A Work Project, presented as part of the requirements for the Award of a Master’s degree in Management from the Nova School of Business and Economics.

# PRIORITISING HEALTH MARKERS FROM WEARABLES FOR WORKING ADULTS: SLEEP, ACTIVITY AND HEART RATE

YAN BECK

Work project carried out under the supervision of:

João Castro

12/12/2025

**Abstract**

Consumer wearables measure many metrics, but it is unclear which ones should be prioritised for health decisions. This thesis focuses on six markers in non-shift working adults: sleep duration and regularity, resting heart rate, heart rate variability, daily steps, and $VO_2$max. A rapid scoping review and an online survey of working adults inform a tiering framework. Sleep duration, daily steps and resting heart rate emerge as Tier 1 markers, while sleep regularity, heart rate variability and $VO_2$max are placed in Tier 2. The thesis translates these findings into concise, device-agnostic operational guidance.

**Keywords:** consumer wearables, health markers, working adults, sleep duration, sleep regularity, resting heart rate, heart rate variability, steps, $VO_2$max

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

## 1. Introduction

Consumer wearables like smartwatches or rings have made continuous, cheap and low-effort measurements of sleep, activity, and heart rate based signals possible in daily life. With over half of the population in many countries owning wearable devices like watches or rings (Lambe et al. 2025, 2), reliability and validity of these measurements to support in personal health decisions is important. This data creates an opportunity to observe patterns relevant to personal health without clinic visits or specialized, expensive equipment. At the same time, the sheer amount of device outputs, uneven accuracy, and frequent algorithm changes make it unclear which markers actually matter for health and also which ones are even measurable with sufficient reliability to inform personal tracking.

This thesis addresses that practical problem by pairing outcome evidence (associations with major health outcomes) with measurement evidence (validity and reliability on consumer wearable devices) for a predefined set of markers in adults aged 18-65 who do not work shifts. The aim is to clarify which markers show the strongest associations with health and how accurately consumer wearables are able to measure these markers.

A gap in the literature motivates this approach as reviews often assess associations between markers and health outcomes without considering whether those same markers are accurately measurable on consumer wearables in normal life settings. Conversely, validation studies focus on the accuracy of measurements for a broad set of metrics. Thus, leaving the task of integrating measurement accuracy and health relevance, and therefore also which metrics to prioritise, to works like this thesis.

Also in general, measurement coverage remains limited. Fewer than 5% of consumer wearables have been validated across all the outcomes they report (Lambe et al. 2025, 3). For illustration, only one study has evaluated the $VO_2$max estimation of the Apple Watch to date (Lambe et al.

2025, 3). Furthermore, few syntheses bridge both evidence streams for each marker and then translate the results into operational guidances for non-clinical contexts. This thesis is designed to help close that gap.

The work asks two questions. First, which markers measured by wearables show the strongest associations with health in working adults? Second, which of those markers are even feasible to be measured with sufficient accuracy to justify interpreting trends in everyday life and base health decisions on them? Markers were prespecified as sleep duration, sleep regularity, resting heart rate, heart-rate variability, steps, and cardiorespiratory fitness ($VO_2max$). The outcomes were restricted to public-health relevance. Mainly, all-cause mortality, cardiometabolic events, and mental health.

Methodologically, this thesis separates evidence into two streams per marker. Outcome (importance) and measurement (feasibility). Study selection and reporting were guided by PRISMA principles (Tricco et al. 2018), using a simplified flow diagram adapted for this rapid scoping review. Searches were conducted in PubMed within a defined time window. Exact search strings and filters, the screening log with decisions and reasons, and the final evidence set are provided in the appendices for transparency. Results present the simplified PRISMA flow and per-marker findings for both streams. Discussion derives operational implications and provides a simple prioritization (tiering) based on outcome evidence, measurability on consumer devices, consistency across studies/devices, and practical actionability.

The contributions of this thesis are threefold. First, it offers a integrated, marker-wise synthesis that pairs health outcome relevance with real life measurability. Second, it develops a tiering framework that identifies a minimal marker set for daily use and positions the remaining markers as a lower tier, indicating optional and less frequent monitoring. Third, it translates this framework into concise, device-agnostic operational implications for each marker.

## 2. Methods

This thesis examined which health markers measurable with consumer wearables in working adults should be prioritised, based on the combination of their associations with major health outcomes and the reliability with which wearables capture them in everyday conditions.

In total six health markers were prespecified: sleep duration, sleep regularity, resting heart rate, heart rate variability, steps, and cardiorespiratory fitness ($VO_2$max).

The health markers were chosen based on: (1) biological plausibility and population wide use (2) measurability on consumer wearables and (3) practical actionability for users.

"Health" in this thesis was looked at through three major health outcomes: all-cause mortality, cardiometabolic events, and mental health.

A rapid, marker-wise scoping review was conducted with two streams per marker:

- Outcome (importance): association of the marker with the prespecified health outcomes.
- Measurement (feasibility): validity/reliability of consumer wearables to capture the marker reliably under real-life conditions.

Searches were run in PubMed from 14-17 October 2025. A simplified PRISMA flow diagram summarizing study selection is presented in the results chapter. The target population was adults aged 18-65 in non-shift employment.

Eligible study designs for the outcome stream were umbrella reviews, systematic reviews and meta-analyses. Eligible designs for the measurement stream were validation studies comparing consumer wearables with reference standards. Polysomnography for sleep duration/timing, electrocardiography for RHR/HRV, StepWatch or manual counting for steps and indirect calorimetry for $VO_2$max/CRF. Free-living studies were preferred but lab studies were accepted as upper bound accuracy.

Common exclusions were exercise HR instead of resting HR, HRV obtained during exercise and stress tests, wearable versus wearable comparisons without a reference standard, device marketing/engineering reports without human validation, shift-worker populations.

As result sets for the search queries were large, a prespecified rapid scoping heuristic was applied before screening: per query, the top 30 “Best Match” results plus the 15 “Most Recent” results were looked at. Screening proceeded in two stages: title/abstract and then full text. Full-text exclusions required a recorded reason. The PRISMA flow in the Results chapter reports the selection counts and the complete search strategies, screening log with decisions/reasons, and the final evidence set (anchors with DOIs) are provided in Appendices A-C.

To avoid redundancy, 1-3 anchor studies per stream per marker were retained (the most recent and relevant umbrella/systematic review for outcome, the most recent and relevant validation for measurement).

Study characteristics and key results were extracted into marker wise outcome and measurement evidence tables. For each marker, a short narrative synthesis summarised outcome, measurement, and operational implications. Markers were then assigned to tiers using a four-axis rubric (chapter 4.2).

Data was collected via a cross-sectional, self-administered online survey (Google Forms) of working adults. Participation was voluntary and anonymous. No names, email addresses, or IP addresses were recorded and branching logic hid interpretation items from non-wearers. Recruitment used convenience sampling, and data were exported to a spreadsheet for analysis. The data collection stopped  on 9 December 2025. As this survey was not designed to offer statistical validation but rather descriptive context, the findings of the survey are stated and discussed in chapter 4 to enable immediate comparison with the literature.

Survey responses were cleaned in a stepwise manner. First, all raw submissions from Google Forms were exported (N = 133). Entries without informed consent (n = 3) were removed. Second, an attention-check was used to identify careless responding. Submissions with an incorrect response to this question (n = 12) were excluded. Third, respondents who reported working regular night-shifts were removed as they fell outside the target population of non-shift working adults (n = 33).

These steps yielded cohort A (N = 85), which served as the main analytic sample for descriptive statistics and importance plots. Within cohort A, cohort B (N = 26) was defined as the subset of respondents who reported owning a wearable device and was used for analyses of marker adoption and usage.

## 3. Results

A total of 3274 records were identified from the database search and 3 via citation chasing. After prespecified sub-sampling of large result sets (2787 removed before screening), 490 titles/abstracts (487 from the database plus 3 from citation chasing) were screened, with 469 excluded at this stage. 21 reports were assessed in full text, 6 were excluded with reasons, and 15 studies were included in the final evidence set (Figure 1).

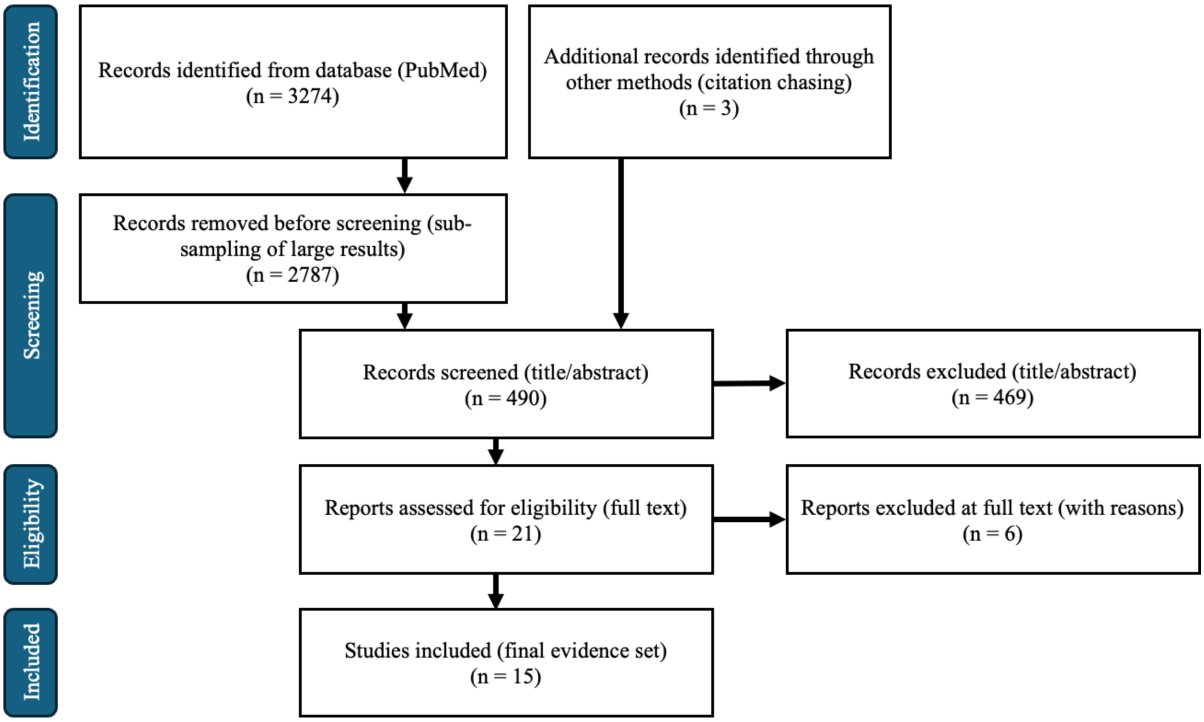



*Figure 1: PRISMA flow diagram of study selection*

**Table 1**: Overview of outcome and measurement results

| Markers | Outcome anchors | Outcome findings | Measurement anchor | Measurement findings |
|---|---|---|---|---|
| Sleep duration | Ungvari et al. 2025; Li et al. 2022; Chaput et al. 2020a | 7-8 h most favorable; increased risk at short/long extremes. | Kainec et al. 2024 | Newer wearables estimate sleep duration with acceptable accuracy. |
| Sleep regularity | Chaput et al. 2020b | Greater day to day sleep variability relates to worse health; no firm targets established yet. | Labie et al. 2025 | Wearables capture timing/regularity accurately; sleep stages often misclassified. |
| Resting heart rate | Zhang et al. 2016 | Higher RHR relates to higher all-cause and cardiovascular mortality. | Dial et al. 2025 | Nocturnal RHR readings are most accurate. |
| Heart rate variability | Jarczok et al. 2022 | Lower HRV relates to higher risk; higher HRV favorable. | Georgiou et al. 2018 | HRV reasonable at rest/night; diminishes with movement and intense activity. |
| Steps | Ding et al. 2025; Bizzozero-Peroni et al. 2024 | More daily steps relate to lower mortality risk and less depression; big gains at ≈7000 and diminishing returns at ≈10000 steps per day | Höchsmann et al. 2020 | Total steps can be underestimated by ~20%. |
| $VO_2max$ / CRF | Han et al. 2022 | Higher $VO_2max$ inversely associated with all-cause and cardiovascular mortality. | Lambe et al. 2025 | Wearables underestimate $VO_2max$; it is a notable error but adequate for within-person trends. |

### 3.1 Outcome results

**Sleep duration.** Short (<7 h) and long (>9 h) sleep durations are linked to higher all-cause mortality versus 7-8 h (Ungvari et al. 2025, 4545). Furthermore, these inappropriate sleep durations may significantly elevate the risk for other adverse health outcomes like cardiovascular disease, cognitive decline and depression (Li et al. 2022, 1495). Notably, a 1-h reduction per 24 hours is associated with 3-11% increased risk of all-cause mortality and coronary heart disease in short sleepers. On the other hand, a 1-h increment in long sleepers is associated with a 7-17% higher risk of stroke mortality and coronary heart disease (Li et al. 2022, 1479). Overall, 7-8 hours per day appears most favorable (Chaput et al. 2020a, 230).

**Sleep regularity.** A greater day to day variability in the timeframe a person goes to sleep is associated with worse health outcomes (Chaput et al. 2020b, 233). However, the currently available evidence on sleep regularity is generally low to moderate regarding quality and lacks the ability to support quantifiable targets (Chaput et al. 2020b, 242). Nonetheless, the evidence still suggests that consistent sleep and wake-up times are associated with good health outcomes (Chaput et al. 2020b, 232) and should therefore be encouraged for healthy sleep (Chaput et al. 2020b, 243).

