# Alma Score and everyday health signals

Longer sleep, lower weight and lower walking heart rate among score improvers over 12 weeks, with an early signal by day 30

Alma | September 9, 2026 | Exploratory retrospective study

## Abstract

Over twelve weeks, 82 of 221 sustained food loggers increased their reconstructed food-only Alma Score. Among score improvers with paired Apple Health measurements, sleep increased 21.5 minutes per recorded night (n=33), weight decreased 1.25 kg (n=34), and walking heart rate decreased 2.05 bpm (n=25). A post hoc first-month extension identified an earlier resting-heart-rate decline of 1.55 bpm among 85 score improvers, compared with 0.04 bpm among 118 other active users. The within-group resting-heart-rate result survived correction across 104 new tests (q=0.014); the between-group result did not survive that broader correction. The original 12-week primary findings did not meet their multiple-testing threshold. These exploratory observations identify measurable early outcomes for prospective evaluation. Both groups used Alma, and the study was not randomized.

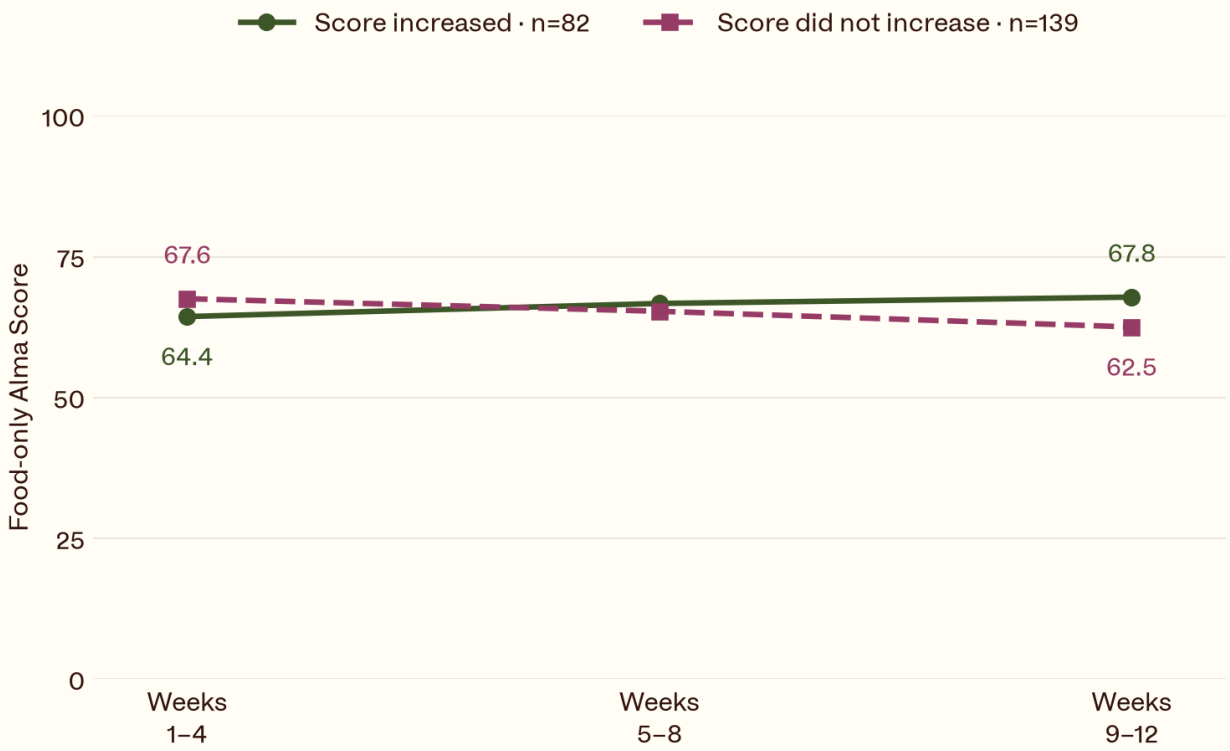
| Entry accounts | Sustained loggers | Score improvers | With health records |
|---|---|---|---|
| 6,884 | 221 (3.2%) | 82 | 73 of 82 |

Original window: weeks 1-4 versus weeks 9-12. Early window: days 1-7 versus days 24-30. Each horizon and marker has its own paired sample. All results concern recorded measurements in active Alma users.

# The question

When recorded food quality improves, what changes beyond the food log? This study follows food-only Alma Score trajectories alongside sleep, body measurements and cardiovascular signals recorded in Apple Health. It examines 12 weeks of sustained logging, then looks for early changes in the first 30 days.

The original cohort was selected for logging duration and completeness before examining HealthKit outcomes. Of 221 sustained loggers, 82 had a higher final-period score; 139 did not. Both groups were active Alma users. The analysis asks what accompanied score improvement, with the other active users providing an observational comparison.




Score trajectories define the groups. Mean food-only Alma Score in three 28-day periods. Improvers: 82; non-improvers: 139. The full 0-100 score range is shown. The increase in the selected group follows from its definition and is not an efficacy result.

# What changed alongside the score

Among score improvers with sufficient paired observations, sleep duration increased by 0.36 hours per night, or 21.5 minutes (n=33; 95% CI, 6.5 to 37.3 minutes). Weight decreased by 1.25 kg (n=34; 95% CI, -2.17 to -0.49), and walking heart rate decreased by 2.05 bpm (n=25; 95% CI, -4.01 to -0.26). These are unadjusted, within-person associations.

