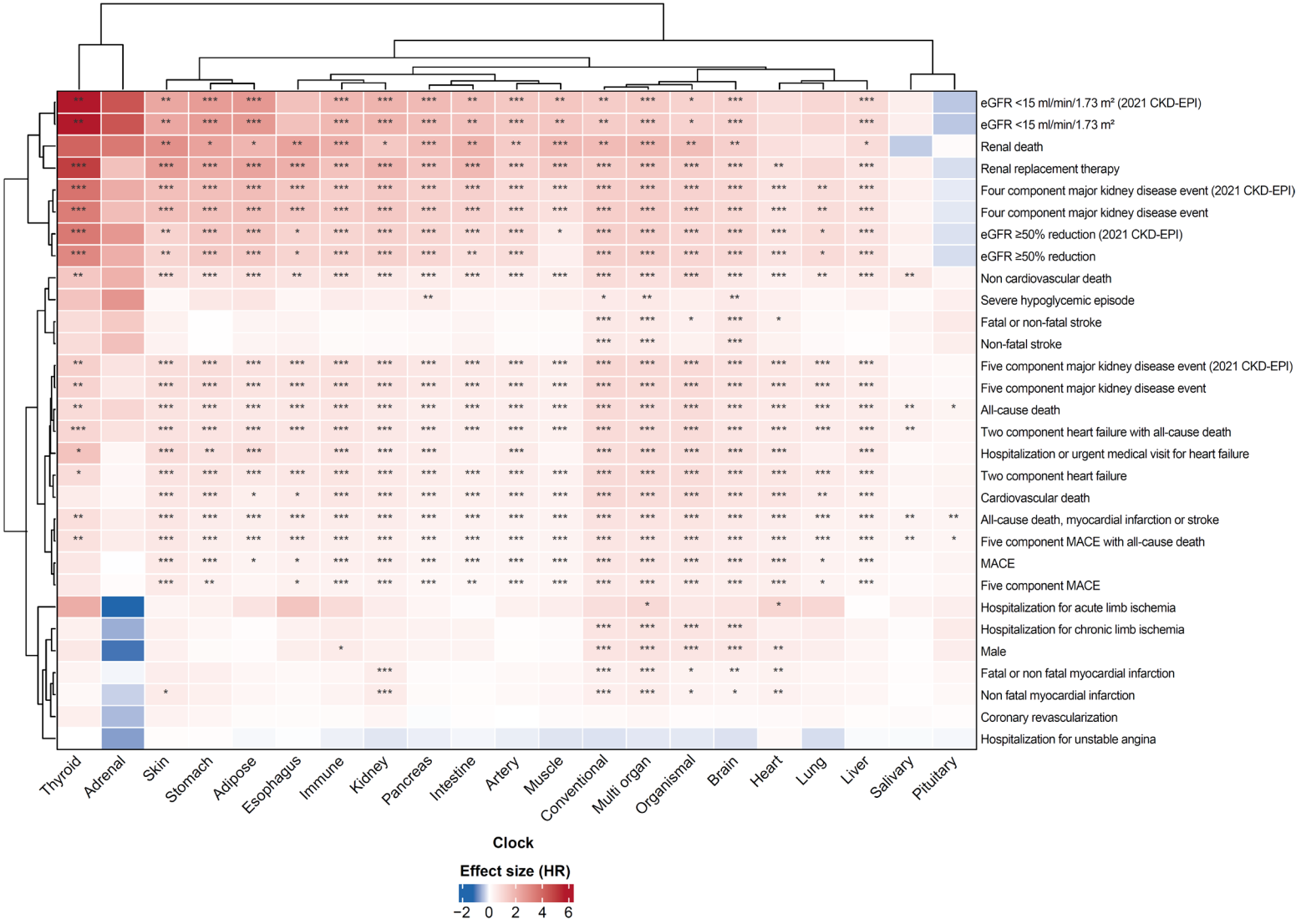
d


STEP-HFpEF DM
EQ-5D-5L
Anxiety or depression
EQ VAS score
Mobility
Pain or discomfort
Self-care
Usual activities
KCCQ
Activity: Climbing a flight of stairs without stopping
Activity: Doing yardwork, housework or carrying groceries
Activity: Dressing yourself
Times forced to sleep in a chair/with at least 3 pillows because of shortness of breath
Have your symptoms of heart failure changed?
How much has swelling in your feet, ankles or legs bothered you?
How much has your heart failure limited your enjoyment of life?
How often have you felt discouraged or down in the dumps?
Activity: Hurrying or jogging
How much has your heart failure limited intimate relationships with loved ones?
How much has your heart failure limited visiting family or friends out of your home?
How much has your heart failure limited your hobbies, recreational activities?
How much has your heart failure limited your working or doing household chores?
How much has your fatigue bothered you?
How much has shortness of breath bothered you?
Activity: showering/bathing
How would you feel spending the rest of your life with your heart failure as it is right now?
Times fatigue limited your ability to do what you want
Times shortness of breath limited your ability to do what you wanted
Times you have swelling in your feet, ankles or legs that woke you up in the morning
Understand things you are able to do to keep symptoms getting worse
Activity: Walking one block on level ground
Know what to do/whom to call if your heart failure gets worse
KCCQS
Clinical summary score (calculated)
Overall summary score
Physical limitation score
Quality-of-life score
Social limitation score
Symptom burden score
Symptom frequency score
Symptom self-efficacy score
Symptom stability score
Total symptom score (score)
PGI
Change - symptoms
Change - walk quickly
Status - symptoms
Status - walk quickly
QOL item or domain
Adipose
Adrenal
Artery
Brain
Esophagus
Heart
Immune
Intestine
Kidney
Liver
Lung
Muscle
Pancreas
Pituitary
Salivary
Skin
Stomach
Thyroid
Organismal
Multi organ
Conventional
Clock
Effect size
1.5 1.0 0.5 0.0 −0.5 −1.0 −1.5