**Resting Heart Rate.** It is a good predictor of all-cause and cardiovascular mortality (Dial et al. 2025, 2). Higher resting heart rate shows a mostly linear association with adverse health outcomes as each 10-bpm increase (compared to 45 beats per minute) is linked to a 9% and 8% higher risks of all-cause and cardiovascular mortality (Zhang et al. 2016, 60). Increased resting heart rate is therefore associated with a higher risk regarding all-cause and cardiovascular mortality (Zhang et al. 2016, 53).

**Heart Rate Variability.** Lower heart rate variability is a good predictor of mortality as lower values correspond to a higher risk across all-cause and cardiovascular health outcomes (Jarczok et al. 2022, 25). Thus, higher heart rate variability can be seen as indication of a better general health in an individual (Georgiou et al. 2018, 7).

**Steps.** A higher step count is consistently associated with a lower risk in major health outcomes like cardiovascular and all-cause mortality with decreasing risk with each additional 1000 steps. However, 7000 steps per day is associated with significant risk reductions compared to 2000 steps per day but with diminishing returns for some health outcomes above that threshold. Therefore, 7000 steps per day may be a more practical target instead of the commonly proposed 10000 steps (Ding et al. 2025, 677). For mental health, 7000 or more steps per day leads to a greatly decreased risk of depression (Bizzozero-Peroni et al. 2024, 7) up to 10000 steps a day where increasing the steps further would not decrease depressive symptoms significantly (Bizzozero-Peroni et al. 2024, 9).

**$VO_2$max**. The most common way to measure cardiorespiratory fitness is the maximal oxygen consumption ($VO_2$max) (Molina-Garcia et al. 2022, 1578). A higher cardiorespiratory fitness shows associations with reduced risk of cardiovascular and all-cause mortality (Han et al. 2022, 739). And is thus a key indicator of overall health (Lambe et al. 2025, 1). Some studies even show that it is a better predictor of cardiovascular and all-cause mortality than commonly known risk factors such as obesity and hypertension (Lambe et al. 2025, 2).

### 3.2 Measurement results

**Sleep duration.** Especially wearable devices of new generations are capable of accurately estimating sleep time within clinically acceptable levels (Kainec et al. 2024, 11). In practice, this means that consumer sleep-tracking devices are cost-effective, convenient, and accurate alternative sleep measurement tools (Kainec et al. 2024, 10).

**Sleep regularity.** Wearable devices are useful tools to monitor sleep. Especially duration, timing, and regularity (Labie et al. 2025, 15). However, wearables struggle to identify some specific sleep stages as calm wakefulness periods and deep- or REM-sleep are often misclassified as light-sleep, thus making wearables less fitting for accurate sleep architecture assessments (Labie et al. 2025, 16). These inaccuracies in sleep stage measurements could lead to false reassurance or increased anxiety regarding sleep as continuous, data-based feedback may lead to excessive daily monitoring and ultimately to orthosomnia which is the obsessive, unhealthy pursuit of perfect sleep (Labie et al. 2025, 16-17).

**Resting heart rate.** Wearables produce nocturnal resting heart rate readings that align closely with an ECG reference (Dial et al. 2025, 6) and therefore sleeping heart rate metrics from wearables can be trusted (Dial et al. 2025, 9). Differences between devices were statistically detectable but negligible, with average error well within commonly cited thresholds for RHR deviations of 5 to 7 beats per minute (Dial et al. 2025, 6-7).

**Heart rate variability.** Generally, wearables are deemed acceptable for personal trend monitoring and health feedback especially when tracked longitudinally (Dial et al. 2025, 9). However, HRV monitoring for individual trends is done best under sleeping conditions as accuracy diminishes with movement or intense activity (Georgiou et al. 2018, 17).

**Steps.** Most consumer wearables generally underestimate the number of total daily steps by around 20% due to intermittent walking bouts which are filtered out by the algorithm of the device (Höchsmann et al. 2020, 6). Also, because every day is different, monitoring several days is recommended (Höchsmann et al. 2020, 2).

**$VO_2$max.** Compared with the goldstandard of calorimetry, $VO_2$max of wearables is generally underestimated (Lambe et al. 2025, 5). Consumer wearables that use exercise tests to estimate $VO_2$max are more accurate than wearables which are using resting tests (Molina-Garcia et al. 2022, 1590). However, the random error i.e. the individual scatter is still large even when estimating with exercise tests on an individual level (Molina-Garcia et al. 2022, 1590) thus making wearables inappropriate to adequately monitor $VO_2max$ changes in sport settings (Molina-Garcia et al. 2022, 1593) or research settings but are acceptable for general fitness monitoring (Lambe et al. 2025, 8). Multiple factors like skin contact, additional weight other than bodyweight or caffeine makes accurate estimations even more difficult (Lambe et al. 2025, 2).

## 4. Discussion

### 4.1 Principal findings

Across six health markers measurable with common wearables in adults aged 18-65 who do not work shifts, the results show consistent associations between those measured markers and health.

Sleep duration follows a U-shaped pattern with the most favorable range of around 7 to 8 hours per day (Ungvari et al. 2025; Li et al. 2022; Chaput et al. 2020a). Wearable devices of newer generations are accurate in estimating sleep time (Kainec et al. 2024).

Sleep regularity indicates that greater day to day variability in timing relates to less favorable health outcomes. However, in general the evidence is low to moderate so specific numeric targets are not yet supported (Chaput et al. 2020b). Wearables accurately monitor timing and regularity but accuracy diminished regarding sleep stages (Labie et al. 2025).

Resting heart rate shows a linear relation in so far that higher levels correspond to higher risk (Zhang et al. 2016; Dial et al. 2025). Heart rate metrics measured during sleep are accurate and difference between devices is negligible (Dial et al. 2025).

Heart rate variability is lower in higher risk states and therefore a higher heart rate variability is a sign of better general health (Jarczok et al. 2022; Georgiou et al. 2018). Wearables are accurate for personal trend monitoring but measurements should be done while sleeping (Dial et al. 2025; Georgiou et al. 2018).

Steps are inversely related, resulting in lower risk with higher steps. However, benefits plateau at higher volumes (Ding et al. 2025; Bizzozero-Peroni et al. 2024). Wearables generally underestimate total daily steps by around 20% as intermittent walking bouts are automatically filtered out (Höchsmann et al. 2020).

$VO_2max$/cardiorespiratory fitness is also inversely associated with risk and therefore being a sign of good general health (Han et al. 2022; Lambe et al. 2025). Wearables generally underestimate $VO_2max$ but estimation based on exercise tests are more accurate than resting tests and are generally suitable for personal tracking (Lambe et al. 2025; Molina-Garcia et al. 2022).

Three measurement themes recur across markers.

First, nocturnal or resting measurement windows produce the most stable heart rate derived metrics, which improves the usability of RHR and HRV (Dial et al. 2025; Georgiou et al. 2018).

Second, consumer devices show systematic biases such as step undercount or $VO_2max$ underestimation. However, personal reliability is generally sufficient when data is interpreted as trend rather than isolated values (Höchsmann et al. 2020; Lambe et al. 2025).

Third, device hardware and algorithm changes  limit cross-device comparisons and favor interpretation that focuses on the direction and the magnitude of change for the same person over a longer period (Lambe et al. 2025).

### 4.2 Prioritization and tiering

Tiers summarize how strongly each marker matters for health and how reliably it can be tracked on consumer wearables for personal decisions. Higher tier = higher practical priority.

Tiering followed a prespecified rubric with four axes scored from 0 (worst) to 2 (best):

- Outcome evidence
- Measurability on mainstream wearables
- Consistency across studies / evidence quality
- Practical actionability

Tier 1 required a total score >6. Tier 2 required a total score of 4-6 with no axis scored 0.

**Table 2**: Tier list

| Marker | Outcome | Measurability | Consistency | Actionability | Total | Tier |
|---|---|---|---|---|---|---|
| Sleep duration | 2 | 2 | 2 | 2 | 8 | **1** |
| Steps | 2 | 2 | 2 | 2 | 8 | **1** |
| Resting heart rate | 2 | 2 | 2 | 1 | 7 | **1** |
| Heart rate variability | 2 | 1 | 2 | 1 | 6 | **2** |
| $VO_2max$ | 2 | 1 | 1 | 1 | 5 | **2** |
| Sleep regularity | 2 | 1 | 1 | 1 | 5 | **2** |

Sleep duration outcomes are strong, accurate consumer measurability of total sleep time, consistent and strong findings, and simple actionability.

Steps show strong health associations, reliable within-person trends despite 20% undercount bias, good evidence and easy daily application.

Resting heart rate combines strong outcome associations with close agreement to ECG during sleep, evidence is reliable but actionability is not straight forward.

Heart rate variability has strong inverse relations with risk, wearable capture during sleep is accurate but daytime readings require caution regarding movement, evidence is good but actions should be based on longitudinal data.

$VO_2max$ shows strong inverse outcome relations, wearable estimates are reliable but have a high variation and wearable testing methods vary, evidence is moderate because more studies need to be done and practical use is limited to personal trend tracking.

Sleep regularity is related to outcomes, timing is measurable on wearables but sleep stages are often misclassified, evidence quality is low to moderate and current guidance remains to maintain consistent schedules as quantifiable targets are not established yet.

## 4.3 User perspective and usage

### 4.3.1 Perceived importance of markers

As of 9. December 2025, N = 85 valid respondents were included for the importance analysis (cleaning rules in Methods). Respondents rated the perceived importance of six wearable markers (1 to 7) which are then rescaled to 0-1 for display.

Figure 2 shows the rescaled means. Horizontal reference lines indicate the evidence benchmarks as visual anchors. The Tier 1 line (1.00) represents the theoretical maximum (ideal Tier 1). The Tier 2 line (0.625) represents the typical score of the Tier 2 markers (median). Markers are thus interpreted based on the proximity of these anchors.

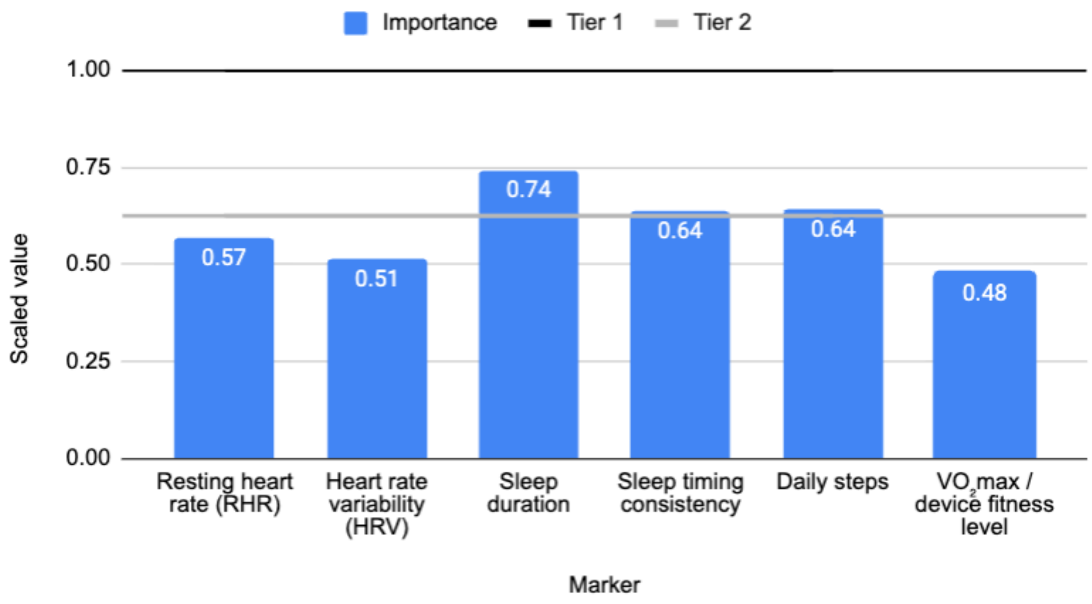



*Figure 2: Perceived importance of health markers*

Sleep duration shows the highest importance (0.74), followed by daily steps (0.64) and sleep timing consistency (0.64). Resting heart rate averages 0.57, heart rate variability 0.51, and $VO_2$max/device fitness level 0.48.

Relative to the reference lines, sleep duration stands out as the single highest priority. Strictly speaking, the score sits just at the boundary of the theoretical Tier 1 threshold (0.75) but it clearly separates from the rest of the markers and sits well above the typical Tier 2 line. This suggests that users prioritize it at a Tier 1 level.

In contrast, daily steps and resting heart rate reveal a notable gap. While the evidence classifies them as Tier 1, users currently perceive them only as Tier 2 priorities. This indicates that they are undervalued. Sleep timing consistency is prioritized within the tier 2 range, aligning well with its evidence classification. HRV and $VO_2$max are at the bottom, which aligns with their Tier 2 classification.