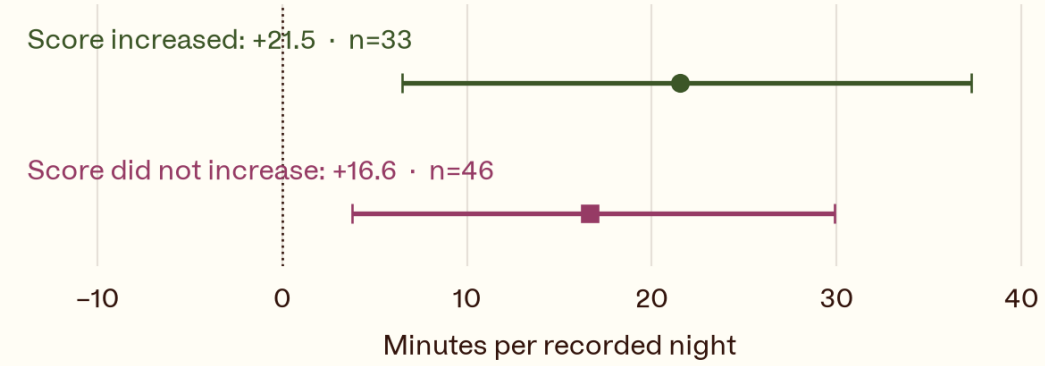
Both groups used Alma. People without score improvement also recorded longer sleep (+16.6 minutes, n=46), lower weight (-1.73 kg, n=52) and lower walking heart rate (-1.53 bpm, n=36). The between-group change intervals include zero for all three. No primary marker or between-group comparison meets a 5% false-discovery-rate threshold across its 18-test family.

Resting heart rate, HRV and estimated VO2 max did not show clear primary changes. Walking steadiness declined by 1.49 percentage points before multiplicity adjustment. Body fat declined more in non-improvers than in improvers; that between-group result also did not survive adjustment. All adequately observed independent markers appear below.

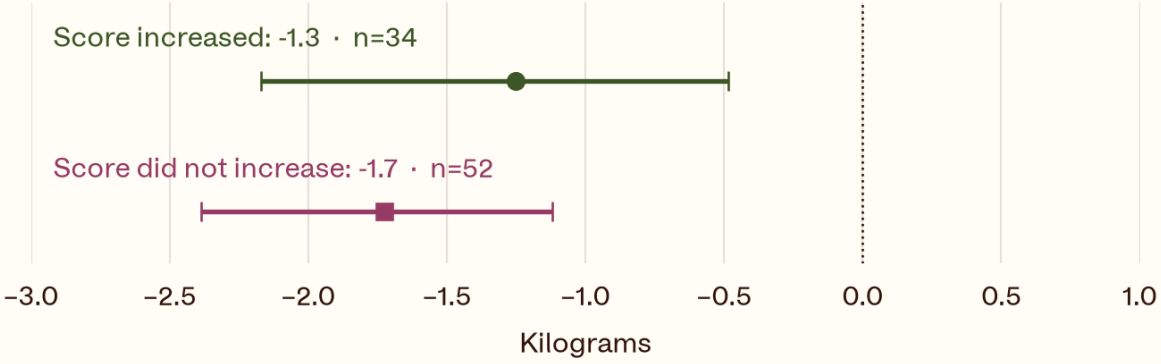
| Measurement | Paired n | Baseline | Endpoint | Change (95% CI) |
|---|---|---|---|---|
| Sleep duration (hr) | 33 | 6.68 | 7.04 | 0.36 (0.11 to 0.62) |
| Walking heart rate (bpm) | 25 | 94.52 | 92.47 | -2.05 (-4.01 to -0.26) |
| Weight (kg) | 34 | 79.90 | 78.65 | -1.25 (-2.17 to -0.49) |

# Changes over 12 weeks

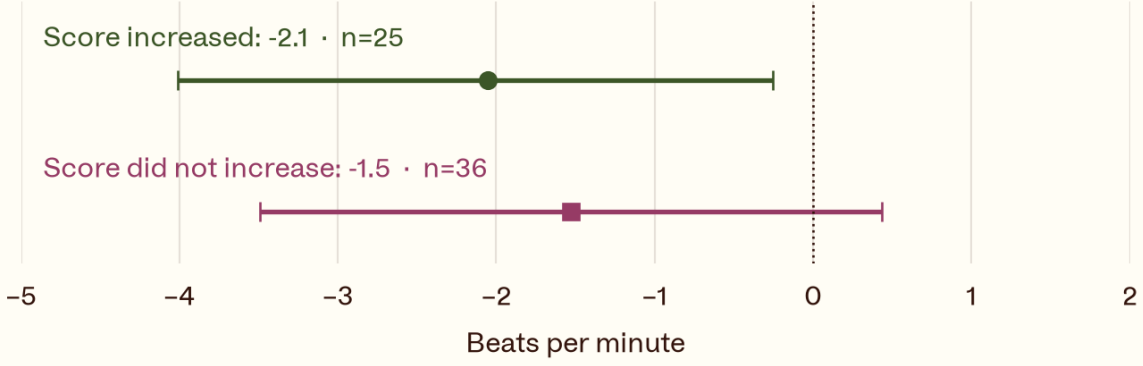
## Sleep duration




## Weight




## Walking heart rate




Changes over 12 weeks. Mean paired change and unadjusted 95% participant-bootstrap intervals. Each panel uses its own stated unit; the vertical reference is no change. Sample sizes differ by metric. The between-group intervals include zero for all three markers.

# An early signal by day 30

A separate first-month analysis considered entry dates from January 1 through August 9, 2026, with follow-up completed before September 8. It included 498 consistent food loggers, of whom 216 increased their food-only Alma Score. Among score improvers with paired measurements, resting heart rate fell 1.55 bpm (n=85; 95% CI, -2.33 to -0.83). Recorded sleep increased 11.9 minutes per night (n=73; 95% CI, 1.8 to 22.0), and weight fell 0.77 kg (n=35; 95% CI, -1.32 to -0.18). These changes were already visible in the final week of the first month.