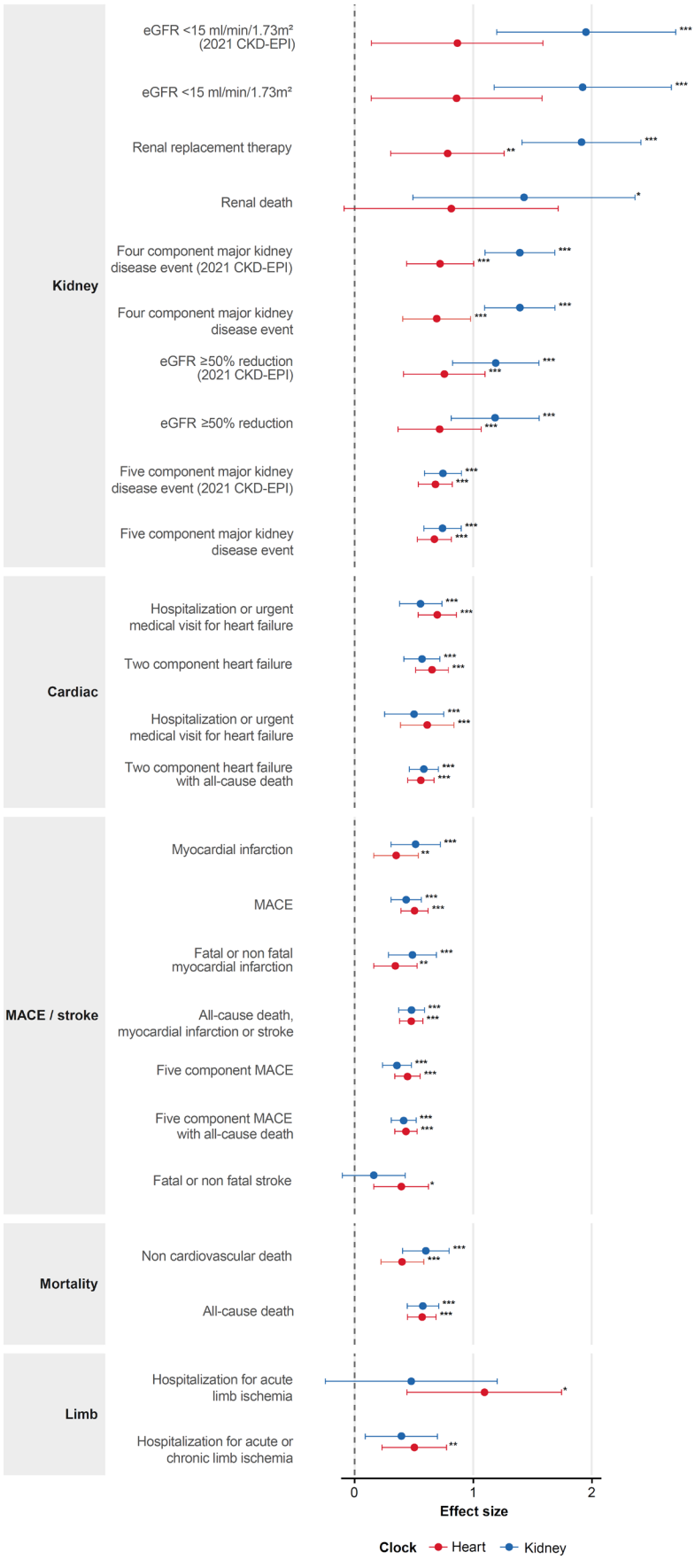
e


eGFR <15 ml/min/1.73 m² (2021 CKD-EPI)
eGFR <15 ml/min/1.73 m²
eGFR ≥50% reduction
eGFR ≥50% reduction (2021 CKD-EPI)
Five component nephropathy
Five component nephropathy (2021 CKD-EPI)
Macroalbuminuria
Cardiovascular death
All-cause death
Two component heart failure
Two component heart failure with all-cause death
Hospitalization or urgent medical visit for heart failure
Hospitalization for unstable angina
MACE
All-cause death, myocardial infarction or stroke
Five component MACE
Five component MACE with all-cause death
HbA1c ≥6.5%
Non cardiovascular death
Fatal or non fatal myocardial infarction
Non fatal myocardial infarction
HbA1c ≥5.7%
Coronary revascularization
Fatal or non fatal stroke
Non fatal stroke
Adrenal
Thyroid
Stomach
Skin
Esophagus
Adipose
Kidney
Intestine
Immune
Lung
Organismal
Multi organ
Conventional
Pancreas
Artery
Brain
Liver
Muscle
Salivary
Heart
Pituitary
Clock
Effect size (HR)
−5 0 5 10 15

f


eGFR <15 ml/min/1.73 m² (2021 CKD-EPI)
eGFR <15 ml/min/1.73 m²
Renal death
Renal replacement therapy
Four component major kidney disease event (2021 CKD-EPI)
Four component major kidney disease event
eGFR ≥50% reduction (2021 CKD-EPI)
eGFR ≥50% reduction
Non cardiovascular death
Severe hypoglycemic episode
Fatal or non-fatal stroke
Non-fatal stroke
Five component major kidney disease event (2021 CKD-EPI)
Five component major kidney disease event
All-cause death
Two component heart failure with all-cause death
Hospitalization or urgent medical visit for heart failure
Two component heart failure
Cardiovascular death
All-cause death, myocardial infarction or stroke
Five component MACE with all-cause death
MACE
Five component MACE
Hospitalization for acute limb ischemia
Hospitalization for chronic limb ischemia
Male
Fatal or non fatal myocardial infarction
Non fatal myocardial infarction
Coronary revascularization
Hospitalization for unstable angina
Thyroid
Adrenal
Skin
Stomach
Adipose
Esophagus
Immune
Kidney
Pancreas
Intestine
Artery
Muscle
Conventional
Multi organ
Organismal
Brain
Heart
Lung
Liver
Salivary
Pituitary
Clock
Effect size (HR)
−2 0 2 4 6