### 4.3.2 Practice index among wearable owners

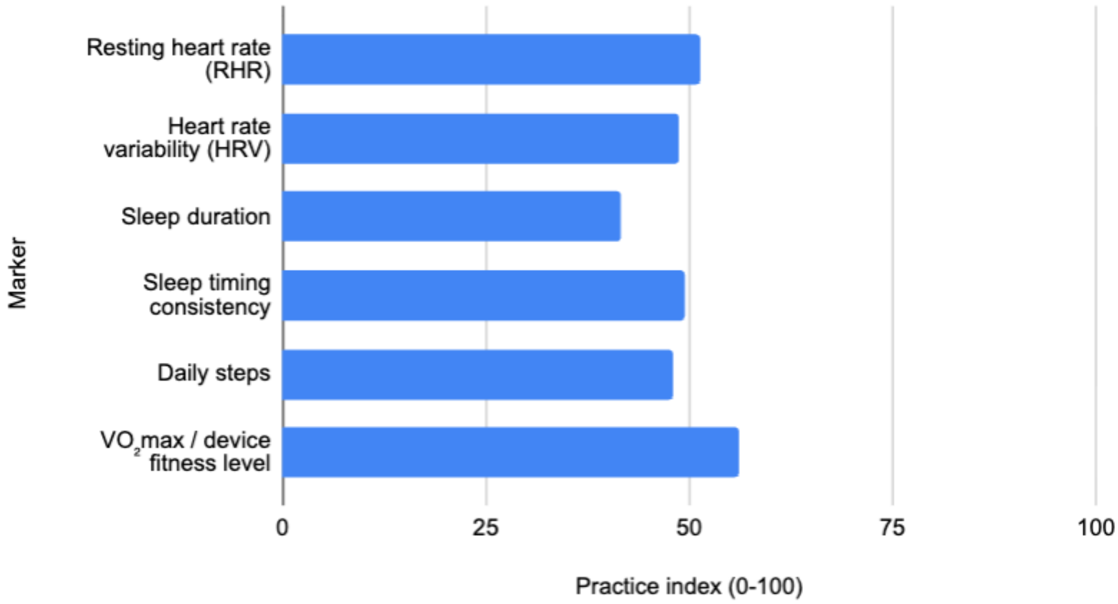



*Figure 3: Practice index*

Based on data collected up to 9. December 2025, figure 3 summarizes with N = 26 respondents how wearable users report interpreting each marker using a 0-100 practice index. Higher means greater reliance on aggregated/multi day readings and lower more on single day readings. For HRV specifically, higher means reliance on nighttime/sleep readings and lower on daytime readings.

Scores cluster around the mid-range. $VO_2$max/device fitness level is highest (56/100) indicating a preference for multi day readings, followed by resting heart rate (51/100), sleep timing consistency (49/100), heart rate variability (49/100) and daily steps (48/100). Sleep duration is lowest (42/100), suggesting a emphasis towards single night readings. No marker approaches either extreme of the 0-100 scale.

### 4.3.3 Adoption table

Table 3 reports self-stated adoption as of 9. December 2025 among wearable owners for the question “Which of these did you look at at least once in the last week?” Responses were multi-select and are summarized as n/N and percent for owners who answered (N = 26). Because this question was multi-select, percentages do not sum up to 100%.

Among respondents, daily steps was most commonly checked (18/26; 70%) together with sleep duration (18/26; 70%) followed by resting heart rate (16/26; 60%). Sleep timing consistency was reported by 10/26 (40%), heart rate variability by 7/26 (30%), and $VO_2max$/device fitness level by 5/26 (20%).

**Table 3**: Adoption rate

| Marker | n/N | Looked last week |
|---|---|---|
| Resting heart rate (RHR) | 16 / 26 | 60% |
| Heart rate variability (HRV) | 7 / 26 | 30% |
| Sleep duration | 18 / 26 | 70% |
| Sleep timing consistency | 10 / 26 | 40% |
| Daily steps | 18 / 26 | 70% |
| $VO_2max$ (or device fitness level) | 5 / 26 | 20% |

### 4.4 Operational implications by marker

Operational implications are derived by triangulating (i) the outcome and measurement syntheses per marker, (ii) the tiering in 4.2, and (iii) the results from the survey.

#### 4.4.1 Sleep duration (Tier 1)

**Survey patterns.** Highest perceived importance (0.74), lowest practice index (42/100), and 70% of owners checked it last week. This means it is valued highest, mostly interpreted for single nights and checked rather often.

**Operational implications.** Use nightly sleep duration from consumer devices (newer generations estimate total sleep time with good accuracy) and interpret it against the recommended ~7-8 h, while avoiding extremes at the short (<7 h) and long (>9 h) end.

Because perceived importance is high but the practice index is low (which indicates a tilt toward single-night readings), the improvement is to shift interpretation toward an averaged, several-night view and apply the 7-8h benchmark to that aggregate, rather than reacting to one night highs or lows. Keep the device constant to preserve comparability.

### 4.4.2 Daily steps (Tier 1)

**Survey patterns.** Moderate/high perceived importance (0.64), moderate practice index (48/100) and high adoption (18/26; 70%). This means steps are seen as somehow important, interpreted as a mix of single day readings and multiple day averages, and rather often checked.

**Operational implications.** Increase personal priority to close the awareness gap and prioritize same-device interpretation because consumer devices undercount steps by about 20%. This means to treat same-device increases/decreases as meaningful. A 7 000 to 10 000 steps per day range aligns with good health outcomes (large gains around 7 000 and diminishing returns above 10 000).

Because adoption is high but the practice index is only moderate, many users likely read day to day totals rather than summarized patterns. The improvement is therefore not more frequent checking but a better method. Shift emphasis from single-day totals toward bias-aware weekly summaries on the same device. Read week over week directions of change, apply the ≈7 000-10 000 range as context rather than a rigid daily target, and explicitly account for systematic undercount when interpreting changes.

### 4.4.3 Resting heart rate (Tier 1)

**Survey patterns.** Moderate perceived importance (0.57), moderate practice index (51/100) and moderate/high adoption (16/26; 60%). This means RHR is seen as somehow important, interpreted more through a mix of multiple day and single day readings and checked occasionally.

**Operational implications.** Focus on nocturnal/sleep-window readings over time (close agreement to ECG) and keep to the same device for comparability.

Because perceived importance is moderate, practice is mixed and adoption is high, the overall improved approach for RHR is to focus more on multi night averages. Furthermore, moderate perceived importance indicates a need to foster more awareness for this health metric.

#### 4.4.4 Heart rate variability (Tier 2)

**Survey patterns.** Perceived importance is moderate (0.51) but one of the lowest compared to the other markers, moderate practice index (49/100) (higher = more night/sleep readings, lower = more daytime/movement readings) and low adoption (7/26; 30%). This means that HRV is seen as neither low or high in importance, interpreted as a mix of nighttime readings and daytime readings, and rarely checked.

**Operational implications.** Focus on sleep based HRV collected during the night and avoid daytime readings or readings affected by movement. Interpret the metric longitudinally across sleep periods. Use it as supporting context alongside RHR and sleep markers.

Because the practice index sits mid-range (indicating many users still default to mixed readings and especially daytime) and the adoption is low, the priority is a method upgrade and more frequent checking.

Shift HRV capture into the nightly sleep window, reduce daytime spot checks, read night to night changes as contextual input alongside RHR and sleep, and maintain the same device for comparability.

#### 4.4.5 Sleep regularity (Tier 2)

**Survey patterns.** Moderate/high perceived importance (0.64), moderate practice index (49/100) and moderate/low adoption rate (10/26; 40%). This means sleep regularity is seen as somehow important, interpreted as a mix of single and averaged readings, and occasionally checked.

**Operational implications.** Monitor bed/wake timing consistency across the week (wearables capture timing/regularity accurately). Do not introduce numeric targets as they are not established yet and avoid relying on sleep-stage outputs, because they are often misclassified.

Because importance and practice are moderate but adoption is on the lower side, the lever is integration and cadence. Add a weekly pattern check of bed/wake times alongside sleep duration, keep interpretation directional (more regular vs less regular), and avoid stage-based interpretations or numeric thresholds.

#### 4.4.6 $VO_2max$ / device fitness level (Tier 2)

**Survey patterns.** Moderate but comparatively the lowest perceived importance score (0.48), highest practice index (56/100) and lowest adoption (5/26; 20%). This means that cardiorespiratory fitness is seen as neither high or low regarding importance, interpreted more towards averaged values, and very rarely checked.

**Operational implications.** Prefer exercise-based testing protocols. Keep to the same device and interpret as within-person trends given underestimation and individual scatter.

While outcome evidence links good cardiorespiratory fitness to strong health associations, the comparatively low perceived importance aligns with its category as a Tier 2 supporting marker. Given device underestimation and substantial individual measurement variation, do not reallocate much weekly effort toward $VO_2max$. Keep it in a low-emphasis role and focus on within-person trends only. It is best seen as a slow moving corroborator of long-term conditioning rather than a short-term decision driver.

## 5. Limitations

Given a 25 page limit and time constraints, this thesis prioritized high-leverage sources over exhaustiveness. A predefined pre-screening heuristic (Top 30 “Best Match” plus a “Most Recent” sweep per query) was applied before screening to manage large result sets. This could introduce selection bias and affect representativeness. Searching relied on a single database (PubMed) with English language limits and a short search window in October 2025, so relevant evidence in other databases, languages, or time periods could have been missed. To avoid redundancy, 1-3 anchor reports per stream per marker were retained, improving clarity but potentially under representing heterogeneity and dissenting findings. A small number of additional context sources were also cited outside the structured search, which could introduce confirmation bias, they are also not included in the PRISMA evidence set. The eligibility criteria deliberately focused on adults aged 18-65 in non-shift employment, so the findings and operational implications should not be generalized to older adults, adolescents, shift workers, or other groups without caution. The tiering framework and axis weights were constructed by the author and have not been formally validated. Scores and tiers should therefore not be interpreted as definitive rankings that would necessarily hold across different populations, devices, or review methods. The descriptive survey of working adults was based on a small convenience sample (N = 85; wearable owners N = 26), recruited online and analyzed cross-sectionally, so results are not representative of the wider working population. Finally, the per-marker operational implications translate observational evidence and validation findings into simplified guidance for non-clinical use. The recommendations have not been evaluated and should be seen as decision support rather than prescriptive clinical recommendations.

## 6. Future directions

Further work is most needed to define quantitative targets for sleep regularity, to standardize nocturnal heart-rate metrics across devices and algorithms, and to harmonize wearable $VO_2$max protocols and accuracy. Beyond marker-specific work, future studies should link wearable-derived metrics directly to major health outcomes in working adults, using the same device outputs and aggregation windows that users see in practice. The tiering framework and axis weights developed here could be tested, refined, or re-weighted in other populations and device ecosystems. Larger, more representative surveys could clarify how working adults actually use these markers over time and whether simple, marker focused dashboards or coaching interventions change behaviour and downstream health risk. Finally, as firmware and algorithms evolve, periodic revalidation of key markers will be required.

## 7. Conclusion

This thesis set out to answer a practical question for working adults who use consumer wearables. Among the many metrics that devices now provide, which markers are most worth prioritising because they are both strongly associated to major health outcomes and measurable with sufficient reliability in everyday life? To address this, the work focused on six prespecified markers (sleep duration, sleep regularity, resting heart rate, heart rate variability, daily steps, and $VO_2$max/device fitness level) and combined a rapid scoping review with a tiering framework and a descriptive survey of working adults. Together, these components move from isolated findings towards an integrated, device agnostic prioritisation framework.

Regarding methodology, the thesis separated evidence into two streams for each marker. The outcome stream synthesised how strongly each marker relates to major health endpoints, while the measurement stream assessed how well consumer wearables capture the same marker in real life conditions relative to reference standards. The evidence was then scored on four axes (outcome evidence, measurability on consumer wearables, consistency across devices and studies, and practical actionability in daily life). These scores were then used to classify markers into tiers, while the survey added a behavioural lens by describing how working adults currently perceive and use the same set of markers.

Across the six markers, the outcome synthesis confirm that all are meaningfully linked to health. Sleep duration shows a characteristic U-shaped pattern, with increased risk at short and long durations relative to a broad mid-range (7 to 8h). Resting heart rate and heart rate variability capture aspects that relate to all-cause and cardiovascular outcomes. Daily steps and cardiorespiratory fitness are consistently associated with lower cardiometabolic and mortality risk as levels increase. Sleep regularity is emerging as a complementary dimension to sleep duration, with less consistent but still good links to outcomes. Overall, the outcome stream indicates that all six markers have a legitimate claim to matter for long-term health.

The measurement stream shows that not all of these markers are equally straightforward to obtain and interpret from consumer devices. Sleep duration and daily steps are captured reasonably well by consumer wearables despite device specific differences. Resting heart rate is measured very accurately in resting or nocturnal windows. Heart rate variability can be captured with acceptable accuracy during sleep or rest, but values are worse under movement and depend heavily on device-specific algorithms. For $VO_2$max, devices tend to underestimate absolute levels and are less extensively validated. Sleep regularity is measurable from sleep and wake times, but targets and thresholds are not yet established. These measurement differences are central for deciding which markers should be treated as primary versus supporting indicators.

Using the four axis rubric, the thesis translated this combined evidence into a tiered framework. Sleep duration, daily steps, and resting heart rate emerged as Tier 1 markers. They pair strong outcome evidence with acceptable measurement and straightforward interpretation. Heart rate variability, sleep regularity, and $VO_2$max/device fitness level were classified as Tier 2 markers. For these, outcome associations are generally favourable. But less extensive validation on wearables, greater methodological heterogeneity, device- and algorithm-specific biases, or more complex interpretation suggest a supporting, trend level role rather than a top priority. The framework emphasises personal trends over specific cut offs and encourages multiple day or night patterns rather than isolated values.

The survey results complement these tiers by showing how working adults currently prioritise and use the same markers. Triangulating evidence and survey findings highlights both alignment (for example, the prominence of sleep duration) and gaps (for example, the undervaluation of daily steps and resting heart rate).

Within the constraints acknowledged in the Limitations and Future directions chapters, the thesis makes pragmatic contributions. It brings together outcome and measurement evidence that are often considered separately, packages them into per-marker summaries, and offers a tiering framework. It then connects this to how working adults actually use their devices and translates the combined insights into operational implications.

Overall, the work suggests that making better use of consumer wearables does not require more and more metrics, but a clearer hierarchy among existing ones. For working adults, anchoring attention on a small set of Tier 1 markers (sleep duration, daily steps, and resting heart rate) and reading them primarily through a multiple-day pattern provides a practical, evidence based starting point, with Tier 2 markers layered in to provide longer-term context as devices and evidence evolve.