Resting heart rate provides the clearest early finding. Its decline meets the 5% false-discovery-rate threshold across the 16 paired 30-day markers (q=0.002) and across the 104 new primary tests in this extension (q=0.014). In other active Alma users, resting heart rate was nearly unchanged (-0.04 bpm, n=118). The between-group difference was -1.50 bpm (95% CI, -2.49 to -0.57); it met the within-horizon comparison threshold (q=0.039), but not the broader extension threshold (q=0.128).

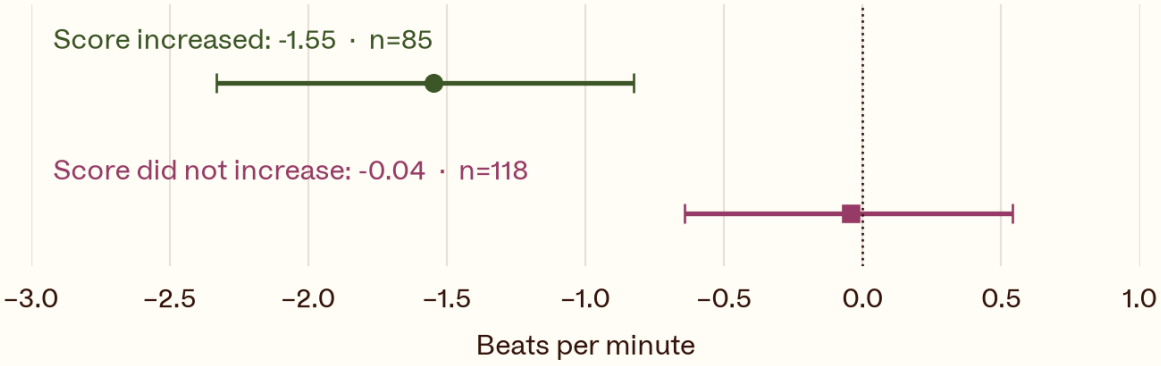
Other active users also recorded more sleep (+7.7 minutes, n=100) and lower weight (-0.41 kg, n=49). The sleep and weight differences between groups were uncertain. Walking heart rate changed by -1.18 bpm among 52 score improvers, with an interval spanning no change. The first-month signal concerns resting heart rate; the 12-week highlighted measure is walking heart rate. They are different measurements.

Day 1-7 averages were compared with day 24-30 averages. Food logging had to meet the substantial-day criteria on at least three days in each of four blocks: days 1-7, 8-14, 15-23 and 24-30. Daily health measures required at least three observed days per endpoint week, and intermittent measures required two. This includes people without 12 weeks of continued logging. The first week is already after entry, so the analysis does not measure the immediate effect of starting Alma.

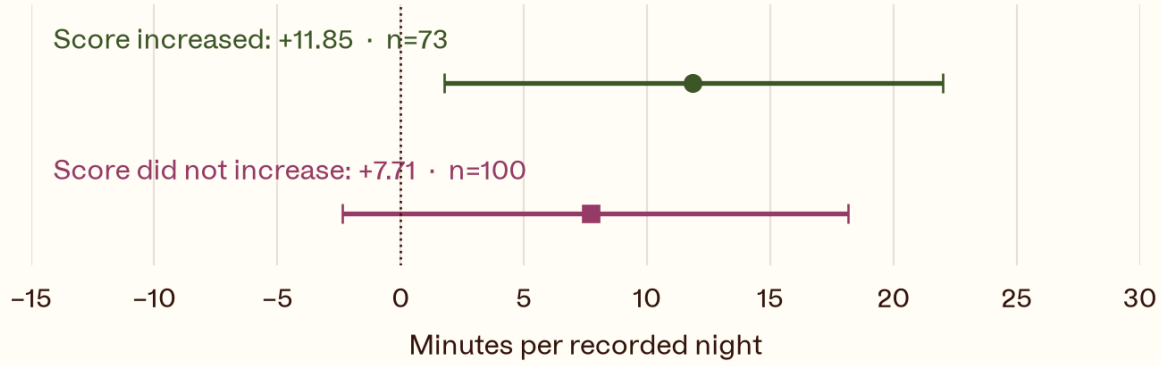
| Measurement | Paired n | Baseline | Endpoint | Change (95% CI) |
|---|---|---|---|---|
| Resting heart rate (bpm) | 85 | 59.57 | 58.02 | -1.55 (-2.33 to -0.83) |
| Sleep duration (hr) | 73 | 6.78 | 6.98 | 0.20 (0.03 to 0.37) |
| Weight (kg) | 35 | 88.43 | 87.66 | -0.77 (-1.32 to -0.18) |