g

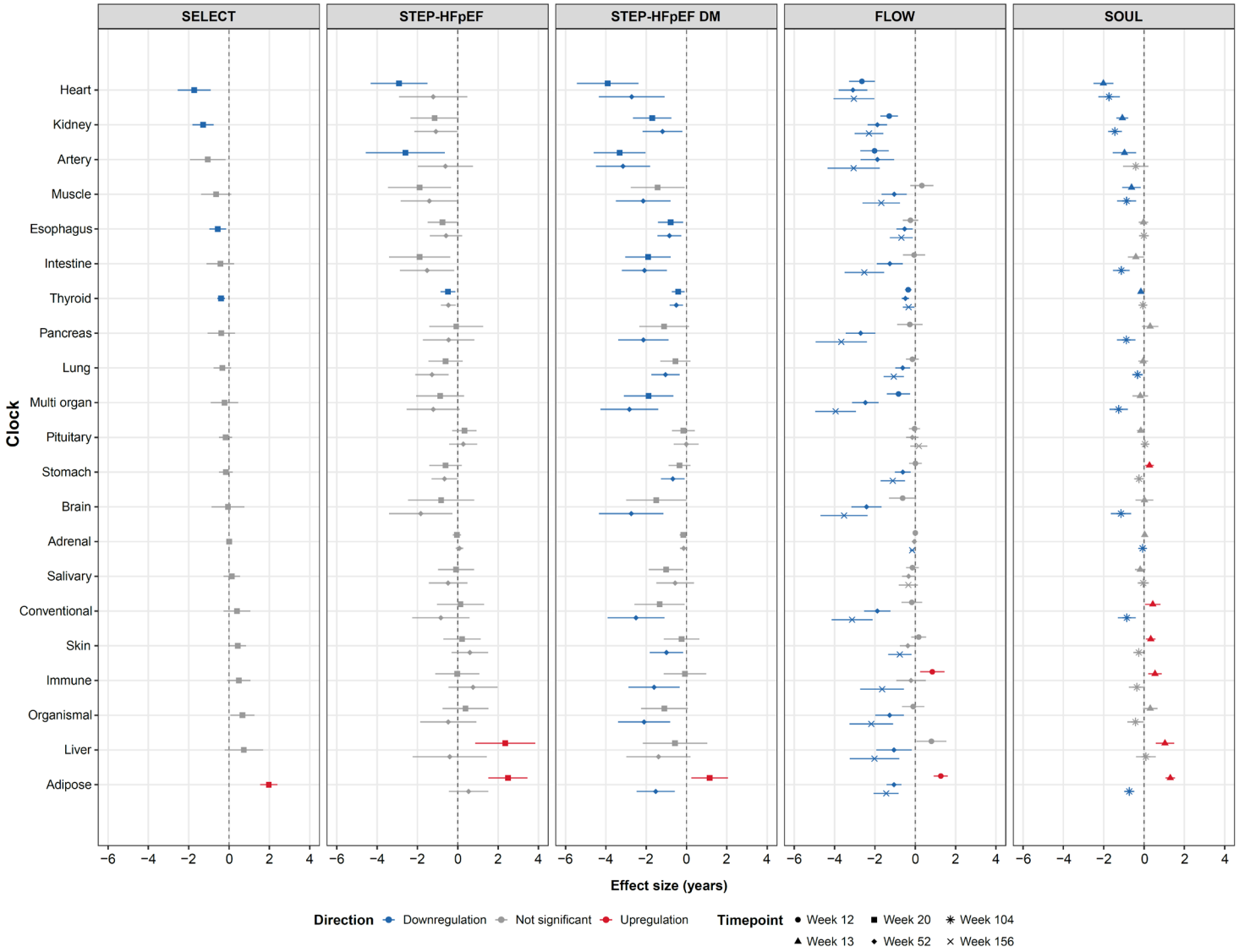

Kidney
eGFR <15 ml/min/1.73m² (2021 CKD-EPI)
eGFR <15 ml/min/1.73m²
Renal replacement therapy
Renal death
Four component major kidney disease event (2021 CKD-EPI)
Four component major kidney disease event
eGFR ≥50% reduction (2021 CKD-EPI)
eGFR ≥50% reduction
Five component major kidney disease event (2021 CKD-EPI)
Five component major kidney disease event
Cardiac
Hospitalization or urgent medical visit for heart failure
Two component heart failure
Hospitalization or urgent medical visit for heart failure
Two component heart failure with all-cause death
MACE / stroke
Myocardial infarction
MACE
Fatal or non fatal myocardial infarction
All-cause death, myocardial infarction or stroke
Five component MACE
Five component MACE with all-cause death
Fatal or non fatal stroke
Mortality
Non cardiovascular death
All-cause death
Limb
Hospitalization for acute limb ischemia
Hospitalization for acute or chronic limb ischemia
0
1
2
Effect size
Clock
Heart
Kidney

Supplementary Fig. 2 |Sensitivity and supportive analyses of semaglutide effects on organ-specific biological-age clocks.

