Smartwatch Health Metrics and Health Anxiety Among Middle-Aged Urban Professionals in India: An Evidence-Based Metric Anxiety Hierarchy for Dashboard Design

Ayushi Saxena[1], Pranjal Sinha[2], Prashanth Cherian Kochuveetil[3]
[1,2] *Student, Department of Design, National Institute of Fashion Technology (NIFT), Bengaluru*
[3] *Professor, Department of Design, National Institute of Fashion Technology (NIFT), Bengaluru*

***Abstract*—Smartwatches have evolved into continuous health monitoring systems, exposing middle-aged urban professionals to real-time biometric data that may amplify health anxiety and compulsive checking behaviours. This study investigates which smartwatch health metrics — Heart Rate, Heart Rate Variability (HRV), SpO2, Sleep Score, and Stress Score — are the strongest predictors of health anxiety intensity among urban Indian professionals aged 40–55, with the goal of producing an evidence-based metric anxiety hierarchy to guide psychologically informed dashboard design. A quantitative descriptive survey was administered to 46 eligible participants across India, using adapted versions of the Short Health Anxiety Inventory (SHAI), Cyberchondria Severity Scale (CSS-12), and Commitment to Self-Tracking scale (C2ST). Spearman Rank Correlation served as the primary inferential analysis, supplemented by exploratory Multiple Linear Regression. Four of five metrics demonstrated statistically significant positive associations with health anxiety after Bonferroni correction. The resulting metric anxiety hierarchy ranked Stress Score first (avg $|r_s|$ = 0.562), followed by Sleep Score (0.525), HRV (0.518), and SpO2 (0.507). Heart Rate, despite being the most frequently monitored metric, produced no significant anxiety association, a finding attributed to its cultural familiarity and interpretive clarity. Regression analysis revealed that the five metrics collectively explained 56.4% of variance in health anxiety outcomes. This study introduces the metric anxiety hierarchy as an empirical construct, challenges the reassurance hypothesis of self-tracking, and extends the cyberchondria framework to wearable-generated biometric data. Findings provide actionable design recommendations, including progressive disclosure of complex metrics and contextual anchoring of algorithmically derived scores, to reduce monitoring-induced psychological distress.**

***Index Terms*—Smartwatch health metrics; Health anxiety; Middle-aged urban professionals; Health dashboard design**

I. AIM

The primary aim of this research is to quantitatively understand how the smartwatch health metrics impact health anxiety levels among middle-aged urban professionals in India and establish a hierarchical (anxiety-induced) framework to aid psychologically informed health dashboard design.

II. OBJECTIVES

1) To identify the frequency and patterns of metric interaction: Determine which health metric data points (HR, HRV, SpO2, Sleep Scores, etc.) are most frequently monitored by the 40–55 age demographic in Bangalore.
2) To assess perceived frequency and self-reported anxiety associated with metric variation.
3) To develop a Metric Anxiety Hierarchy: Rank biometric metrics based on their predictive power for health anxiety, distinguishing between "low-stress" (informative) and "high-stress" (triggering) data.

III. INTRODUCTION

The interest in using smartwatch as a health wearable has shown rapid growth in the past years. Smartwatches are a category of health wearables that can help individuals take charge of their well-being and improve their health [1]. They have rapidly evolved from fitness accessories into continuous self-monitoring health systems, providing users with real-time access to cardiac, respiratory, and sleep-related metrics.

Global adoption studies show increasing wearable use among adults aged 40–55, a demographic simultaneously entering higher cardiometabolic risk years and becoming more proactive about preventive health [2]. However, smartwatch usage can induce certain negative effects, like anxiety and dependence on the middle-aged demographic [3]. Research in digital health and human–computer interaction indicates that frequent exposure to physiological data increases self-surveillance behaviors and perceived health vigilance. While self-tracking can enhance self-efficacy, it can also intensify bodily monitoring and lead to wrong health metrics interpretation, particularly when data is ambiguous or fluctuating within normal and poor ranges [4].