# Changes by day 30

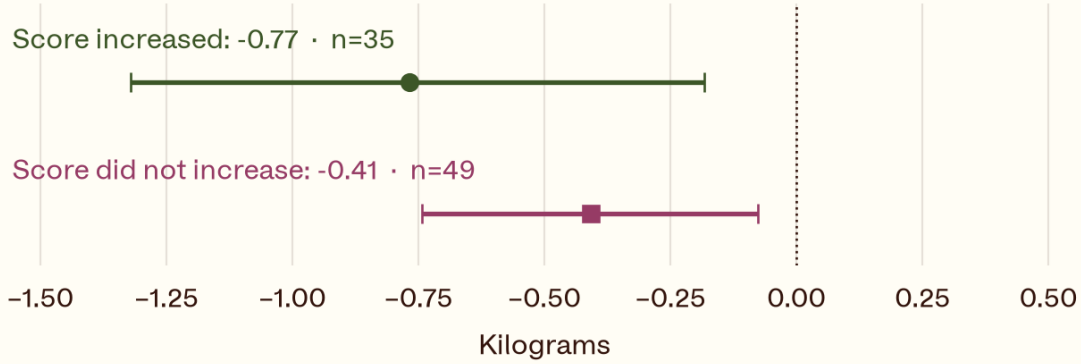
## Resting heart rate




## Sleep duration




## Weight



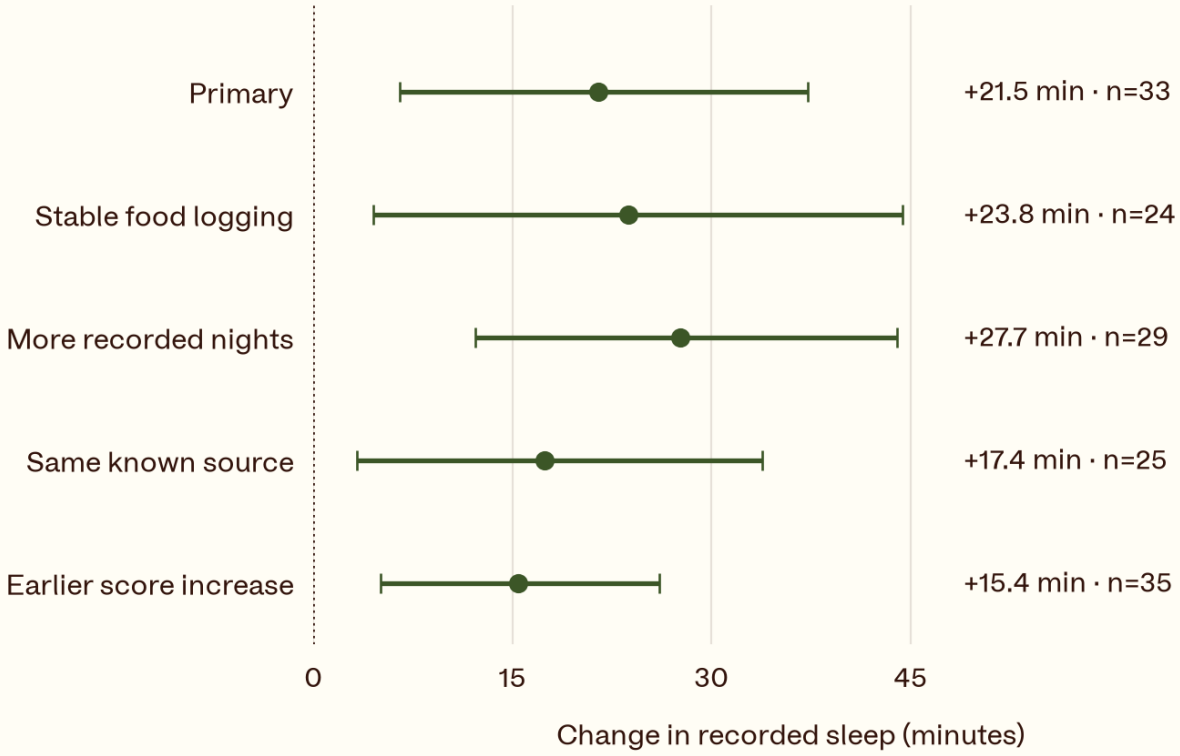

Changes visible in the first month. First-week averages versus averages in days 24-30. Score improvers and other active Alma users, with mean paired changes and unadjusted 95% bootstrap intervals. Resting heart rate is distinct from the walking heart rate highlighted at 12 weeks. These are different samples, not points on one continuous cohort trajectory.

# How sensitive are the findings

The sleep increase remained positive when food logging was comparatively stable (+23.8 minutes, n=24), when at least 14 sleep days were required in each window (+27.7 minutes, n=29), and when both windows used the same known source identifier (+17.4 minutes, n=25). These overlapping analyses are robustness checks, not independent replications. Source identifiers identify software or recording sources, not necessarily an unchanged physical device.

A temporal sensitivity selected score improvement by weeks 5-8, before the final health-measurement window. In that group, weight fell 1.48 kg (n=38; 95% CI, -2.26 to -0.82; within-family q=0.007). However, the weight-change difference versus the corresponding non-improvers was +0.10 kg (95% CI, -0.93 to +1.12). This supports a weight-change observation within the group, not a score-specific benefit. The secondary analysis followed the primary analysis; its q value does not correct across every analysis in this investigation.

Requiring a five-point score increase left only 22 people, with fewer than ten paired observations for each highlighted 12-week health marker. Those effects are suppressed. The broader time-horizon extension below evaluates continuous score changes and the same five-point sensitivity.




Sleep duration across recording-quality checks. Mean change in minutes per recorded night, with unadjusted 95% bootstrap intervals. The samples overlap. The last row selects improvement in weeks 5-8, before the final measurement window. These are sensitivity analyses, not independent replications.

# What longer follow-up adds

The six-month check required 168 elapsed days and at least ten substantial food-log days in each of six 28-day blocks. Fifty people qualified, including 15 whose score increased. None of those score improvers had sufficient paired sleep or heart-rate measurements, and fewer than ten had paired weight records. Longer food histories did not supply enough linked health history for a score-improver outcome estimate.

Across all six-month loggers with paired weight records, regardless of score direction, mean weight changed by -4.01 kg (n=13; 95% CI, -7.43 to -1.02). This is a small descriptive sample, not a continuation curve for the 30-day or 12-week samples. It is not used to claim a score-related benefit.

# Does a larger score rise track a larger health change

The extension also examined score change continuously across all paired participants, rather than splitting only at zero. No rank correlation or adjusted score-change coefficient met the within-horizon or combined extension threshold. A larger score increase was not consistently associated with a larger health-measurement change in these data.