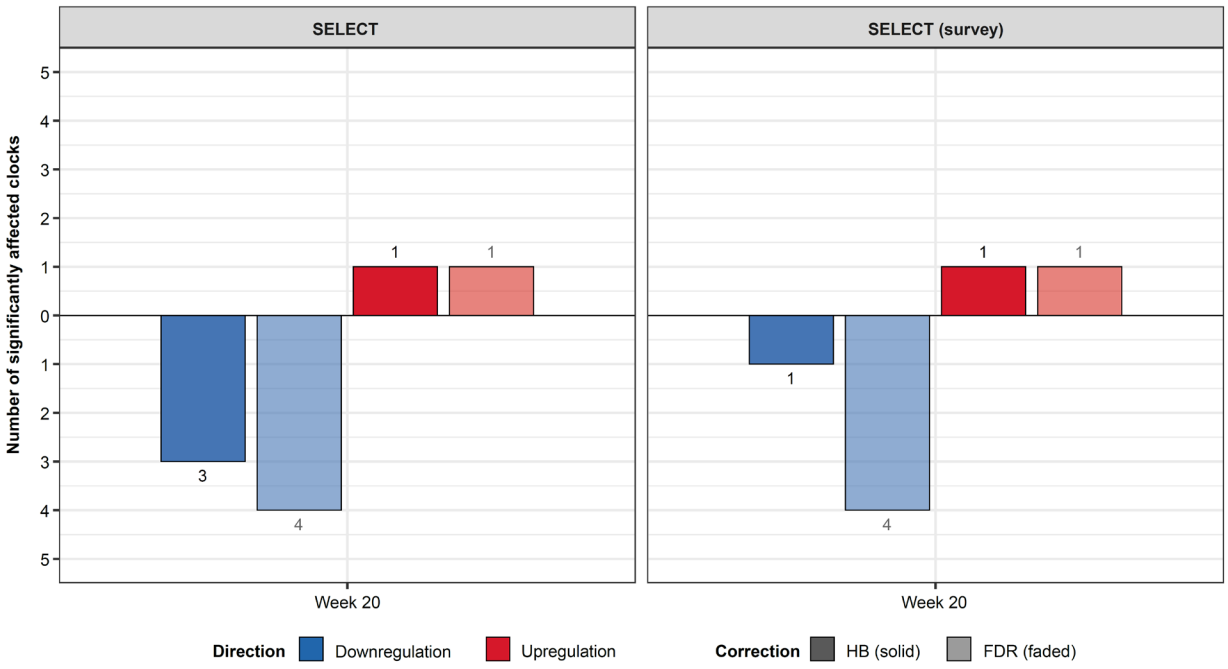
a, Effects of semaglutide on organ-specific biological age models, like Fig. 2a, but with less stringent FDR correction. Forest plots showing the semaglutide treatment-associated change in organ-specific biological-age models across five clinical trials: SELECT, STEP-HFpEF, STEP-HFpEF DM, FLOW and SOUL. Effect sizes represent the difference in estimated biological age between semaglutide treatment and placebo. Error bars denote 95% confidence intervals. Colors indicate significance and direction after FDR correction: blue, significantly downregulated; red, significantly upregulated; gray, non-significant. **b**, Number of significantly affected organ-specific biological-age clock models in standard and survey-weighted SELECT analyses. Bars show counts under Holm–Bonferroni (HB, solid) and FDR (faded) correction. Significance threshold: adjusted $P < 0.05$. **c**, Metformin-adjusted effects of semaglutide biological-age clock models. Forest plots showing semaglutide treatment-associated changes in organ-specific biological-age models across three semaglutide trials with substantial metformin usage: 71.0% (333/469) in STEP-HFpEF DM, 52.4% (1,190/2,271) in FLOW and 74.9% (3,937/5,255) in SOUL. Effect sizes represent the difference in estimated biological age between semaglutide and placebo at each timepoint, adjusted using the baseline metformin use model: CHG ~ BASE + TRT01P + med.metformin. Negative values indicate a reduction in predicted biological age (i.e. a younger clock estimate), whereas positive values indicate an increase. Each timepoint is represented by a distinct point shape (triangle: week 13; square: week 20; diamond: week 52; asterisk: week 104; cross: week 156). Error bars denote 95% confidence intervals. *P* values were corrected for multiplicity using Holm–Bonferroni. Colors indicate significance and direction after correction: blue, significantly downregulated; red, significantly upregulated; gray, non-significant. **d**, Forest plots showing the interaction between speed of aging and semaglutide treatment effect on organ-specific biological-age clock changes across five clinical trials: SELECT, STEP-HFpEF, STEP-HFpEF DM, FLOW and SOUL. Effect sizes represent the interaction term from a linear model in which the change in biological age (BioAge change) is regressed on BioAge at baseline, speed of aging and their interaction with treatment (TRT01P). A negative interaction term indicates that semaglutide preferentially reduces biological age in individuals with a greater baseline age deviation (i.e. greater treatment benefit in faster agers), whereas a positive term indicates the reverse. Each timepoint is represented by a distinct point shape (circle, week 12; triangle, week 13; square, week 20; diamond, week 52; asterisk, week 104; cross, week 156). Error Error bars denote 95% confidence intervals. Significance was assessed at $P < 0.05$ after Holm–Bonferroni correction for multiple comparisons across organs and

timepoints within each trial. Colors indicate direction and significance after correction: blue, significantly negative interaction; red, significantly positive interaction; grey, non-significant. **e,** Effects of semaglutide on organ-specific biological-age clock models. Forest plots showing the semaglutide treatment-associated change in organ-specific biological-age models across two clinical trials with SomaScan data: STEP 1 and STEP 2 (2.4 mg and 1.0 mg semaglutide). Effect sizes represent the difference in estimated biological age between treatment and placebo at week 68.