## References

Bizzozero-Peroni, Bruno, Valentina Díaz-Goñi, Estela Jiménez-López, Eva Rodríguez-Gutiérrez, Irene Sequí-Domínguez, Sergio Núñez de Arenas-Arroyo, José Francisco López-Gil, Vicente Martínez-Vizcaíno, and Arthur Eumann Mesas. 2024. "Daily Step Count and Depression in Adults: A Systematic Review and Meta-Analysis." JAMA Network Open, 7(12): e2451208.

Chaput, Jean-Philippe, Caroline Dutil, Ryan Featherstone, Robert Ross, Lora Giangregorio, Travis J. Saunders, Ian Janssen, et al. 2020a. "Sleep Duration and Health in Adults: An Overview of Systematic Reviews." Applied Physiology, Nutrition, and Metabolism, 45: S218-S231.

Chaput, Jean-Philippe, Caroline Dutil, Ryan Featherstone, Robert Ross, Lora Giangregorio, Travis J. Saunders, Ian Janssen, et al. 2020b. "Sleep Timing, Sleep Consistency, and Health in Adults: A Systematic Review." Applied Physiology, Nutrition, and Metabolism, 45: S232-S247.

Dial, Michael B., Margaret E. Hollander, Emaly A. Vatne, Angela M. Emerson, Nathan A. Edwards, and Joshua A. Hagen. 2025. "Validation of Nocturnal Resting Heart Rate and Heart Rate Variability in Consumer Wearables." Physiological Reports, 13: e70527.

Ding, Ding, Binh Nguyen, Tracy Nau, Mengyun Luo, Borja del Pozo Cruz, Paddy C. Dempsey, and Zachary Munn, et al. 2025. "Daily Steps and Health Outcomes in Adults: A Systematic Review and Dose-Response Meta-Analysis." The Lancet Public Health, 10: e668-e681.

Georgiou, Konstantinos, Andreas V. Larentzakis, Nehal N. Khamis, Ghadah I. Alsuhaibani, Yasser A. Alaska, and Elias J. Giallafos. 2018. "Can Wearable Devices Accurately Measure Heart Rate Variability? A Systematic Review." Folia Medica (Plovdiv), 60(1): 7-20.

Han, Minghui, Ranran Qie, Xuezhong Shi, Yongli Yang, Jie Lu, Fulan Hu, Ming Zhang, Zhenzhong Zhang, Dongsheng Hu, and Yang Zhao. 2022. "Cardiorespiratory Fitness and

Mortality from All Causes, Cardiovascular Disease and Cancer: Dose-Response Meta-Analysis of Cohort Studies." British Journal of Sports Medicine, 56: 733-739.

Höchsmann, Christoph, Raphael Knaier, Denis Infanger, and Arno Schmidt-Trucksäss. 2020. "Validity of Smartphones and Activity Trackers to Measure Steps in a Free-Living Setting over Three Consecutive Days." Physiological Measurement, 41: 015001.

Jarczok, Marc N., Katja Weimer, Christin Braun, DeWayne P. Williams, Julian F. Thayer, Harald O. Gündel, and Elisabeth M. Balint. 2022. "Heart Rate Variability in the Prediction of Mortality: A Systematic Review and Meta-Analysis of Healthy and Patient Populations." Neuroscience & Biobehavioral Reviews, 143: 104907.

Kainec, Kyle A., Jamie Caccavaro, Morgan Barnes, Chloe Hoff, Annika Berlin, and Rebecca M. C. Spencer. 2024. "Evaluating Accuracy in Five Commercial Sleep-Tracking Devices Compared to Research-Grade Actigraphy and Polysomnography." Sensors, 24: 635.

Labie, Céline, Nils Runge, Zosia Goossens, Olivier Mairesse, Jo Nijs, Anneleen Malfliet, Dieter Van Assche, et al. 2025. "Can a Commercially Available Smartwatch Device Accurately Measure Nighttime Sleep Outcomes in Individuals with Knee Osteoarthritis and Comorbid Insomnia? A Comparison with Home-Based Polysomnography." Sensors, 25: 4813.

Lambe, Rory, Ben O'Grady, Maximus Baldwin, and Cailbhe Doherty. 2025. "Investigating the Accuracy of Apple Watch VO2 Max Measurements: A Validation Study." PLOS ONE, 20(5): e0323741.

Li, Jin, Dehong Cao, Yin Huang, Zeyu Chen, Ruyi Wang, Qiang Dong, Qiang Wei, and Liangren Liu. 2022. "Sleep Duration and Health Outcomes: An Umbrella Review." Sleep and Breathing, 26: 1479-1501.

Molina-Garcia, Pablo, Hannah L. Notbohm, Moritz Schumann, Rob Argent, Megan Hetherington-Rauth, Julie Stang, Wilhelm Bloch, Sulin Cheng, Ulf Ekelund, Luis B. Sardinha, Brian Caulfield, Jan Christian Brønd, Anders Grøntved, and Francisco B. Ortega. 2022. "Validity of Estimating the Maximal Oxygen Consumption by Consumer Wearables: A Systematic Review with Meta-analysis and Expert Statement of the INTERLIVE Network." Sports Medicine, 52: 1577-1597.

Tricco, Andrea C., Erin Lillie, Wasifa Zarin, Kelly K. O'Brien, Heather Colquhoun, Danielle Levac, David Moher, et al. 2018. "PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation." Annals of Internal Medicine, 169(7): 467-473.

Ungvari, Zoltan, Mónika Fekete, Péter Varga, János Tibor Fekete, Andrea Lehoczki, Annamaria Buda, Ágnes Szappanos, György Purebl, Anna Ungvari, and Balázs Győrffy. 2025. "Imbalanced Sleep Increases Mortality Risk by 14-34%: A Meta-Analysis." GeroScience, 47: 4545-4566.

Zhang, Dongfeng, Xiaoli Shen, and Xin Qi. 2016. "Resting Heart Rate and All-Cause and Cardiovascular Mortality in the General Population: A Meta-Analysis." CMAJ, 188(3): E53-E63.

# Appendix

## Appendix A. Search strategies

**Table 4**: PRISMA search queries

| Nr | Marker | Search purpose | Database | Search Query | Filters | Date run | Results count | Notes | Total Scanned |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Sleep duration | Outcome - SR/MA | PubMed | (sleep duration OR total sleep time) AND (mortality OR all-cause mortality OR cardiovascular OR cardiometabolic OR diabetes OR depression OR anxiety) | 10 years; Meta-Analysis, Systematic Review; English; Humans; Adult: 19+ years | 14.10.2025 | 192 | | 30+15 scanned |
| 2 | Sleep duration | Outcome - Cohort | PubMed | (sleep duration OR total sleep time) AND (mortality OR all-cause mortality OR cardiovascular OR cardiometabolic OR diabetes OR depression OR anxiety) AND (cohort OR prospective) | 10 years; English; Humans; Adult: 19+ years | 14.10.2025 | 1712 | | 30+15 scanned |
| 3 | Sleep duration | Measurement / validation | PubMed | ("sleep duration" OR "total sleep time" OR "time in bed" OR "sleep onset" OR "sleep offset") AND (wearable OR smartwatch OR ring OR "consumer device" OR Fitbit OR Garmin OR "Apple Watch" OR Oura) AND (validation OR validity OR accuracy OR agreement OR reliability) AND (polysomnography OR PSG OR actigraphy) AND (free-living OR home OR "at home" OR ambulatory OR nocturnal OR overnight) | 10 years; English; Humans; Adult: 19+ years | 15.10.2025 | 52 | | 30+15 scanned |
| 4 | Sleep regularity | Outcome - SR/MA | PubMed | ("sleep regularity" OR "sleep variability" OR "sleep timing" OR "sleep regularity index" OR "social jetlag") AND (mortality OR "all-cause mortality" OR cardiovascular OR cardiometabolic OR diabetes OR depression OR anxiety) | 10 years; Meta-Analysis, Systematic Review; English; Humans; Adult: 19+ years | 15.10.2025 | 6 | | 6 scanned |

| Nr | Marker | Search purpose | Database | Search Query | Filters | Date run | Results count | Notes | Total Scanned |
|---|---|---|---|---|---|---|---|---|---|
| 5 | Sleep regularity | Measurement / validation | PubMed | ("sleep timing" OR "sleep onset" OR "sleep offset" OR "bedtime" OR "wake time" OR "midsleep" OR "sleep regularity index" OR "social jetlag") AND (wearable OR smartwatch OR ring OR "consumer device" OR Fitbit OR Garmin OR "Apple Watch" OR Oura) AND (validation OR validity OR accuracy OR agreement OR reliability OR "Bland-Altman" OR ICC) AND (polysomnography OR PSG OR actigraphy OR diary) AND (free-living OR home OR "at home" OR ambulatory OR overnight) | 10 years; English; Humans; Adult: 19+ years | 15.10.2025 | 33 | | 33 scanned |
| 6 | Resting Heart Rate | Outcome - SR/MA | PubMed | ("resting heart rate" OR "nocturnal heart rate" OR "night heart rate" OR "sleep heart rate") AND (mortality OR "all-cause mortality" OR cardiovascular OR "coronary heart disease" OR stroke OR "heart failure" OR cardiometabolic OR diabetes OR depression OR anxiety) | 10 years; Meta-Analysis, Systematic Review; English; Humans; Adult: 19+ years | 16.10.2025 | 34 | | 34 scanned |
| 7 | Resting Heart Rate | Outcome - SR/MA | citation chasing | Backward reference check of: Ada N Nordeidet, 2024, doi.org/10.1093/eurjpc/zwae092 | - | 16.10.2025 | 1 | | 1 scanned |
| 8 | Resting Heart Rate | Outcome - SR/MA | citation chasing | Backward reference check of: Marie Therese Cooney, 2010, 10.1016/j.ahj.2009.12.029 | - | 16.10.2025 | 1 | | 1 scanned |
| 9 | Resting Heart Rate | Outcome - SR/MA | citation chasing | Backward reference check of: Ester Puig, 2021, https://doi.org/10.1093/eurjpc/zwab148 | - | 16.10.2025 | 1 | | 1 scanned |
| 10 | Resting Heart Rate | Measurement / validation | PubMed | ("resting heart rate" OR "nocturnal heart rate" OR "night heart rate" OR "sleep heart rate") AND (wearable OR smartwatch OR ring OR "consumer device" OR Fitbit OR Garmin OR "Apple Watch" OR Oura OR WHOOP) AND (validation OR validity OR accuracy OR agreement OR reliability) AND (ECG OR electrocardiogram) AND (free-living OR home OR overnight OR nocturnal) | 10 years; English; Humans; Adult: 19+ years | 16.10.2025 | 3 | | 3 scanned |

| Nr | Marker | Search purpose | Database | Search Query | Filters | Date run | Results count | Notes | Total Scanned |
|---|---|---|---|---|---|---|---|---|---|
| 11 | Heart Rate Variability | Outcome - SR/MA | PubMed | ("heart rate variability" OR HRV) AND (mortality OR "all-cause mortality" OR cardiovascular OR "coronary heart disease" OR stroke OR "heart failure" OR cardiometabolic OR diabetes OR depression OR anxiety) | 10 years; Meta-Analysis, Systematic Review; English; Humans; Adult: 19+ years | 16.10.2025 | 47 | No good studies after screening | 30 scanned |
| 12 | Heart Rate Variability | Outcome - SR/MA | PubMed | ("heart rate variability" OR HRV) AND (mortality OR "all-cause mortality" OR "cardiovascular disease" OR "coronary heart disease" OR stroke OR "heart failure") | 10 years; Meta-Analysis, Systematic Review; English; Humans; Adult: 19+ years | 16.10.2025 | 16 | Still no good studies after screening | 16 scanned |
| 13 | Heart Rate Variability | Outcome - SR/MA | PubMed | ("heart rate variability" OR HRV) AND (mortality OR "all-cause mortality" OR death OR "cardiovascular disease" OR "coronary heart disease" OR stroke OR "heart failure") | 10 years; Meta-Analysis, Systematic Review; English; Humans | 16.10.2025 | 93 | | 30+15 scanned |
| 14 | Heart Rate Variability | Measurement / validation | PubMed | ("heart rate variability" OR HRV) AND (wearable OR smartwatch OR ring OR "consumer device" OR Fitbit OR Garmin OR "Apple Watch" OR Oura) AND (validation OR validity OR accuracy OR agreement OR reliability) AND (ECG OR electrocardiogram) AND (nocturnal OR night OR sleep OR resting OR rest) AND (free-living OR home OR ambulatory OR overnight) | 10 years; English; Humans; Systematic Review; Meta-Analysis | 17.10.2025 | 2 | | 2 scanned |
| 15 | Steps | Outcome - SR/MA | PubMed | (steps OR "step count" OR "daily steps") AND (mortality OR "all-cause mortality" OR cardiovascular OR cardiometabolic OR diabetes OR depression OR anxiety) | 10 years; English; Humans; Systematic Review; Meta-Analysis | 17.10.2025 | 371 | | 30+14 scanned |

| **Nr** | **Marker** | **Search purpose** | **Database** | **Search Query** | **Filters** | **Date run** | **Results count** | Notes | Total Scanned |
|---|---|---|---|---|---|---|---|---|---|
| 16 | Steps | Measurement / validation | PubMed | (steps OR "step count" OR "daily steps") AND (validation OR validity OR accuracy OR agreement OR reliability) AND (ActiGraph OR "research accelerometer" OR StepWatch OR "manual count") AND (free-living OR "free living" OR home OR community OR ambulatory OR "real world") AND (Fitbit OR Garmin OR "Apple Watch" OR Oura OR Withings OR Polar OR Samsung OR WHOOP) | 10 years; English; Humans | 17.10.2025 | 57 | | 30+11 scanned |
| 17 | $VO_2max$ | Outcome - SR/MA | PubMed | ("cardiorespiratory fitness" OR $VO_2max$ OR "maximal oxygen uptake" OR "maximal oxygen consumption" OR "aerobic fitness") AND (mortality OR "all-cause mortality" OR cardiovascular OR "coronary heart disease" OR stroke OR "heart failure" OR diabetes OR depression OR anxiety) | 10 years; English; Humans; Systematic Review; Meta-Analysis | 17.10.2025 | 348 | | 30+15 scanned |
| 18 | $VO_2max$ | Outcome - SR/MA | PubMed | ("cardiorespiratory fitness" OR $VO_2max$ OR "maximal oxygen uptake" OR "aerobic fitness") AND (mortality OR "all-cause mortality" OR cardiovascular OR "cardiovascular disease" OR CVD OR cancer) | 10 years; English; Humans; Systematic Review; Meta-Analysis | 17.10.2025 | 298 | | 31+15 scanned |
| 19 | $VO_2max$ | Measurement / validation | PubMed | ($VO_2max$ OR "cardiorespiratory fitness" OR "maximal oxygen uptake") AND ("Apple Watch" OR smartwatch) AND (validation OR validity OR accuracy OR agreement OR reliability) AND ("cardiopulmonary exercise test" OR CPET OR "indirect calorimetry") | 10 years; English; Humans | 17.10.2025 | 7 | | 7 scanned |