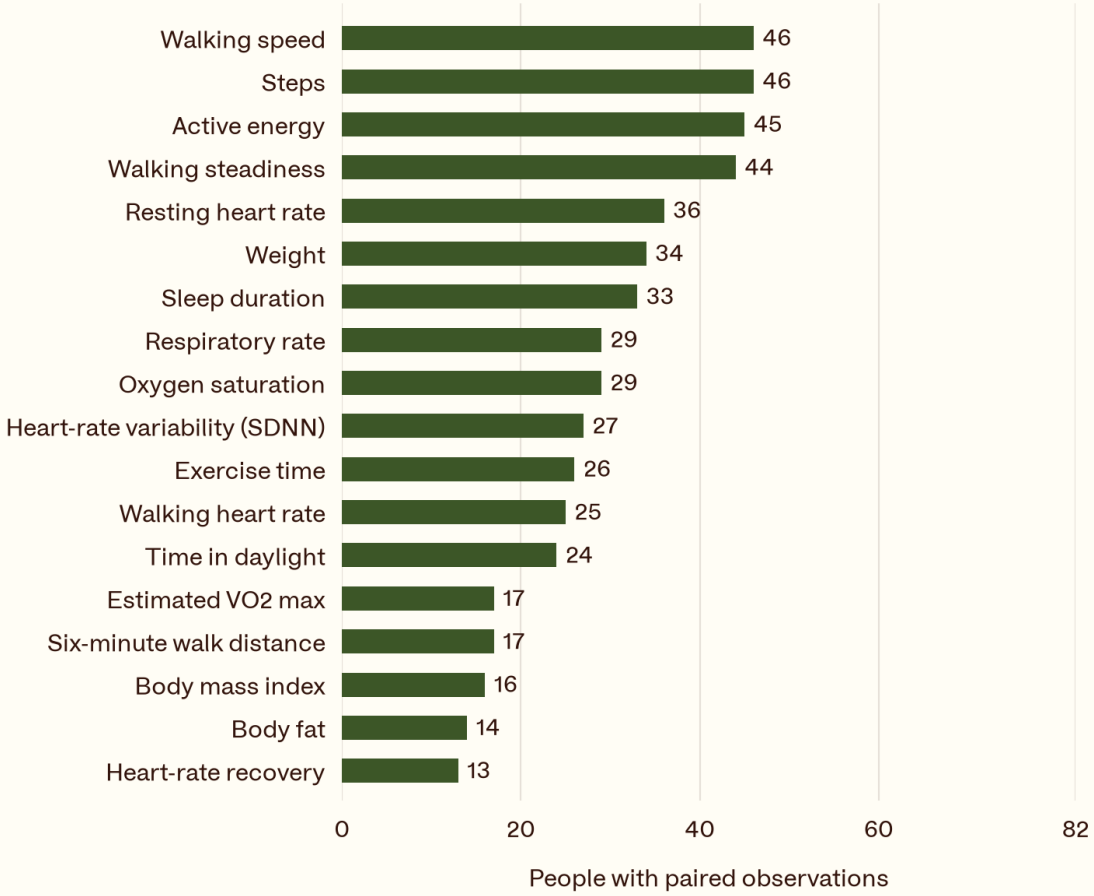
At 30 days, the adjusted resting-heart-rate estimate was -0.31 bpm per five score points (n=203; 95% CI, -0.79 to +0.16). This model accounts for starting score, starting resting heart rate and entry month. It does not confirm a graded score-response relationship. The corresponding weight model could not provide reliable HC3 uncertainty because an entry-month category contained a single observation; it is marked not estimated rather than replaced with a more favorable specification.

Rank correlations used every adequately observed marker at 30, 84 and 168 days. Adjusted models required at least 30 paired people, a full-rank design and estimable robust standard errors. All new primary tests were retained: 64 tests at 30 days, 35 continuous-association tests at 84 days and five correlations at 168 days. The combined 104-test correction does not include all earlier analyses or secondary sensitivity checks.

# Who contributed health measurements

The entry population contained 6,884 eligible accounts; 221 (3.2%) met the sustained logging rules. Of the 82 score improvers, 73 had at least one usable HealthKit record in the 84-day window. Measurement-specific sample sizes were much smaller. The analysis inventoried 34 observed measurement streams, including dietary inputs, alternate energy representations and workout records. Eighteen independent continuous markers had at least ten paired score improvers.

Daily metrics required at least seven observed days per 28-day window; intermittent metrics required at least two. The latter included weight, body composition, waist, estimated VO2 max, heart-rate recovery, six-minute walk distance and walking steadiness. The chart shows paired participation, not the number of measurements. Blood glucose and waist lacked sufficient paired coverage; no outcome estimate is presented for them.




How many score improvers had paired measurements. Participant counts for every adequately observed primary independent marker, out of 82 score improvers. Intermittent metrics require two observed days per window; daily metrics require seven. Sparse and descriptive streams are listed in the complete data inventory.

# Study design and methods

This company-run retrospective study used records extracted on September 9, 2026. Entry was the first retained local food-log day between January 1 and June 15, 2026, among accounts created in that interval. Known minors and internal or test accounts were excluded. Unknown age was permitted by the entry rule; no unknown-age participant met the sustained cohort criteria in the prior score audit. No research-specific consent or ethics determination was established in this analysis, and no peer review has occurred.

A substantial logged day required at least two meals, 800-5,000 recorded kcal, and usable scores for at least 90% of completed food items. Sustained logging required ten such days in each 28-day block over 84 days. Day 0 is entry; baseline is days 0-27 and the endpoint is days 56-83. Score improvement is any increase in the participant's mean food-only score between those windows. This is an analytic definition, not a validated clinical threshold. Available telemetry did not establish repeated use of score guidance, so the group is described as score improvers rather than adherent users.

The food-only score sums retained category contributions using the backend's ingredient-versus-parent authority rules, caps each of ten daily categories between -10 and +10, adds 40, and caps and rounds the total to 0-100. Supplements and exercise-related sodium credits are excluded. Excluding exercise prevents activity data from directly determining both the score and an activity outcome. Retained food estimates may reflect historical processing changes; the reconstructed score is not a versioned archive of what appeared on every screen.