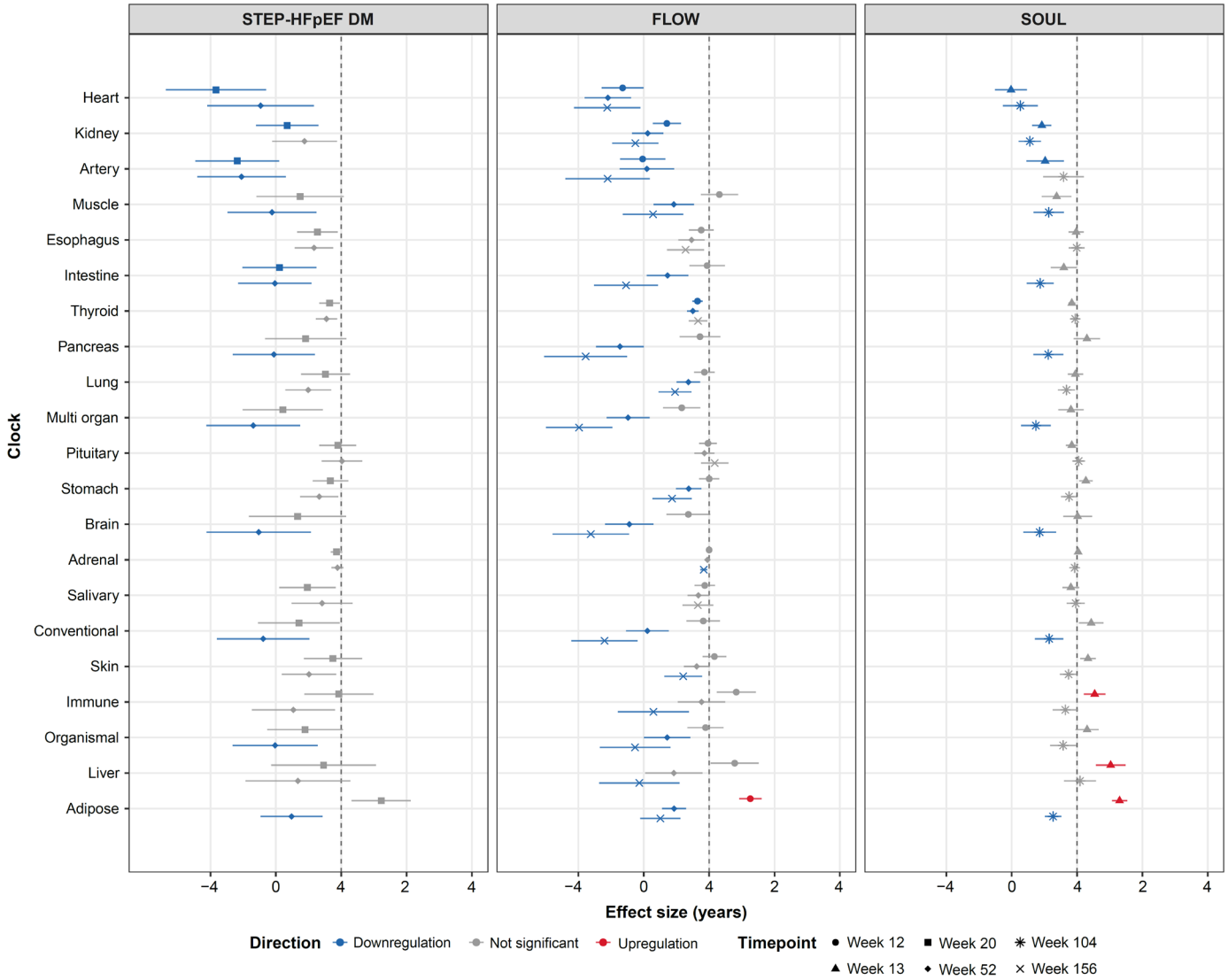
a


SELECT
STEP-HFpEF
STEP-HFpEF DM
FLOW
SOUL
Clock
Heart
Kidney
Artery
Muscle
Esophagus
Intestine
Thyroid
Pancreas
Lung
Multi organ
Pituitary
Stomach
Brain
Adrenal
Salivary
Conventional
Skin
Immune
Organismal
Liver
Adipose
−6 −4 −2 0 2 4
Effect size (years)
Direction Downregulation Not significant Upregulation
Timepoint Week 12 Week 20 Week 104
Week 13 Week 52 Week 156

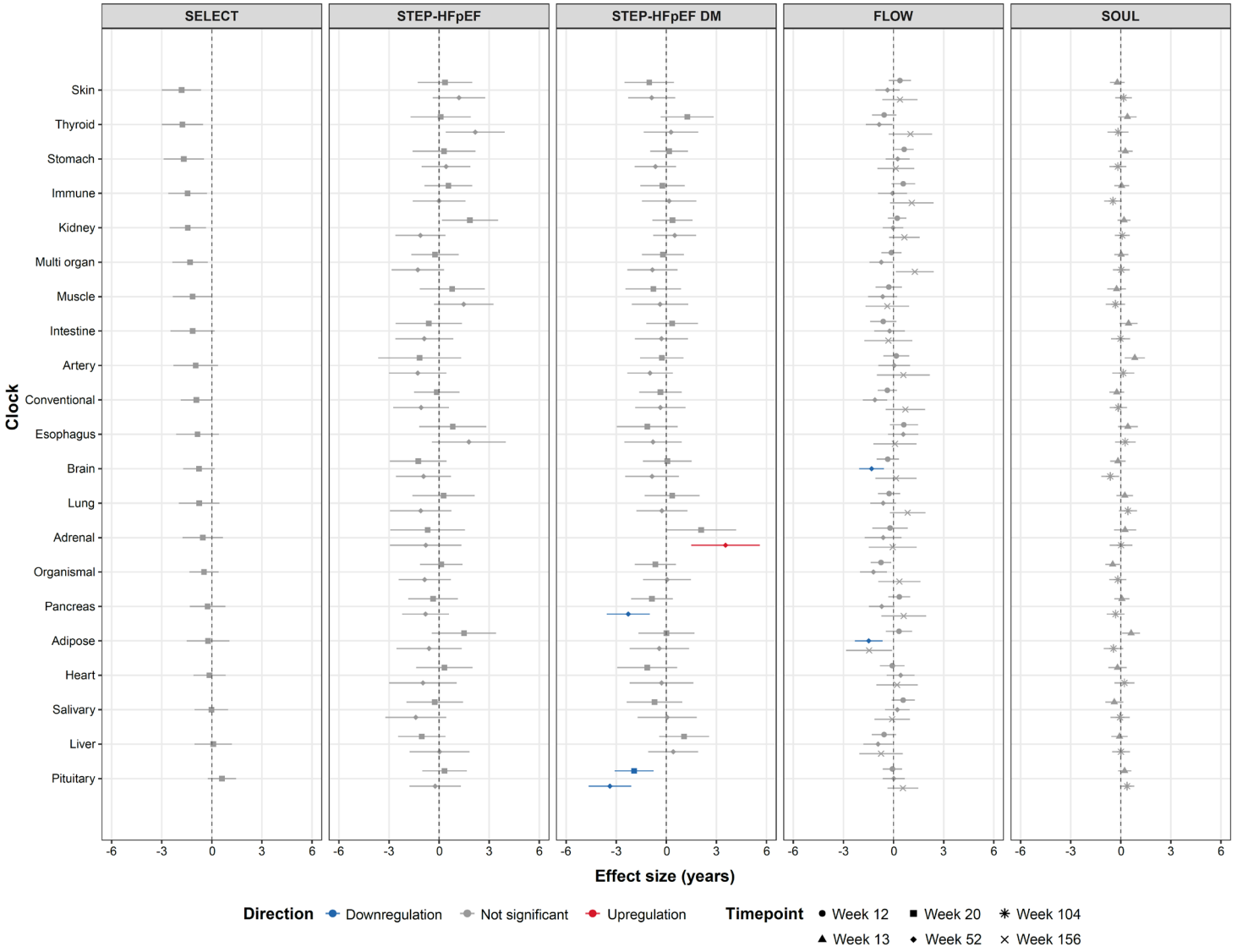
b


SELECT
SELECT (survey)
Number of significantly affected clocks
5
4
3
2
1
0
1
2
3
4
5
1
1
3
4
1
1
1
4
Week 20
Week 20
Direction
Downregulation
Upregulation
Correction
HB (solid)
FDR (faded)