## Appendix B. Screening

**Table 5**: PRISMA screening

| Nr | DOI | Marker | Study Type | Design | Title | Abstract Screening? | Full-text reading? | Full-text decision | Reason |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 10.1016/j.biopsych.2015.05.014 | Sleep Duration | Outcome | Meta-Analysis | Sleep Disturbance, Sleep Duration, and Inflammation: A Systematic Review and Meta-Analysis of Cohort Studies and Experimental Sleep Deprivation | Yes | No | - | Too much focus on sleep disturbance and inflammation |
| 1 | 10.1038/s41467-023-41249-y | Sleep Duration | Outcome | Meta-Analysis | Genome-wide association studies and cross-population meta-analyses investigating short and long sleep duration | Yes | No | - | Too focused on genomes |
| 1 | 10.1007/s11325-021-02458-1 | Sleep Duration | Outcome | Umbrella-Review | Sleep duration and health outcomes: an umbrella review | Yes | **Yes** | **Include** | - |
| 1 | 10.1016/j.psyneuen.2022.105843 | Sleep Duration (Steps) | Outcome | Systematic Review | The effects of physical activity on cortisol and sleep: A systematic review and meta-analysis | Yes | No | - | Population issues and not relevant for sleep duration |
| 1 | 10.1007/s11357-025-01592-y | Sleep duration | Outcome | Meta-Analysis | Imbalanced sleep increases mortality risk by 14-34%: a meta-analysis | Yes | **Yes** | **Include** | - |
| 2 | 10.1139/apnm-2020-0034 | Sleep Duration | Outcome | Overview of systematic reviews | Sleep duration and health in adults: an overview of systematic reviews | Yes | **Yes** | **Include** | - |
| 2 | 10.2105/AJPH.2016.303326 | Sleep Duration | Outcome | Narrative review | Determinants and Consequences of Obesity | Yes | **Yes** | **Exclude** | Did not meet criteria (context) |

| Nr | DOI | Marker | Study Type | Design | Title | Abstract Screening? | Full-text reading? | Full-text decision | Reason |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 10.2337/dc24-1208 | Sleep Duration (Sleep regularity) | Outcome | Cohort study | Sleep Irregularity and the Incidence of Type 2 Diabetes: A Device-Based Prospective Study in Adults | Yes | No | - | Focus on T2 Diabetes only |
| 2 | 10.3389/fpubh.2022.880276 | Sleep Duration | Outcome | Cohort study | Association of Sleep Duration With All-Cause and Cardiovascular Mortality: A Prospective Cohort Study | Yes | No | - | This cohort study does not anything to the SR/MA |
| 3 | 10.3390/s24206532 | Sleep duration | Measurement | | Accuracy of Three Commercial Wearable Devices for Sleep Tracking in Healthy Adults | Yes | **Yes** | **Exclude** | Did not meet criteria (context) |
| 3 | 10.1080/15402002.2017.1300587 | Sleep duration | Measurement | | The Sleep of the Ring: Comparison of the ŌURA Sleep Tracker Against Polysomnography | Yes | No | - | Just OURA and only lab |
| 3 | 10.2196/20465 | Sleep duration | Measurement | | Sleep Tracking of a Commercially Available Smart Ring and Smartwatch Against Medical-Grade Actigraphy in Everyday Settings: Instrument Validation Study | Yes | No | - | Just OURA and one watch, not recent enough |
| 3 | 10.3390/s24237475 | Sleep duration | Measurement | | Deriving Accurate Nocturnal Heart Rate, rMSSD and Frequency HRV from the Oura Ring | Yes | No | - | Just OURA and not recent enough |
| 3 | 10.1016/j.sleep.2024.01.020 | Sleep duration | Measurement | | Validity and reliability of the Oura Ring Generation 3 (Gen3) with Oura sleep staging algorithm 2.0 (OSSA 2.0) when compared to multi-night ambulatory polysomnography: A validation study of 96 participants and 421,045 epochs | Yes | **Yes** | **Exclude** | Did not meet criteria (context) |

| Nr | DOI | Marker | Study Type | Design | Title | Abstract Screening? | Full-text reading? | Full-text decision | Reason |
|---|---|---|---|---|---|---|---|---|---|
| 3 | 10.3390/s24020635 | Sleep duration | Measurement | | Evaluating Accuracy in Five Commercial Sleep-Tracking Devices Compared to Research-Grade Actigraphy and Polysomnography | Yes | **Yes** | **Include** | - |
| 3 | 10.2196/13084 | Sleep duration | Measurement | | Validity Evaluation of the Fitbit Charge2 and the Garmin vivosmart HR+ in Free-Living Environments in an Older Adult Cohort | Yes | **Yes** | **Exclude** | Did not meet criteria (context) |
| 3 | 10.1016/j.sleh.2024.02.007 | Sleep duration | Measurement | | Performance of wearable sleep trackers during nocturnal sleep and periods of simulated real-world smartphone use | Yes | No | - | Focus on bedtime routines |
| 4 | 10.1139/apnm-2020-0032 | Sleep regularity | Outcome | Systematic Review | Sleep timing, sleep consistency, and health in adults: a systematic review | Yes | **Yes** | **Include** | - |
| 5 | 10.1371/journal.pone.0243214 | Sleep regularity | Measurement | | Assessing the performance of a commercial multisensory sleep tracker | Yes | No | - | Not recent enough |
| 5 | 10.2196/13384 | Sleep regularity | Measurement | | Accuracy of Fitbit Wristbands in Measuring Sleep Stage Transitions and the Effect of User-Specific Factors | Yes | No | - | Not recent enough |
| 5 | 10.3390/s25154813 | Sleep regularity | Measurement | | Can a Commercially Available Smartwatch Device Accurately Measure Nighttime Sleep Outcomes in Individuals with Knee Osteoarthritis and Comorbid Insomnia? A Comparison with Home-Based Polysomnography | Yes | **Yes** | **Include** | - |
| 6 | 10.1136/bjsports-2015-095841 | Resting Heart Rate | Outcome | Systematic Review | Effects of high-intensity interval training on cardiometabolic health: a systematic review and meta-analysis of intervention studies | Yes | No | - | Focus not on RHR |

| Nr | DOI | Marker | Study Type | Design | Title | Abstract Screening? | Full-text reading? | Full-text decision | Reason |
|---|---|---|---|---|---|---|---|---|---|
| 6 | 10.1002/dmrr.3095 | Resting Heart Rate | Outcome | Systematic Review | Resting heart rate and risk of type 2 diabetes: A prospective cohort study and meta-analysis | Yes | **Yes** | **Exclude** | Did not meet criteria (context) |
| 6 | 10.1093/eurjpc/zwae092 | Resting Heart Rate | Outcome | Systematic Review | Sex-specific and polygenic effects underlying resting heart rate and associated risk of cardiovascular disease | Yes | No | - | Focus on genomes |
| 6 | 10.1007/s00592-024-02369-z | Resting Heart Rate | Outcome | Systematic Review | Effect of resting heart rate on the risk of metabolic syndrome in adults: a dose-response meta-analysis | Yes | No | - | Focus on metabolic syndrome |
| 6 | 10.1016/j.cardfail.2020.11.003 | Resting Heart Rate | Outcome | Systematic Review | Resting Heart Rate as an Important Predictor of Mortality and Morbidity in Ambulatory Patients With Heart Failure: A Systematic Review and Meta-Analysis | Yes | No | - | Focus on ambulatory patients with heart failure |
| 6 | 10.1177/10998004221099556 | Resting Heart Rate | Outcome | Systematic Review | The Effect of Wearable and Smartphone Applications on Physical Activity, Quality of Life, and Cardiovascular Health Outcomes in Overweight/Obese Adults: A Systematic Review and Meta-analysis of Randomized Controlled Trials | Yes | No | - | Focus on physical activity |
| 6 | 10.1007/s13679-025-00645-9 | Resting Heart Rate | Outcome | Systematic Review | Additional Effect of Exercise to Intermittent Fasting on Body Composition and Cardiometabolic Health in Adults With Overweight/obesity: A Systematic Review and Meta-analysis | Yes | No | - | Focus on intermittent fasting |

| Nr | DOI | Marker | Study Type | Design | Title | Abstract Screening? | Full-text reading? | Full-text decision | Reason |
|---|---|---|---|---|---|---|---|---|---|
| 7 | 10.1016/j.ahj.2009.12.029 | Resting Heart Rate | Outcome | Cohort study | Elevated resting heart rate is an independent risk factor for cardiovascular disease in healthy men and women | Yes | No | - | Prefer a SR/MA |
| 8 | 10.1093/eurjpc/zwab148 | Resting Heart Rate | Outcome | Cohort study | Resting heart rate, cardiovascular events, and all-cause mortality: the REGICOR study | Yes | No | - | Prefer a SR/MA |
| 9 | 10.1503/cmaj.150535 | Resting Heart Rate | Outcome | Meta-Analysis | Resting heart rate and all-cause and cardiovascular mortality in the general population: a meta-analysis | Yes | **Yes** | **Include** | |
| 10 | 10.14814/phy2.70527 | Resting Heart Rate | Measurement | Cohort study | Validation of nocturnal resting heart rate and heart rate variability in consumer wearables | Yes | **Yes** | **Include** | - |
| 13 | 10.1371/journal.pone.0299793 | Heart Rate Variability | Outcome | Meta-Analysis | Beneficial impacts of physical activity on heart rate variability: A systematic review and meta-analysis | Yes | No | - | focus on physical activity and its impact on HRV |
| 13 | 10.1016/j.neubiorev.2022.104907 | Heart Rate Variability | Outcome | Systematic Review | Heart rate variability in the prediction of mortality: A systematic review and meta-analysis of healthy and patient populations | Yes | **Yes** | **Include** | - |
| 14 | 10.1007/s11325-025-03476-z | Heart Rate Variability | Measurement | Meta-Analysis | Electrocardiogram heart rate variability for machine learning diagnosis of obstructive sleep Apnoea: A bayesian meta-analysis | Yes | No | - | Focus on accuracy of machine learning based models |
| 14 | 10.2478/folmed-2018-0012 | Heart Rate Variability | Measurement | Systematic Review | Can Wearable Devices Accurately Measure Heart Rate Variability? A Systematic Review | Yes | **Yes** | **Include** | |

| Nr | DOI | Marker | Study Type | Design | Title | Abstract Screening? | Full-text reading? | Full-text decision | Reason |
|---|---|---|---|---|---|---|---|---|---|
| 15 | 10.1093/eurjpc/zwad229 | Steps | Outcome | Meta-Analysis | The association between daily step count and all-cause and cardiovascular mortality: a meta-analysis | Yes | No | - | Just looks at all-cause mortality and CVD |
| 15 | 10.1016/S2468-2667(21)00302-9. | Steps | Outcome | Meta-Analysis | Daily steps and all-cause mortality: a meta-analysis of 15 international cohorts | Yes | No | - | Just looks at all-cause mortality |
| 15 | 10.1007/s40279-021-01536-4 | Steps | Outcome | Meta-Analysis | Daily Step Count and All-Cause Mortality: A Dose-Response Meta-analysis of Prospective Cohort Studies | Yes | No | - | Not recent enough |
| 15 | 10.1016/j.ypmed.2024.108047 | Steps | Outcome | Umbrella-Review | Daily steps and all-cause mortality: An umbrella review and meta-analysis | Yes | No | - | Just looks at all-cause mortality |
| 15 | 10.1001/jamanetworkopen.2024.51208 | Steps | Outcome | Systematic Review | Daily Step Count and Depression in Adults: A Systematic Review and Meta-Analysis | Yes | **Yes** | **Include** | |
| 15 | 10.1016/S2468-2667(25)00164-1 | Steps | Outcome | Systematic Review | Daily steps and health outcomes in adults: a systematic review and dose-response meta-analysis | Yes | **Yes** | **Include** | |
| 15 | 10.1016/j.ijnurstu.2023.104563 | Steps | Outcome | Systematic Review | Sedentary behavior patterns and the risk of non-communicable diseases and all-cause mortality: A systematic review and meta-analysis | Yes | No | - | Focus on sedentary behaviour |
| 15 | https://doi.org/10.1093/eurjpc/zwad293 | Steps | Outcome | Meta-Analysis | The association between daily step count and all-cause and cardiovascular mortality: a meta-analysis | Yes | No | - | Focus on life expectancy |
| 15 | 10.1016/j.amjcard.2022.06.056 | Steps | Outcome | Meta-Analysis | Meta-analysis of Per-Day Step Count and All-Cause Mortality | Yes | No | - | Only focus on all-cause mortality and CVD |