Among all other health metrics that middle-aged urban demographic tracks, the most common ones are Heart Rate, HRV, SpO2, Stress Scores and Sleep Scores. These are tracked by almost every smartwatch, irrespective of the model or brand [5]. Individuals prone to health anxiety tend to misinterpret benign bodily sensations as indicators of illness (phenomenon called Hypochondriasis). Continuous biometric feedback through smartwatches may amplify this interpretative bias by making normal physiological variability highly visible and numerically salient. Emerging empirical studies on wearable use report that abnormal alerts and metric fluctuations can provoke worry, reassurance-seeking, and repeated checking behaviors, especially in populations already attentive to cardiovascular health risk [6].

Over-reliance on smartwatches can lead the older demographic to constantly seek out new information about their health concerns, leading to increased anxiety and distress. By quantitatively identifying which smartwatch health and fitness metrics are the strongest predictors of health anxiety intensity in this demographic, we may produce an evidence-based metric hierarchy that can guide more psychologically informed dashboard design.

The paper explores possible links between frequent tracking of smartwatch health metrics and its impact on the physiological well-being of a user. The research question was: Which smartwatch health metrics (HR, HRV, SpO2, sleep score, stress score) are the strongest predictors of health anxiety intensity among middle-aged urban professionals (40–55) in India? The purpose of the study is to produce evidence of an anxiety hierarchy of health metrics that will give designers (1) justification to soften or contextualize high-anxiety metrics (e.g., adding reassuring language) and (2) a framework for progressive disclosure (gate high-anxiety metrics behind intentional user action).

IV. HYPOTHESIS

A. Theoretical Grounding

Traditional health monitoring is infrequent and clinically supervised, but consumer-grade wearables have introduced continuous, on-demand access to physiological data — a structural shift that heightens somatic self-surveillance without providing the interpretive scaffolding that clinical settings offer [7]. While self-tracking can enhance self-efficacy, technologies that make normal physiological fluctuations numerically visible can intensify bodily monitoring and lead to misinterpretation of data, particularly when readings fall within ambiguous or borderline ranges [4].

The existing literature offers a specific basis for predicting which metrics are most likely to amplify anxiety. HRV is the most empirically studied metric in this context. Jo et al. [8], in a four-week longitudinal study using smartwatch-derived PPG signals, demonstrated statistically significant associations between HRV parameters — specifically SDNN, SDSD, and LF — and self-reported anxiety scores measured by the GAD-7. This association emerged from passive monitoring of consumer-grade devices, establishing that the act of observing HRV data itself co-occurs with anxiety outcomes. This finding is further supported by a systematic review by Hickey et al. [9], which identified HRV as the primary physiological parameter for anxiety detection in wearable technology research, noting that its interpretive complexity contributes to its psychological salience.

SpO2 carries comparable interpretive demand. Unlike heart rate, which users contextualise through decades of cultural familiarity with BPM as a health concept, SpO2 readings require users to apply a clinical threshold — the 95–100% normal range — to make sense of any given value. Minor deviations below this threshold carry genuine clinical significance (hypoxemia), and most smartwatch displays do not provide contextual framing that disambiguates clinically meaningful drops from sensor noise or positional artefacts. This absence of interpretive context, combined with SpO2's clinical salience for respiratory and cardiovascular conditions, establishes it as a plausible anxiety trigger alongside HRV.

Sleep Score introduces a different but equally relevant mechanism. Baron et al. [10], who coined the clinical term orthosomnia in the Journal of Clinical Sleep Medicine, documented the first case series of patients seeking treatment for self-diagnosed sleep disorders arising directly from sleep tracker data. Their patients demonstrated compulsive checking behaviors and sleep-related perfectionism driven by the numerical output of wearable scores. A subsequent cross-sectional prevalence study [11] confirmed this pattern in a general population sample of 523 participants, finding that sleep tracker ownership co-occurred with elevated anxiety and sleep preoccupation scores.