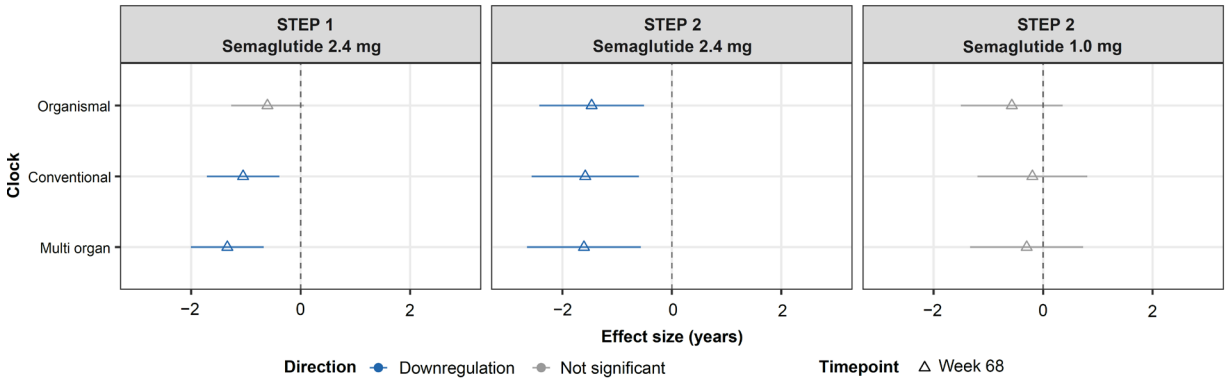
c


STEP-HFpEF DM
FLOW
SOUL
Clock
Heart
Kidney
Artery
Muscle
Esophagus
Intestine
Thyroid
Pancreas
Lung
Multi organ
Pituitary
Stomach
Brain
Adrenal
Salivary
Conventional
Skin
Immune
Organismal
Liver
Adipose
−4
0
4
2
4
Effect size (years)
Direction
Downregulation
Not significant
Upregulation
Timepoint
Week 12
Week 20
Week 104
Week 13
Week 52
Week 156

d


SELECT
STEP-HFpEF
STEP-HFpEF DM
FLOW
SOUL
Clock
Skin
Thyroid
Stomach
Immune
Kidney
Multi organ
Muscle
Intestine
Artery
Conventional
Esophagus
Brain
Lung
Adrenal
Organismal
Pancreas
Adipose
Heart
Salivary
Liver
Pituitary
-6
-3
0
3
6
Effect size (years)
Direction
Downregulation
Not significant
Upregulation
Timepoint
Week 12
Week 20
Week 104
Week 13
Week 52
Week 156