| Nr | DOI | Marker | Study Type | Design | Title | Abstract Screening? | Full-text reading? | Full-text decision | Reason |
|---|---|---|---|---|---|---|---|---|---|
| 15 | 10.1161/CIRCULATIONAHA.122.061288 | Steps | Outcome | Meta-Analysis | Prospective Association of Daily Steps With Cardiovascular Disease: A Harmonized Meta-Analysis | Yes | No | - | Only focus on CVD |
| 15 | 10.1016/j.jshs.2021.09.004 | Steps | Outcome | Meta-Analysis | The relationships between step count and all-cause mortality and cardiovascular events: A dose-response meta-analysis | Yes | No | - | Not recent enough |
| 16 | 10.3390/s24144658 | Steps | Measurement | | Apple Watch 6 vs. Galaxy Watch 4: A Validity Study of Step-Count Estimation in Daily Activities | Yes | No | - | Just 2 devices and focus also on phone apps |
| 16 | 10.1186/s12874-023-01868-x | Steps | Measurement | | Validation of Oura ring energy expenditure and steps in laboratory and free-living | Yes | No | - | Just one device |
| 16 | 10.1016/j.gaitpost.2022.08.004 | Steps | Measurement | | Step-count outcomes of 13 different activity trackers: Results from laboratory and free-living experiments | Yes | **Yes** | **Exclude** | Wrong methodology |
| 16 | 10.1371/journal.pone.0251975 | Steps | Measurement | | Comprehensive comparison of Apple Watch and Fitbit monitors in a free-living setting | Yes | No | - | Just three devices |
| 16 | 10.2196/16674 | Steps | Measurement | | Low-Cost Consumer-Based Trackers to Measure Physical Activity and Sleep Duration Among Adults in Free-Living Conditions: Validation Study | Yes | No | - | Not recent enough |
| 16 | 10.1186/s12966-015-0314-1 | Steps | Measurement | | Systematic review of the validity and reliability of consumer-wearable activity trackers | Yes | No | - | Not recent enough |
| 16 | 10.1088/1361-6579/ab635f | Steps | Measurement | | Validity of smartphones and activity trackers to measure steps in a free-living setting over three consecutive days | Yes | **Yes** | **Include** | |

| Nr | DOI | Marker | Study Type | Design | Title | Abstract Screening? | Full-text reading? | Full-text decision | Reason |
|---|---|---|---|---|---|---|---|---|---|
| 17 | 10.1136/bjsports-2024-108748 | $VO_2$max | Outcome | Systematic Review and Meta-Analysis | Cardiorespiratory fitness, body mass index and mortality: a systematic review and meta-analysis | Yes | No | - | Focus on BMI |
| 17 | 10.1136/bjsports-2023-107849 | $VO_2$max | Outcome | Systematic Review | Cardiorespiratory fitness is a strong and consistent predictor of morbidity and mortality among adults: an overview of meta-analyses representing over 20.9 million observations from 199 unique cohort studies | Yes | No | - | Prefer more outcomes |
| 18 | 10.1016/j.atherosclerosis.2021.06.904 | $VO_2$max | Outcome | Meta-Analysis | Is estimated cardiorespiratory fitness an effective predictor for cardiovascular and all-cause mortality? A meta-analysis | Yes | No | - | Just focus on CV- and all-cause mortality |
| 18 | 10.1136/bjsports-2021-104876 | $VO_2$max | Outcome | Meta-Analysis | Cardiorespiratory fitness and mortality from all causes, cardiovascular disease and cancer: dose-response meta-analysis of cohort studies | Yes | **Yes** | **Include** | |
| 19 | 10.1371/journal.pone.0323741 | $VO_2$max | Measurement | | Investigating the accuracy of Apple Watch VO2 max measurements: A validation study | Yes | **Yes** | **Include** | |
| 19 | 10.2196/43340 | $VO_2$max | Measurement | | Smartwatch-Based Maximum Oxygen Consumption Measurement for Predicting Acute Mountain Sickness: Diagnostic Accuracy Evaluation Study | Yes | No | - | Focus on altitude, measurements at low and high altitude |

## Appendix C. Evidence

**Table 6**: PRISMA evidence

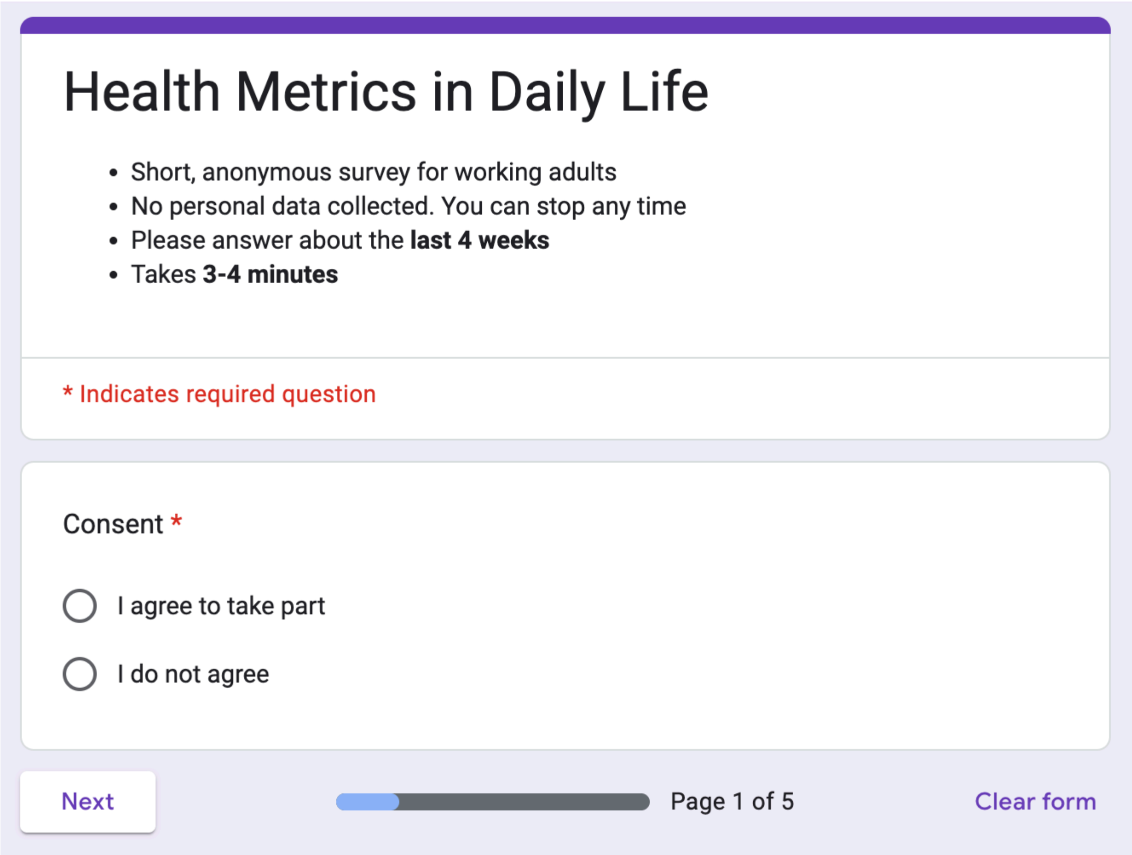
| Nr | Marker | Study Type | Author & Year | Design | Population / Devices | Outcomes / References | Findings | Limitations | DOI |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Sleep duration | Outcome | Jin Li, 2022 | Overview of Systematic reviews | 69 meta-analyses | cardiovascular disease, cognitive decline, coronary heart disease, depression, falls, frailty, lung cancer, metabolic syndrome, and stroke | Inappropriate sleep duration is a risk factor for developing non-cancer conditions. Decreasing and increasing sleep hours towards extreme sleep durations are associated with poor health outcomes | Self-reporting | 10.1007/s11325-021-02458-1 |
| 1 | Sleep duration | Outcome | Zoltan Ungvari, 2025 | Meta Analysis of cohort studies | 79 studies | all-cause mortality | Sleep below 7h results in 14% increase in mortality risk, above 8h 34% higher risk | Diverse range of populations, settings and methodologies; sleep duration self-reporting | 10.1007/s11357-025-01592-y |
| 2 | Sleep duration | Outcome | Jean-Philippe Chaput, 2020 | Overview of Systematic reviews | 11 Systematic reviews | mortality, cardiovascular, T2 diabetes, mental health, brain health, cognitive function, falls, injuries and other health outcomes | Seven to 8 h of sleep per day was most favourably associated with health. | 96% of studies relied on self-reported sleep, many reviews had also no subgroups | 10.1139/apnm-2020-0034 |
| 3 | Sleep duration | Measurement | Kyle A. Kainec et al. 2024 | Validation | 5 devices | Reference = polysomnography and research-grade actigraphy | All but one were accurate regarding sleep duration | Devices worn for one night, validity of devices may improve with upgrades | 10.3390/s24020635 |

| Nr | Marker | Study Type | Author & Year | Design | Population / Devices | Outcomes / References | Findings | Limitations | DOI |
|---|---|---|---|---|---|---|---|---|---|
| 4 | Sleep regularit y | Outcome | Jean-Philippe Chaput, 2020 | Systematic review | 41 studies | mortality, cardiovascular, T2 diabetes, mental health, brain health, cognitive function, falls, injuries and other health outcomes | Later sleep timing and greater sleep variability were generally associated with adverse health outcomes | 63% of studies relied on self-reported sleep | 10.1139/apnm-2020-0032 |
| 5 | Sleep regularit y | Measurem ent | Celine Labie, 2025 | Validation | 1 device (Fitbit) | Reference = polysomnography | The device shows limited ability to differentiate quiet wakefulness from sleep and less precision in classifying specific sleep stages. Sleep stage and fragmentation data should be interpreted with caution | Findings specific to FitBit Sense, just nighttime sleep and no napping | 10.3390/s2515481 3 |
| 9 | Resting Heart Rate | Outcome | Dongfeng Zhang, 2016 | Meta Analysis of cohort studies | 46 studies | all-cause mortality, cardiovascular mortality | Higher resting heart rate was independently associated with increased risks of all-cause and cardiovascular mortality. | Nighttime HR better than RHR, possibility of HR-lowering medications during follow-up, publication bias | https://doi.org/10. 1503/cmaj.15053 5 |
| 10 | Resting Heart Rate | Measurem ent | Michael B Dial, 2025 | Validation | 5 devices | Reference = ECG | Oura devices showed the highest agreement for RHR and HRV, and WHOOP showed acceptable agreement, whereas Garmin Fenix and Polar demonstrated lower concordance | atrial fibrillation interferes with normal heart rhythms and thus should be considered, individual device algorithms | https://doi.org/10. 14814/phy2.7052 7 |

| Nr | Marker | Study Type | Author & Year | Design | Population / Devices | Outcomes / References | Findings | Limitations | DOI |
|---|---|---|---|---|---|---|---|---|---|
| 13 | Heart Rate Variability | Outcome | Marc N Jarczok, 2022 | Systematic review | 32 studies | all-cause mortality, cardiovascular mortality | Lower HRV parameter values were significant predictors of higher mortality | Most studies report only selective HRV parameters, limiting comparability of those parameters across studies | 10.1016/j.neubiorev.2022.104907 |
| 14 | Heart Rate Variability | Measurement | Konstantinos Georgiou, 2018 | Systematic review | 18 studies | Reference = ECG for most of them (two used PPG) | Resting HRV measured with wearables shows good to excellent agreement with ECG but accuracy declines with exercise | Only a few PPG studies | 10.2478/folmed-2018-0012 |
| 15 | Steps | Outcome | Ding Ding, 2025 | Systematic review and meta-analysis | 57 studies for SR and 31 for MA | all-cause mortality, cardiovascular disease, cancer, type 2 diabetes, cognitive outcomes, mental health outcomes, physical function, falls | Modest daily step counts were associated with health benefits but 7000 steps per day was associated with sizeable risk reductions across most outcomes, compared with the reference of 2000 steps per day. Even though risk continued to decrease beyond 7000 steps per day, it plateaued for some outcomes | small number of studies on outcomes, lack of age specific analysis, biases on the individual study level | 10.1016/S2468-2667(25)00164-1 |
| 15 | Steps | Outcome | Bruno Bizzozero-Peroni, 2024 | Systematic review and meta-analysis | 33 studies for SR (and 29 for MA) | Depression (reported as diagnosis or number of symptoms) | Higher daily step counts were associated with fewer depressive symptoms in the general adult population. a daily step count of 7000 or higher was associated with lower risk of depression in prospective studies. | most studies lacked robust methods, some publication biases | 10.1001/jamanetworkopen.2024.51208 |
| 16 | Steps | Measurement | Christoph Höchsmann, 2020 | Validation | 1 device (Garmin Vivofit) | reference = research grade accelerometer | Most devices underestimated steps by >20% => good enough for trends but daily values can be off | small n, 3 days | 10.1088/1361-6579/ab635f |

| Nr | Marker | Study Type | Author & Year | Design | Population / Devices | Outcomes / References | Findings | Limitations | DOI |
|---|---|---|---|---|---|---|---|---|---|
| 18 | $VO_2$max | Outcome | Minghui Han, 2022 | Meta Analysis of cohort studies | 34 studies | all-cause mortality, cardiovascular disease, cancer | inverse dose-response associations of CRF with all-cause, CVD and cancer mortality. Per one-metabolic equivalent increase in CRF was associated with 12%, 13% and 7% decreased risk of all-cause, CVD and cancer mortality, respectively | No thresholds of low, moderate and high CRF, selection bias | 10.1136/bjsports-2021-104876 |
| 19 | $VO_2$max | Measurement | Rory Lambe, 2025 | Validation | 1 device (apple Watch) | reference = indirect calorimetry | Apple Watch underestimated VO2 max, with a mean difference of 6.07 mL/kg/min. However, Apple Watch as an alternative to conventional VO2 max prediction from submaximal exercise is warranted | sample bias because all had good cardiorespiratory fitness and also young => limits generalisability | 10.1371/journal.pone.0323741 |