Only records marked healthkit entered the primary health analysis. Manual and other sources were excluded. Pounds were converted to kilograms. Broad plausibility bounds, unit checks, row exclusions and source code are supplied with the aggregate download. Values outside the bounds were discarded; missing days were never imputed as zero. Every participant contributes equally through a window-level mean, regardless of measurement frequency.

Paired changes use 10,000 participant bootstrap resamples with a fixed seed and percentile 95% intervals. Paired t tests supplied two-sided p values; Benjamini-Hochberg adjustment covered all 18 sufficiently observed primary independent markers. Between-group changes use independent participant bootstrap resampling, Welch tests and a separate 18-test adjustment. These intervals are not simultaneous intervals, and the study is exploratory rather than preregistered. The original 12-week comparisons were unadjusted for confounders. The extension's continuous-association models adjust for starting score, starting marker and entry month; medication, illness and other concurrent changes remain unmeasured.

# What the stored measurements represent

Most non-cumulative HealthKit metrics use the most recent available sample per local day, not an average over every sample. Thus the HRV result concerns daily retained SDNN samples, and blood glucose would not represent fasting glucose or continuous-glucose time in range. Apple's SDNN definition is documented separately [4]. Estimated VO2 max is a device estimate, not a laboratory exercise test; Apple describes its estimation conditions and limitations [5].

Sleep totals combine recorded asleep intervals into sessions and assign them to the wake day. Cumulative metrics such as steps and exercise minutes omit days with no positive quantity in the current sync implementation. Their results therefore describe recorded days and cannot distinguish true zero activity from missing observations. Weight synchronization reads the latest available measurement, so server history is incomplete and source-device identity is not retained.

The daily-metric table was introduced in May 2026; a 90-day client lookback can backfill earlier measurements. Retained records therefore do not establish continuous connection to Alma throughout follow-up. Aggregate daily values can be updated by later synchronization, and historical code or device changes may affect comparability. A healthkit source label does not establish that every measurement was sensor-generated or clinically validated.

Nutrients written to Apple Health can overlap with food logs and are excluded from the independent marker family. Alternate active-energy totals and recorded workout duration are displayed as descriptive streams, not separate confirmations of benefit. Body temperature, reproductive categories and UV exposure have no universal favorable direction. Weight loss itself is not automatically a favorable outcome without information about intent and health context.

# Where AHEI fits

The Alternate Healthy Eating Index was developed to study dietary patterns and health outcomes [1]. The updated AHEI-2010 was evaluated against major chronic disease in long-running cohorts [2], and subsequent work examined changes in diet quality and mortality [3]. That literature motivates the hypothesis that improving dietary quality may matter for health.

Alma Score adapts dietary-quality concepts into a proprietary ten-category 0-100 score. AHEI-2010 has eleven components and a different scoring system. The original index’s evidence does not transfer automatically to Alma’s scoring algorithm, food estimates or coaching. Correspondence with Harvard-affiliated researchers establishes research interest, not completed validation, institutional endorsement or authorship of this company analysis.

# What this study supports

The study identifies changes beyond the food log within 30 days and over 12 weeks. The clearest early finding is a resting-heart-rate decline among first-month score improvers, which survives the extension's multiple-testing correction. At 12 weeks, longer sleep appears across several recording-quality checks, alongside lower weight and walking heart rate. These observations support prospective evaluation of the measured outcomes; they do not establish that changing Alma Score causes them.

The parent 12-week cohort's mean score declined from 66.4 to 64.5; the score-improver findings describe a selected subgroup. Sustained loggers are a small minority of entrants, and HealthKit availability adds another filter. Selecting score improvers after observing their scores invites regression to the mean; the groups also differ at baseline. For example, paired VO2-max samples start at 41.7 versus 34.6 mL/kg/min. Concurrent exercise, medications, illness, weight-loss intent, seasonal changes and measurement practices remain plausible explanations. Narrow confidence intervals within a selected group do not remove these biases.

A prospective study should specify the exposure, primary endpoint and analysis in advance; document consent and ethics review; capture score-feedback engagement and stable measurement provenance; and assess dietary quality independently. A randomized score-feedback intervention with a suitable comparison would address whether Alma causes benefit. The present data can inform recruitment and measurement feasibility, but its selected effect sizes should not be treated as unbiased inputs for an efficacy claim.

# Transparency and reproducibility

Sponsor and analysis author: Alma. This draft was prepared with AI-assisted analysis, programming and writing. It has not undergone independent scientific review. The earlier score-only paper was withdrawn from the local website at the requester's direction; its numerical audit remains internal. The subgroup question was chosen after that analysis, and the HealthKit analysis rules were written before inspecting paired HealthKit changes. This chronology is disclosed to distinguish exploratory work from a prospectively specified efficacy study.

Only aggregate outputs are included here. Identifiers were joined inside a read-only production database, and the temporary analytical export used arbitrary study keys and relative days. No individual trajectories, names, contact details, device identifiers or exact participant dates are published. Numerical outcomes below ten participants are suppressed. An independent pandas reconstruction checked 281 numerical conditions, and eight focused analytical tests passed. The analysis plan, aggregate results and computation scripts accompany this draft.

The 30-day, six-month and continuous-association analyses were added after reviewing the original 12-week results. Their rules were recorded before inspecting the new outcome estimates. A fresh read-only extraction reproduced the 221-person cohort and all 82 score improvers. Retained HealthKit rows had minor snapshot differences: the 12-week sleep estimate was +21.6 rather than +21.5 minutes. The original primary estimates remain frozen; the extension download records the newer snapshot and all attempted analyses. An independent reconstruction checked 1,138 additional numerical conditions, and all 14 analytical tests passed.