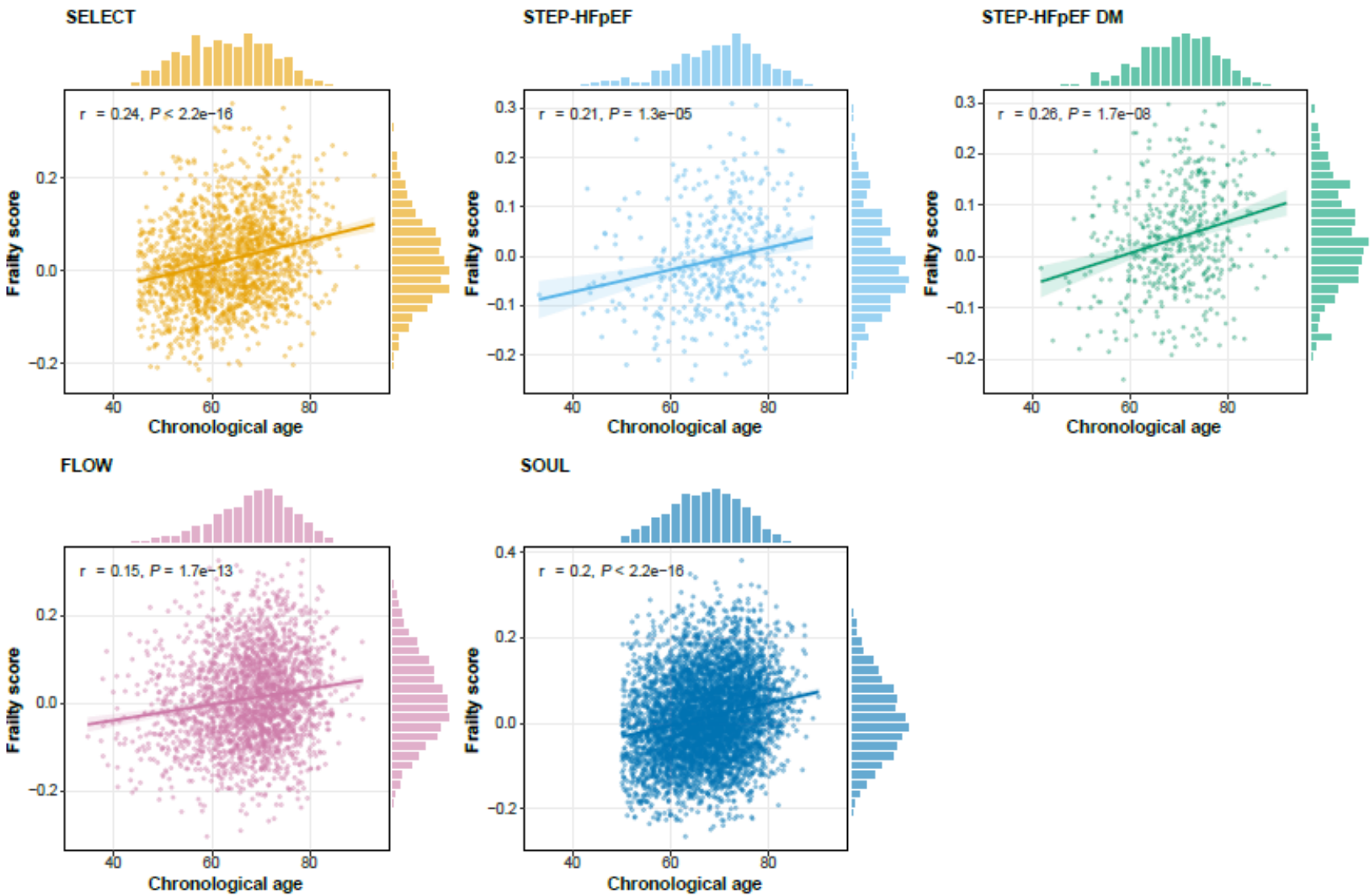
e


STEP 1
Semaglutide 2.4 mg
STEP 2
Semaglutide 2.4 mg
STEP 2
Semaglutide 1.0 mg
Clock
Organismal
Conventional
Multi organ
−2
0
2
Effect size (years)
Direction
Downregulation
Not significant
Timepoint
Week 68

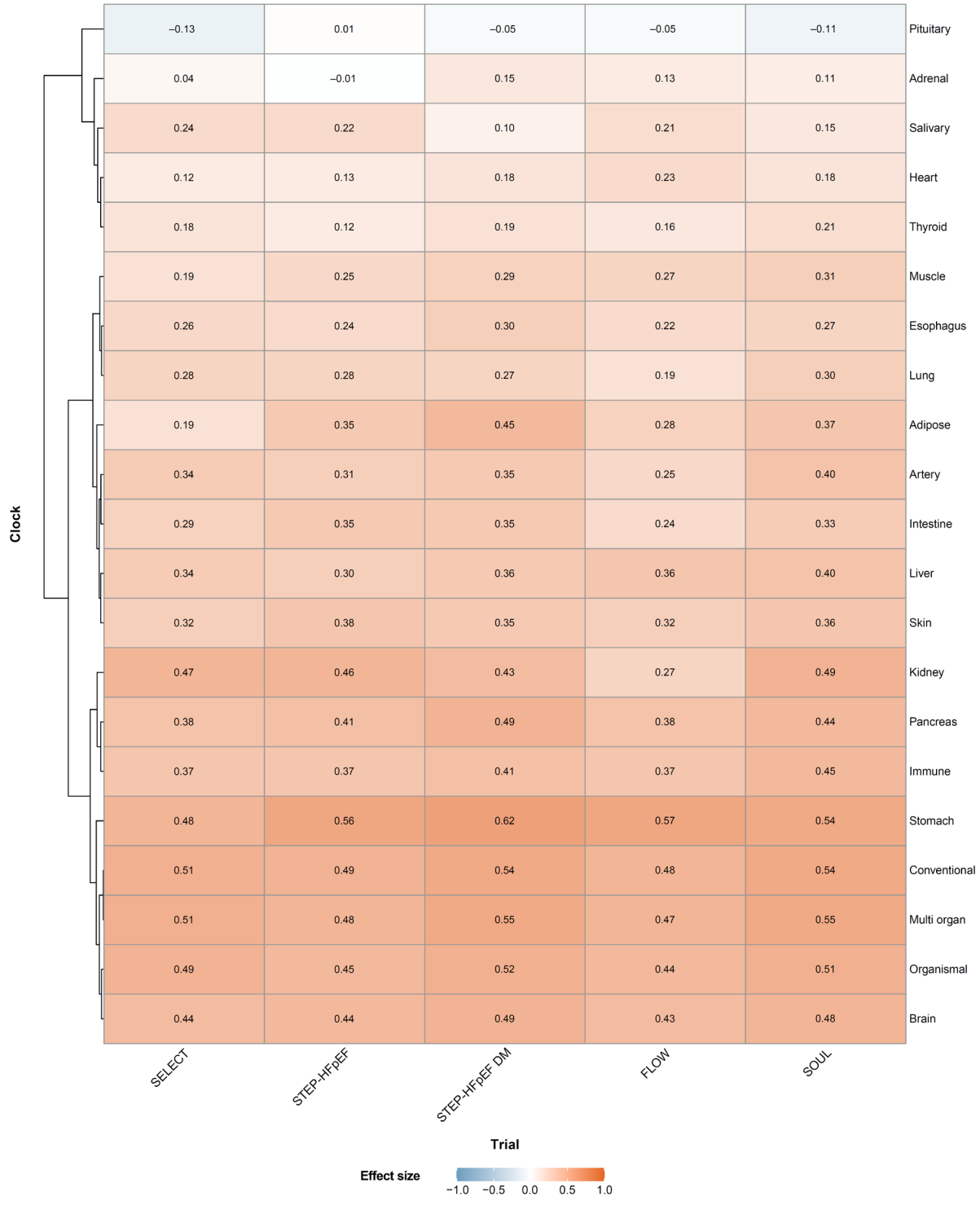
Supplementary Fig. 3 |Cross-sectional characterization of the proteomic frailty score (PFS) across trials.

a, Relation between chronological age and frailty score. Scatterplots show that PFS, from Jia, X. *et al*[22]*,* has a positive association with chronological age. Trial-fitted lines show the linear regression summarizing the observed relationship between frailty measures and chronological age. Shaded areas indicate 95% confidence intervals around the ordinary least squares regression line. r = correlation coefficient. **b**, Cross-sectional correlations between proteomic frailty-score residuals and organ-clock residuals across trials. Heatmap shows Pearson correlation coefficients between PFS and each organ-specific biological-age clock, computed on residuals after regressing out chronological age across SELECT (*n* = 1,641), STEP-HFpEF (*n* = 416), STEP-HFpEF DM (*n* = 469), FLOW (*n* = 2,271) and SOUL (*n* = 5,255). Organs are ordered by hierarchical clustering using Ward.D2 method with Euclidean distance. **c**, Association of sex with PFS residuals across trials. Forest plots show regression coefficients for female versus male participants (male as the reference) in SELECT, STEP-HFpEF, STEP-HFpEF DM, FLOW and SOUL. PFS residuals were age-adjusted, and models included sex, age, smoking status and race. Points represent regression coefficients and error bars denote 95% confidence intervals. Positive values indicate higher PFS residuals in females than in males.