## Appendix D. Survey



# Health Metrics in Daily Life

- Short, anonymous survey for working adults
- No personal data collected. You can stop any time
- Please answer about the **last 4 weeks**
- Takes **3-4 minutes**

* Indicates required question

Consent *

○ I agree to take part

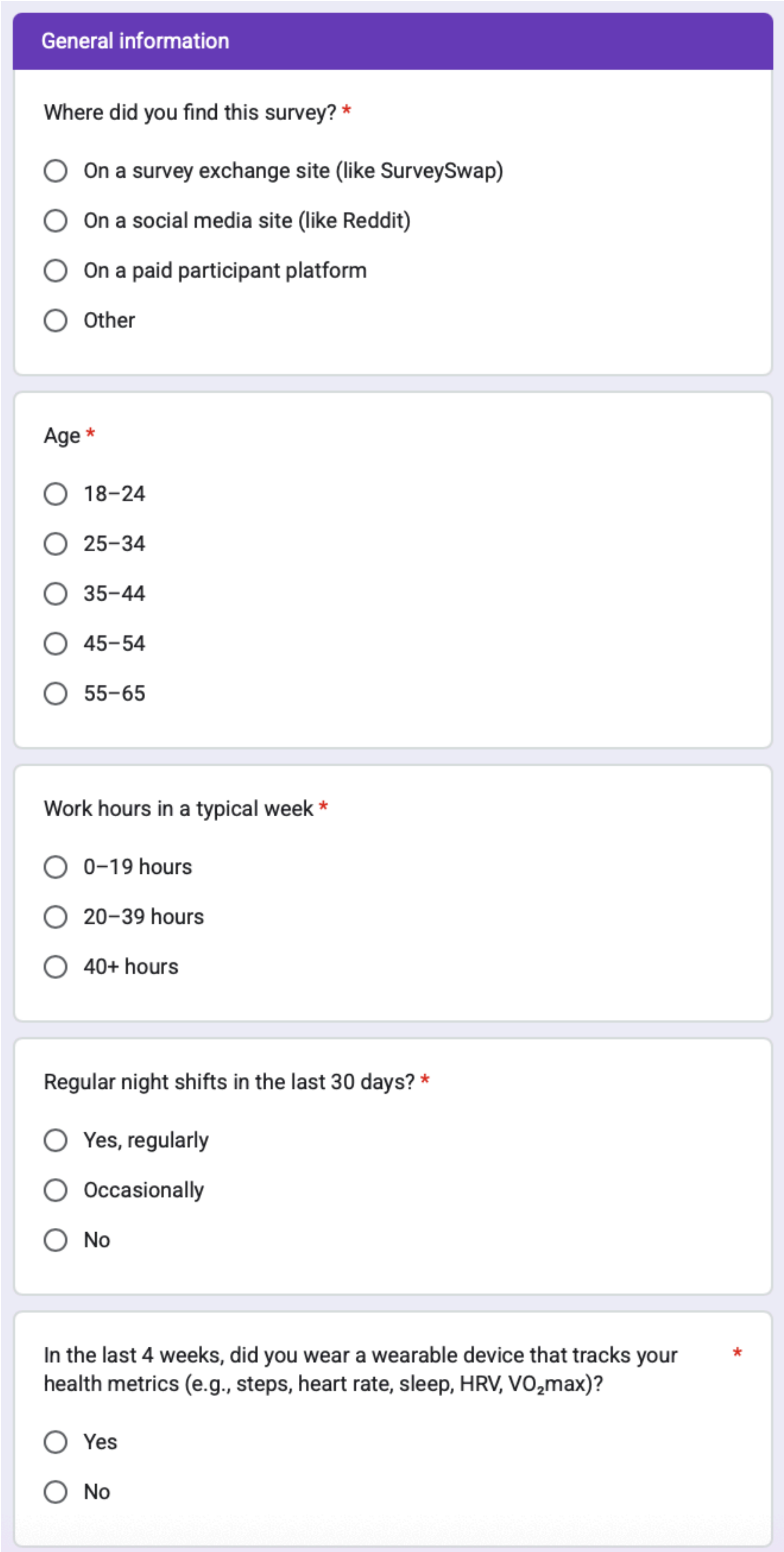
○ I do not agree

Next

Page 1 of 5

Clear form

*Figure 4: Survey questionnaire, page 1*



General information

Where did you find this survey? *

- On a survey exchange site (like SurveySwap)
- On a social media site (like Reddit)
- On a paid participant platform
- Other

Age *

- 18–24
- 25–34
- 35–44
- 45–54
- 55–65

Work hours in a typical week *

- 0–19 hours
- 20–39 hours
- 40+ hours

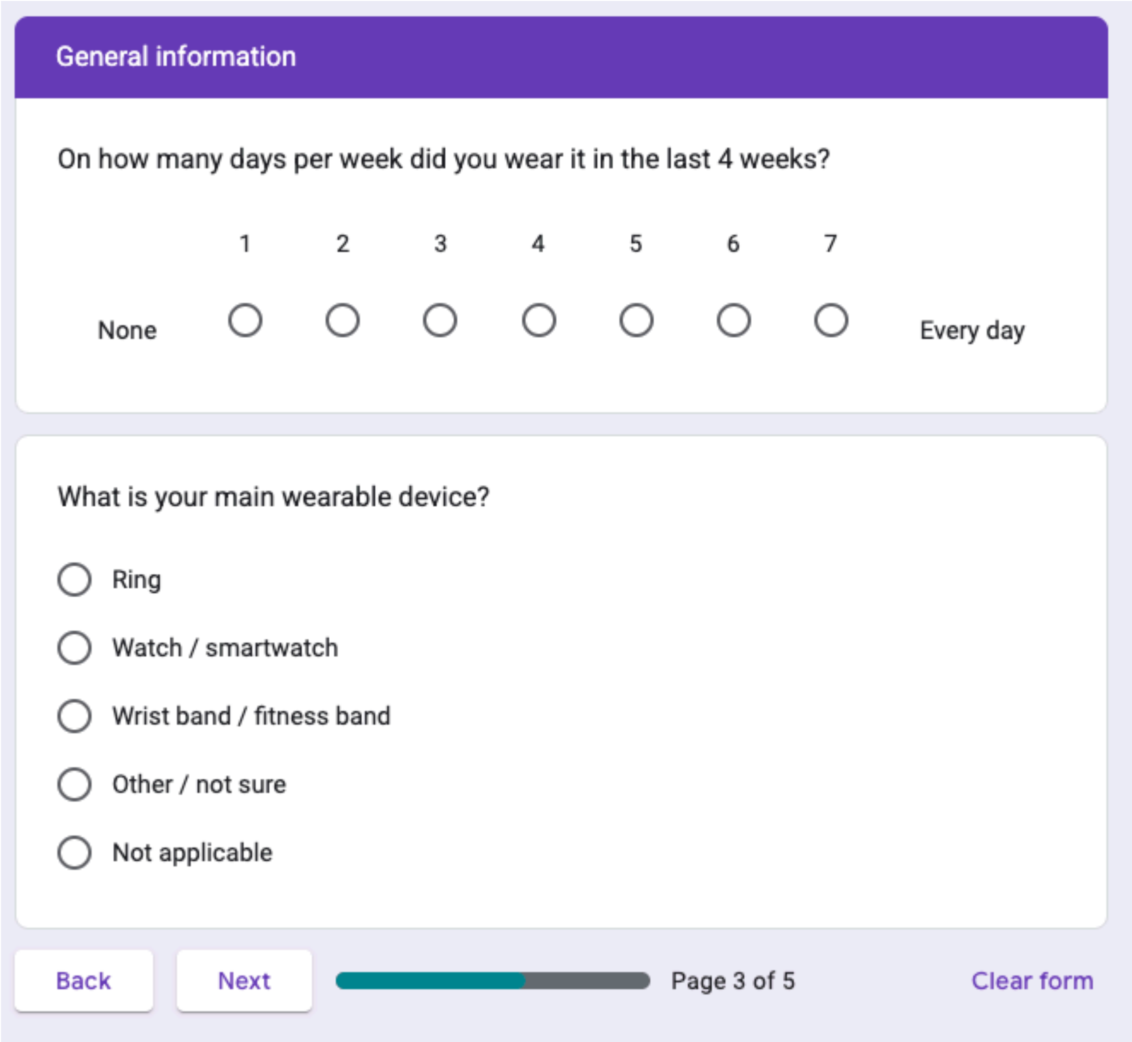
Regular night shifts in the last 30 days? *

- Yes, regularly
- Occasionally
- No

In the last 4 weeks, did you wear a wearable device that tracks your health metrics (e.g., steps, heart rate, sleep, HRV, $VO_2$max)? *

- Yes
- No

*Figure 5: Survey questionnaire, page 2*


General information

On how many days per week did you wear it in the last 4 weeks?

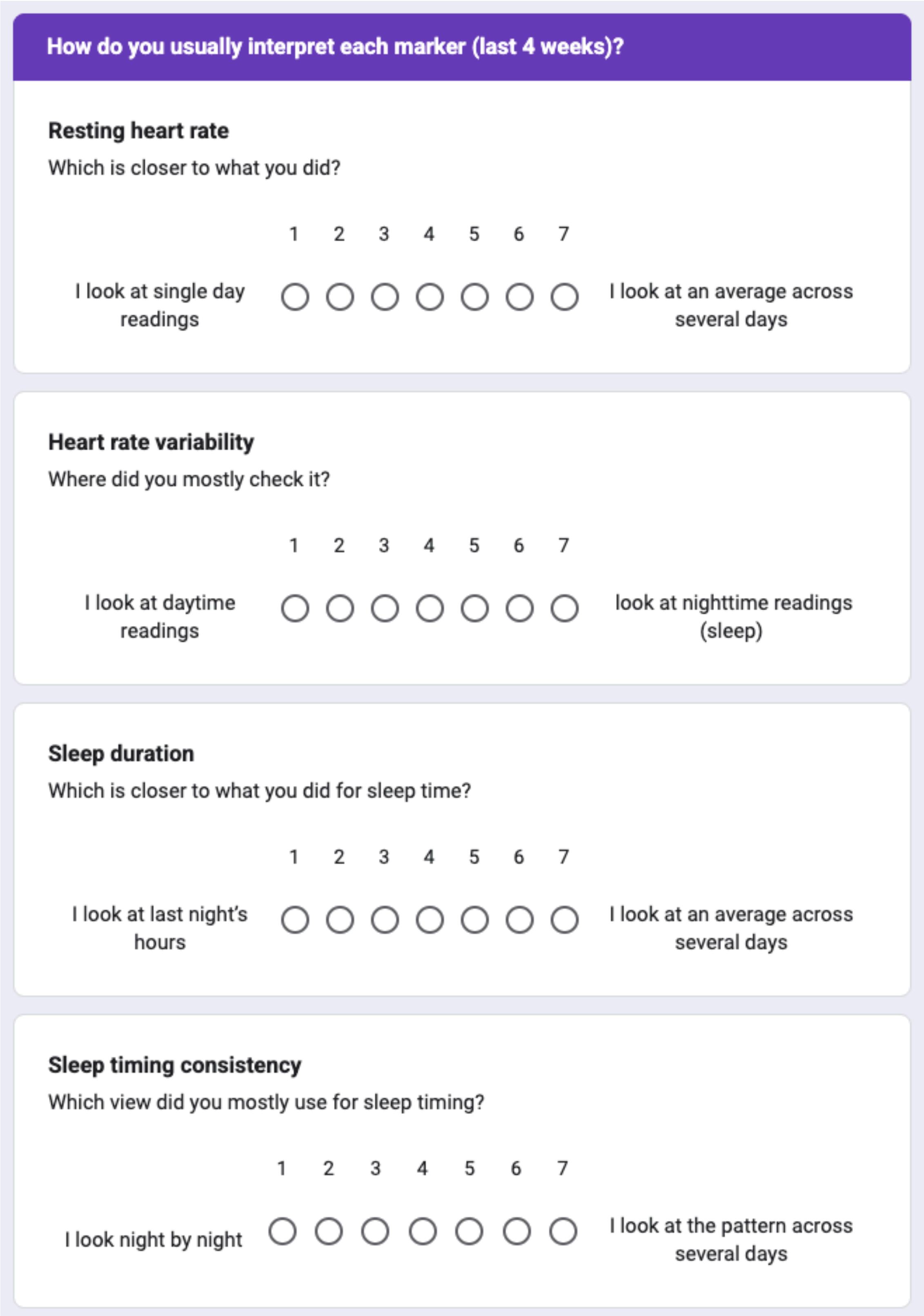
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| None | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Every day |

What is your main wearable device?

- ○ Ring
- ○ Watch / smartwatch
- ○ Wrist band / fitness band
- ○ Other / not sure
- ○ Not applicable

Back Next Page 3 of 5 Clear form

*Figure 6: Survey questionnaire, page 3*



**How do you usually interpret each marker (last 4 weeks)?**

**Resting heart rate**

Which is closer to what you did?

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| I look at single day readings | ○ | ○ | ○ | ○ | ○ | ○ | ○ | I look at an average across several days |

**Heart rate variability**

Where did you mostly check it?