# Complete 12-week marker results

Score improvers. Changes and intervals are in the stated units. Intervals are unadjusted; q values cover the 18 primary paired tests. No result meets q < 0.05. Sparse outcomes are suppressed.

| Measurement | Paired n | Baseline | Endpoint | Change (95% CI) | q |
|---|---|---|---|---|---|
| Active energy (kcal) | 45 | 627.32 | 602.57 | -24.75 (-62.24 to 12.15) | 0.631 |
| Walking steadiness (%) | 44 | 91.79 | 90.30 | -1.49 (-2.97 to -0.21) | 0.225 |
| Blood glucose (mg/dL) | 0 | Unavailable | Unavailable | Insufficient paired data | Not estimated |
| Body fat (%) | 14 | 20.28 | 20.23 | -0.06 (-0.26 to 0.17) | 0.826 |
| Body mass index (BMI) | 16 | 26.09 | 26.01 | -0.07 (-0.25 to 0.10) | 0.819 |
| Exercise time (min) | 26 | 70.95 | 72.01 | 1.07 (-5.23 to 7.83) | 0.847 |
| Heart-rate recovery (bpm) | 13 | 37.12 | 35.99 | -1.13 (-3.63 to 1.91) | 0.819 |
| Heart-rate variability (SDNN) (ms) | 27 | 51.90 | 50.80 | -1.10 (-5.46 to 3.27) | 0.826 |

| Measurement | Paired n | Baseline | Endpoint | Change (95% CI) | q |
|---|---|---|---|---|---|
| Lean mass (kg) | Fewer than 10 | Unavailable | Unavailable | Insufficient paired data | Not estimated |
| Oxygen saturation (%) | 29 | 96.28 | 96.27 | -0.01 (-0.31 to 0.27) | 0.930 |
| Respiratory rate (br/min) | 29 | 15.65 | 15.67 | 0.01 (-0.29 to 0.26) | 0.930 |
| Resting heart rate (bpm) | 36 | 58.40 | 58.21 | -0.19 (-1.32 to 0.93) | 0.847 |
| Six-minute walk distance (m) | 17 | 495.51 | 489.09 | -6.43 (-17.04 to 0.00) | 0.637 |
| Sleep duration (hr) | 33 | 6.68 | 7.04 | 0.36 (0.11 to 0.62) | 0.103 |
| Steps (steps) | 46 | 9542.07 | 9745.99 | 203.93 (-359.80 to 801.25) | 0.819 |
| Time in daylight (min) | 24 | 93.60 | 106.20 | 12.60 (-8.96 to 35.87) | 0.638 |
| Estimated VO2 max (mL/kg/min) | 17 | 41.69 | 42.70 | 1.02 (-0.09 to 2.29) | 0.454 |
| Waist circumference (cm) | 0 | Unavailable | Unavailable | Insufficient paired data | Not estimated |
| Walking heart rate (bpm) | 25 | 94.52 | 92.47 | -2.05 (-4.01 to -0.26) | 0.225 |
| Walking speed (m/s) | 46 | 1.09 | 1.08 | -0.01 (-0.03 to 0.02) | 0.826 |
| Weight (kg) | 34 | 79.90 | 78.65 | -1.25 (-2.17 to -0.49) | 0.103 |
| Workout effort (/10) | Fewer than 10 | Unavailable | Unavailable | Insufficient paired data | Not estimated |

# The comparison groups

Change among participants without score improvement, and the score-improver change minus the non-improver change. Positive differences do not always mean favorable outcomes.

| Measurement | Paired n | Non-improver change (95% CI) | Difference in changes (95% CI) |
|---|---|---|---|
| Active energy (kcal) | 81 | -18.41 (-46.34 to 9.93) | -6.34 (-55.08 to 41.20) |
| Walking steadiness (%) | 73 | -1.48 (-2.45 to -0.54) | -0.01 (-1.81 to 1.60) |
| Blood glucose (mg/dL) | 0 | Insufficient paired data | Insufficient paired data |
| Body fat (%) | 28 | -0.82 (-1.33 to -0.41) | 0.77 (0.30 to 1.35) |
| Body mass index (BMI) | 29 | -0.24 (-0.37 to -0.11) | 0.16 (-0.06 to 0.38) |
| Exercise time (min) | 43 | -4.06 (-8.08 to -0.10) | 5.13 (-2.09 to 12.65) |
| Heart-rate recovery (bpm) | 13 | 2.05 (-0.31 to 4.12) | -3.19 (-6.54 to 0.49) |
| Heart-rate variability (SDNN) (ms) | 37 | -0.59 (-2.74 to 1.64) | -0.51 (-5.33 to 4.29) |
| Lean mass (kg) | 22 | -0.05 (-0.43 to 0.47) | Insufficient paired data |
| Oxygen saturation (%) | 37 | 0.00 (-0.20 to 0.20) | -0.02 (-0.37 to 0.32) |
| Respiratory rate (br/min) | 30 | 0.29 (-0.01 to 0.62) | -0.27 (-0.71 to 0.13) |
| Resting heart rate (bpm) | 57 | 0.23 (-0.43 to 0.88) | -0.42 (-1.71 to 0.88) |
| Six-minute walk distance (m) | 27 | -0.35 (-3.62 to 2.72) | -6.07 (-18.29 to 2.02) |
| Sleep duration (hr) | 46 | 0.28 (0.06 to 0.50) | 0.08 (-0.25 to 0.43) |
| Steps (steps) | 81 | -442.25 (-965.84 to 45.60) | 646.18 (-119.29 to 1422.73) |
| Time in daylight (min) | 32 | 23.21 (7.46 to 39.59) | -10.60 (-36.96 to 16.98) |
| Estimated VO2 max (mL/kg/min) | 24 | 0.62 (-0.34 to 1.62) | 0.39 (-1.10 to 1.99) |

| Measurement | Paired n | Non-improver change (95% CI) | Difference in changes (95% CI) |
|---|---|---|---|
| Waist circumference (cm) | 0 | Insufficient paired data | Insufficient paired data |
| Walking heart rate (bpm) | 36 | -1.53 (-3.49 to 0.44) | -0.53 (-3.32 to 2.25) |
| Walking speed (m/s) | 78 | -0.01 (-0.02 to 0.01) | 0.00 (-0.03 to 0.03) |
| Weight (kg) | 52 | -1.73 (-2.39 to -1.12) | 0.47 (-0.59 to 1.50) |
| Workout effort (/10) | Fewer than 10 | Insufficient paired data | Insufficient paired data |