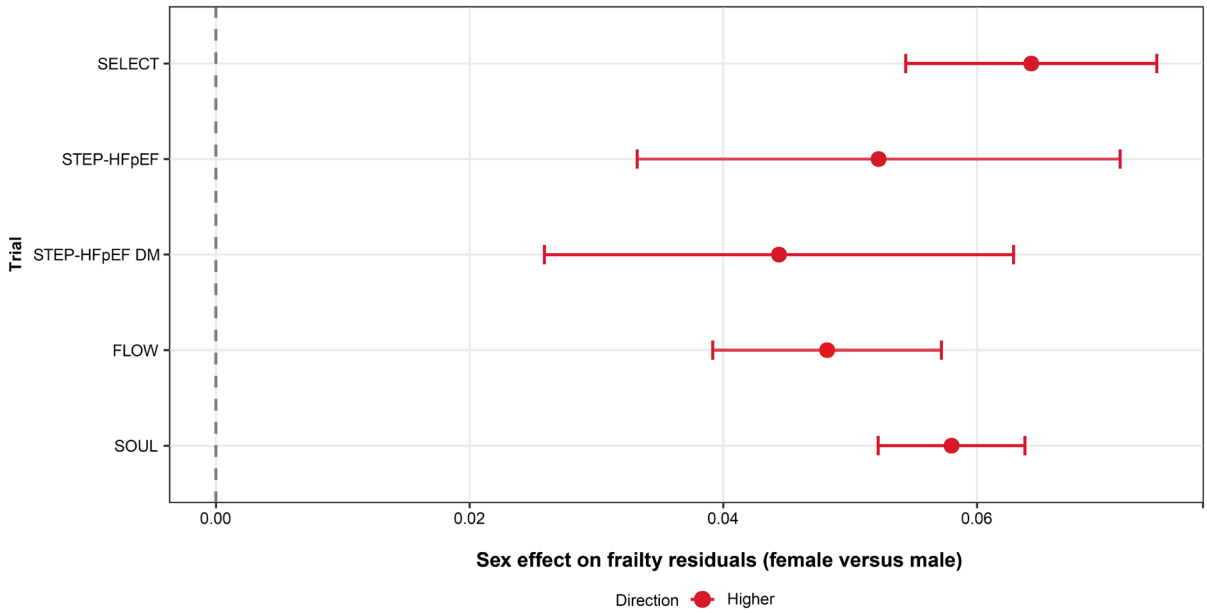
a


SELECT
r = 0.24, P < 2.2e−16
STEP-HFpEF
r = 0.21, P = 1.3e−05
STEP-HFpEF DM
r = 0.26, P = 1.7e−08
FLOW
r = 0.15, P = 1.7e−13
SOUL
r = 0.2, P < 2.2e−16
Frailty score
Chronological age

b


Clock
−0.13 0.01 −0.05 −0.05 −0.11 Pituitary
0.04 −0.01 0.15 0.13 0.11 Adrenal
0.24 0.22 0.10 0.21 0.15 Salivary
0.12 0.13 0.18 0.23 0.18 Heart
0.18 0.12 0.19 0.16 0.21 Thyroid
0.19 0.25 0.29 0.27 0.31 Muscle
0.26 0.24 0.30 0.22 0.27 Esophagus
0.28 0.28 0.27 0.19 0.30 Lung
0.19 0.35 0.45 0.28 0.37 Adipose
0.34 0.31 0.35 0.25 0.40 Artery
0.29 0.35 0.35 0.24 0.33 Intestine
0.34 0.30 0.36 0.36 0.40 Liver
0.32 0.38 0.35 0.32 0.36 Skin
0.47 0.46 0.43 0.27 0.49 Kidney
0.38 0.41 0.49 0.38 0.44 Pancreas
0.37 0.37 0.41 0.37 0.45 Immune
0.48 0.56 0.62 0.57 0.54 Stomach
0.51 0.49 0.54 0.48 0.54 Conventional
0.51 0.48 0.55 0.47 0.55 Multi organ
0.49 0.45 0.52 0.44 0.51 Organismal
0.44 0.44 0.49 0.43 0.48 Brain
SELECT
STEP-HFpEF
STEP-HFpEF DM
FLOW
SOUL
Trial
Effect size
−1.0 −0.5 0.0 0.5 1.0

c


SELECT
STEP-HFpEF
STEP-HFpEF DM
FLOW
SOUL
Trial
0.00
0.02
0.04
0.06
Sex effect on frailty residuals (female versus male)
Direction
Higher

Supplementary Fig. 4 |Analytical frameworks for treatment-effect and mediation analyses of biological-age scores.

Analytical frameworks for assessing the impact of semaglutide on biological-age scores. Left panel: standard treatment analysis illustrates the standard regression approach used to measure the total association (total effect) between semaglutide administration (independent variable, A) and changes in biological age (dependent variable, Y). This model estimates the cumulative effect without accounting for intermediate mechanisms. Right panel: mediation analysis presenting a set of directed acyclic graphs outlining the causal paths for mediation analysis. These models allow for the decomposition of the total effect into a direct effect of semaglutide on biological age (A → Y) and an indirect effect (A → M → Y) mediated through a third variable M. This approach provides critical insight into whether the effects of semaglutide on biological-age scores are mediated by established metabolic changes or operate through independent physiological pathways. Three specific candidate mediators are proposed: weight change; HbA_{1c} change, representing changes in glycemic control; and hsCRP change, representing changes in systemic inflammation.