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| I look at daytime readings | ○ | ○ | ○ | ○ | ○ | ○ | ○ | look at nighttime readings (sleep) |

**Sleep duration**

Which is closer to what you did for sleep time?

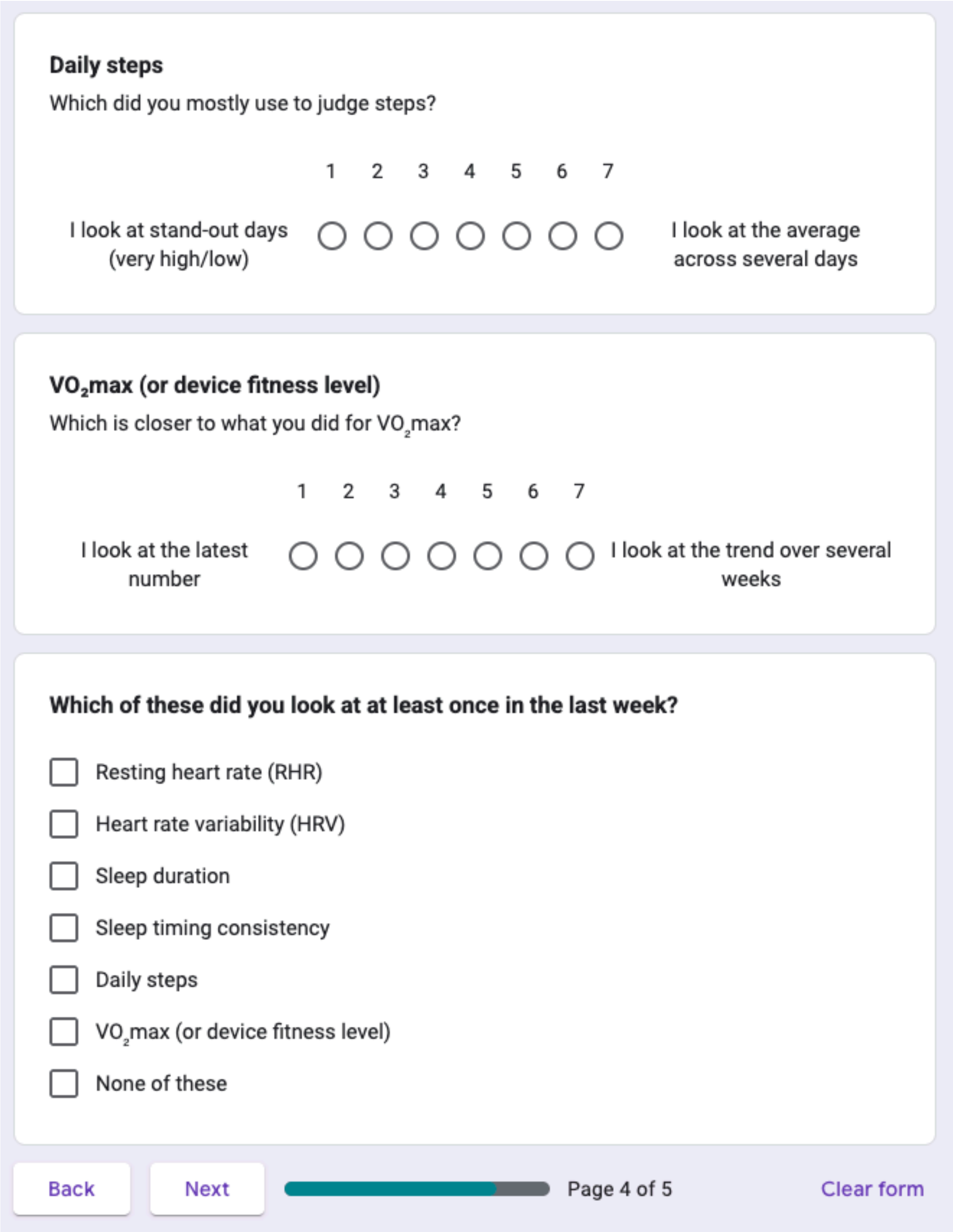
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| I look at last night's hours | ○ | ○ | ○ | ○ | ○ | ○ | ○ | I look at an average across several days |

**Sleep timing consistency**

Which view did you mostly use for sleep timing?

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| I look night by night | ○ | ○ | ○ | ○ | ○ | ○ | ○ | I look at the pattern across several days |

*Figure 7: Survey questionnaire, page 4*



**Daily steps**

Which did you mostly use to judge steps?

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| I look at stand-out days (very high/low) | ○ | ○ | ○ | ○ | ○ | ○ | ○ | I look at the average across several days |

**$VO_2$max (or device fitness level)**

Which is closer to what you did for $VO_2$max?

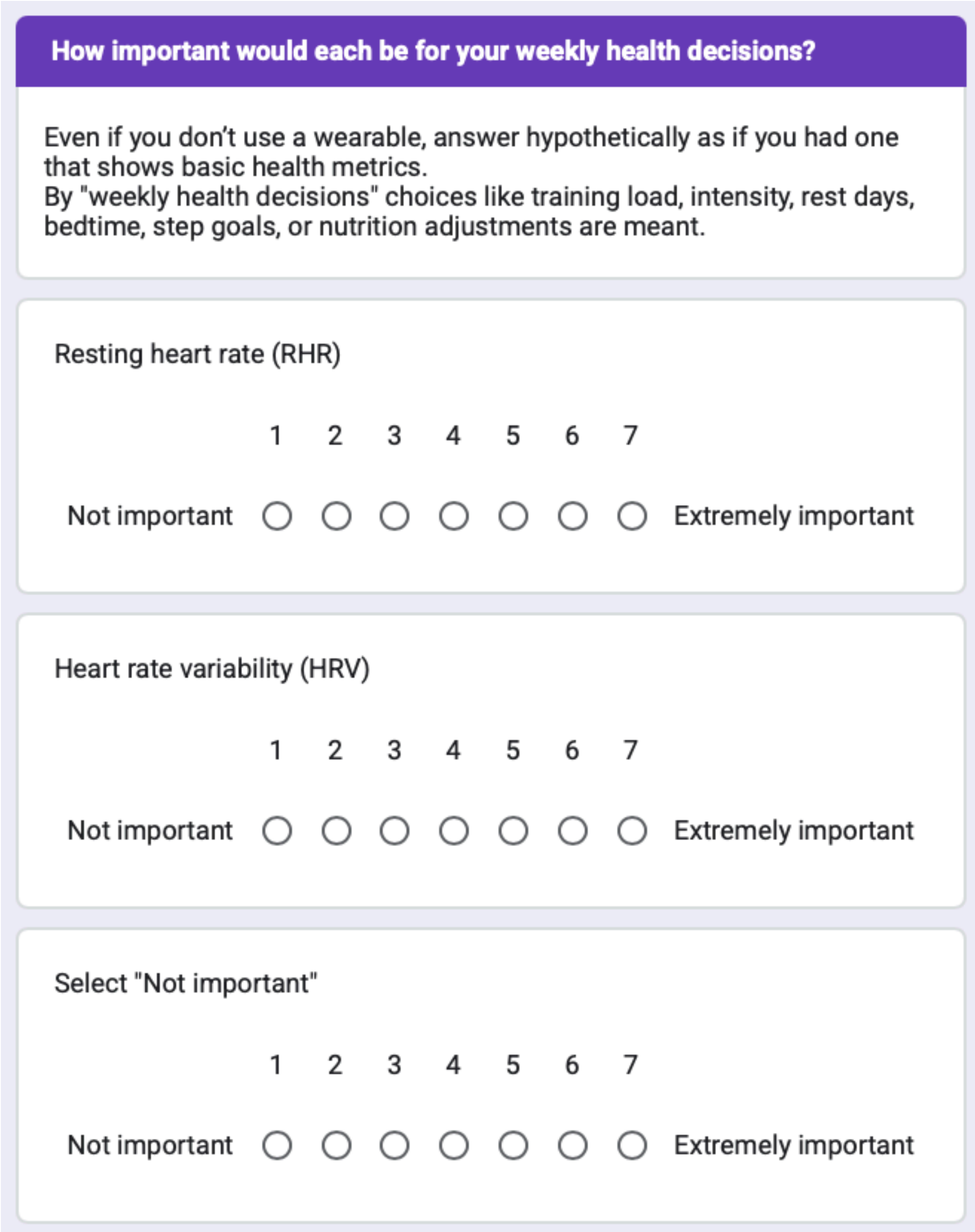
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| I look at the latest number | ○ | ○ | ○ | ○ | ○ | ○ | ○ | I look at the trend over several weeks |

**Which of these did you look at at least once in the last week?**

- ☐ Resting heart rate (RHR)
- ☐ Heart rate variability (HRV)
- ☐ Sleep duration
- ☐ Sleep timing consistency
- ☐ Daily steps
- ☐ $VO_2$max (or device fitness level)
- ☐ None of these

Back Next Page 4 of 5 Clear form

*Figure 8: Survey questionnaire, page 4*



**How important would each be for your weekly health decisions?**

Even if you don't use a wearable, answer hypothetically as if you had one that shows basic health metrics.
By "weekly health decisions" choices like training load, intensity, rest days, bedtime, step goals, or nutrition adjustments are meant.

Resting heart rate (RHR)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Not important | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Extremely important |

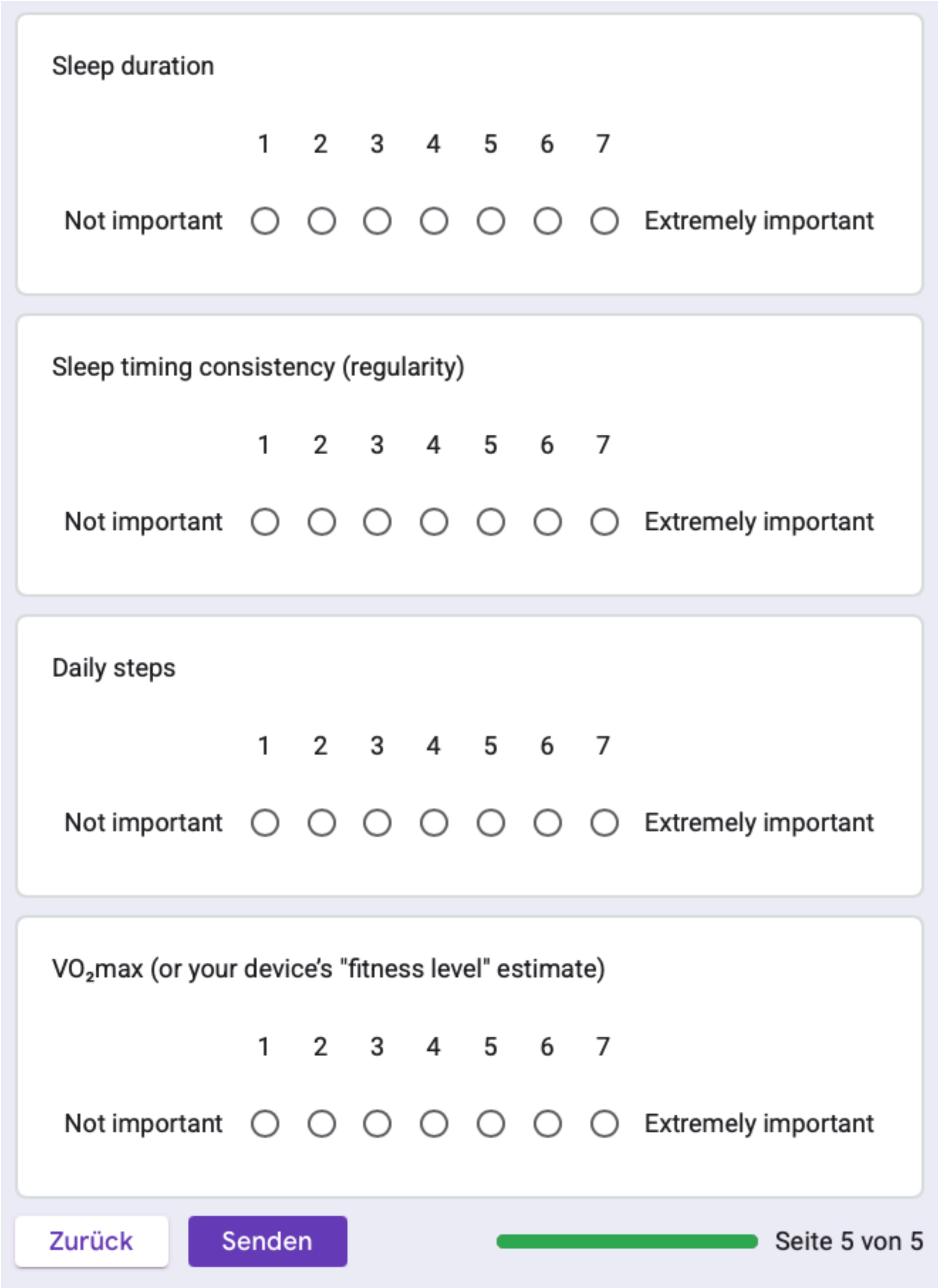
Heart rate variability (HRV)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Not important | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Extremely important |

Select "Not important"

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Not important | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Extremely important |

*Figure 9: Survey questionnaire, page 4*



Sleep duration

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Not important | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Extremely important |

Sleep timing consistency (regularity)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Not important | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Extremely important |

Daily steps

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Not important | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Extremely important |

$VO_2$max (or your device's "fitness level" estimate)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|---|---|---|---|---|---|---|---|---|
| Not important | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Extremely important |

Zurück
Senden
Seite 5 von 5

*Figure 10: Survey questionnaire, page 5*