# Additional measurement streams

These descriptive measurements and dietary inputs do not independently establish health benefit. Workout values concern recorded workout days only. Alternate energy representations are not separate replications.

| Measurement | Paired n | Change (95% CI) |
|---|---|---|
| Activity ring energy (kcal) | 30 | -31.17 (-70.72 to 6.68) |
| Sleeping wrist temperature (°C) | 11 | -0.02 (-0.09 to 0.05) |
| Dietary caffeine (mg) | Fewer than 10 | Insufficient paired data |
| Dietary carbohydrates (g) | 16 | -50.02 (-104.88 to -15.30) |
| Dietary energy consumed (kcal) | 17 | -235.92 (-504.68 to -18.31) |
| Dietary fat total (g) | 16 | -5.76 (-10.26 to -1.34) |
| Dietary protein (g) | 16 | -3.75 (-14.94 to 7.88) |
| Dietary sugar (g) | 0 | Insufficient paired data |
| Dietary water (mL) | Fewer than 10 | Insufficient paired data |
| Hourly active energy total (kcal) | 63 | 19.03 (-9.80 to 47.30) |
| UV exposure (UVI) | Fewer than 10 | Insufficient paired data |
| Recorded workout duration (min) | 46 | 20.33 (4.59 to 41.76) |

# Complete 30-day marker results

First-week averages versus days 24-30 among score improvers. Intervals are unadjusted. The q values below cover all 104 new primary extension tests, including comparisons and continuous score associations. Resting heart rate meets q < 0.05.

| Measurement | Paired n | Change (95% CI) | Extension q |
|---|---|---|---|
| Active energy (kcal) | 118 | 22.04 (-15.55 to 60.18) | 0.756 |
| Walking steadiness (%) | Fewer than 10 | Insufficient paired data | Not estimated |
| Blood glucose (mg/dL) | 0 | Insufficient paired data | Not estimated |
| Body fat (%) | 22 | -0.09 (-0.40 to 0.27) | 0.827 |
| Body mass index (BMI) | 28 | -0.16 (-0.31 to -0.03) | 0.332 |
| Exercise time (min) | 63 | -1.26 (-13.02 to 9.00) | 0.937 |
| Heart-rate recovery (bpm) | Fewer than 10 | Insufficient paired data | Not estimated |
| Heart-rate variability (SDNN) (ms) | 55 | 4.35 (1.32 to 7.36) | 0.274 |
| Lean mass (kg) | 17 | -0.30 (-0.78 to 0.09) | 0.756 |
| Oxygen saturation (%) | 49 | -0.17 (-0.42 to 0.08) | 0.756 |
| Respiratory rate (br/min) | 57 | 0.03 (-0.17 to 0.25) | 0.935 |
| Resting heart rate (bpm) | 85 | -1.55 (-2.33 to -0.83) | 0.014 |
| Six-minute walk distance (m) | 0 | Insufficient paired data | Not estimated |
| Sleep duration (hr) | 73 | 0.20 (0.03 to 0.37) | 0.332 |
| Steps (steps) | 121 | 448.27 (-91.37 to 978.08) | 0.638 |
| Time in daylight (min) | 43 | 8.72 (-13.09 to 30.96) | 0.794 |

| Measurement | Paired n | Change (95% CI) | Extension q |
|---|---|---|---|
| Estimated VO2 max (mL/kg/min) | 11 | 1.25 (0.21 to 2.36) | 0.375 |
| Waist circumference (cm) | 0 | Insufficient paired data | Not estimated |
| Walking heart rate (bpm) | 52 | -1.18 (-3.15 to 0.94) | 0.756 |
| Walking speed (m/s) | 116 | -0.01 (-0.03 to 0.01) | 0.767 |
| Weight (kg) | 35 | -0.77 (-1.32 to -0.18) | 0.307 |
| Workout effort (/10) | Fewer than 10 | Insufficient paired data | Not estimated |

# Full extension tables and models

The website and extension JSON contain every 30-, 84- and 168-day comparison, rank correlation, adjusted model, all-participant estimate and sensitivity. Outcomes below ten participants are suppressed. The 30-day resting-heart-rate difference is -1.50 bpm (95% CI, -2.49 to -0.57), with q=0.039 in the 16-test horizon comparison family and q=0.128 across the extension. No continuous score-change association met its correction threshold. The 168-day score-improver sample lacked sufficient paired coverage for any reported marker.

# References

1. McCullough et al. Dietary quality and major chronic disease risk in men and women (2002).

2. Chiuve et al. Alternative dietary indices both strongly predict risk of chronic disease (2012).

3. Sotos-Prieto et al. Association of changes in diet quality with total and cause-specific mortality (2017).

4. Apple Developer. HealthKit heart-rate variability SDNN.

5. Apple. Using Apple Watch to estimate cardio fitness with VO2 max.

6. STROBE. Reporting guidance for observational research.

# Analysis files

The accompanying aggregate JSON contains every reported comparison and sensitivity estimate. The analysis ZIP contains the frozen analysis plan, cleaning bounds, computation scripts, focused tests and numerical validation. Individual source records are not published.

Website route: alma.food/research/alma-score-health-signals

The article, aggregate results and analysis code are available at the website link above.