The cyberchondria literature provides the overarching theoretical frame. In a systematic review and meta-analysis, Schenkel et al. [12] established that cyberchondria is best conceptualized as a safety behavior in which repeated information-seeking amplifies rather than resolves health concern. A subsequent review of wearable devices [13] explicitly extended this framework to biometric monitoring, noting that minor fluctuations in smartwatch-recorded metrics within normal ranges are perceived as indicators of serious health issues.

B. Hypotheses

Null Hypothesis (H0): The monitoring frequency of no smartwatch health metric — Heart Rate, HRV, SpO2, Sleep Score, or Stress Score — will demonstrate a statistically significant Spearman rank correlation with health anxiety intensity, as measured by SHAI-5 and CSS-12 composite scores, among urban Indian professionals aged 40–55.

Research Hypothesis (H1): Among urban Indian professionals aged 40–55, the monitoring frequency of HRV and SpO2 will demonstrate significantly stronger Spearman rank correlations with health anxiety intensity — as measured by SHAI-5 and CSS-12 composites — compared to aggregated performance metrics such as Sleep Score and Stress Score.

This prediction is grounded in two convergent mechanisms. First, HRV and SpO2 are raw physiological signals whose values fluctuate in ways that are clinically meaningful but interpretively inaccessible to non-clinical users, making them more likely to trigger anxiety when readings appear abnormal [8, 9]. Second, composite scores such as Sleep Score and Stress Score, while algorithmically derived, present a single summarised number that reduces the moment-to-moment volatility of raw biofeedback — and are therefore predicted to carry a lower anxiety burden.

V. METHOD

This study combines a systematic review of current literature with a quantitative survey, using a quantitative descriptive survey design grounded on the following scales: (1) Short Health Anxiety Inventory (SHAI), (2) Cyberchondria (CSS-12), (3) Commitment to Self-Tracking (C2ST). While the quantitative sections aim to gauge the frequency and intensity of checking smartwatch metrics and its impact on the user, the subjective type questions aim to understand how tracking smartwatch health metrics has changed the relationship between one's health or body over a period.

A. Participants

A total of 117 random participants were recruited from all over India, from states like Bangalore, New Delhi, and Jharkhand. Random sampling was employed to minimize selection bias. The sample consisted of urban professionals of varying ages, with no constraints regarding gender or degree of education. An online survey distributed through Google Forms made it possible to reach a wide audience. Of the total 117 participants, only 58 owned a smartwatch, further out of which only 46 measured any of the following health metrics – HRV, Heart Rate, SpO2, Sleep Score & Stress Score. The 46 participants' ages ranged between 40–55. The most frequently checked metric was Heart Rate (M=2.80, SD=1.08), followed by Sleep Score (M=2.38, SD=1.25).

VI. CONSTRUCTS

The questionnaire started with a brief introduction that described the aim of the study and asked three filter questions: (1) Whether the participant is aged between 40–55, (2) Whether the participant owns a smartwatch, (3) Whether the participant measures any of these health metrics – HRV, Heart Rate, SpO2, Sleep Score, Stress Score.

The next section guided participants to report how often they checked each metric on their smartwatches using a five-point Likert Scale (1 = "Never", 5 = "Very Often"). The fifth section uses the Commitment to Self-Tracking (C2ST) scale — a 12-item, behavior-based tool measuring intensity of dedication to self-tracking technologies. The sixth section uses an adaptation of the Short Health Anxiety Inventory (SHAI) to assess health anxiety independently of actual physical health status. The seventh section uses an adaptation of CSS-12 (Cyberchondria Severity Scale, 12-item) measuring: (1) Excessiveness: repetitive and escalating search behavior; (2) Compulsion: extent to which searching interferes with work or social life; (3) Distress: emotional panic triggered by search results; (4) Reassurance Seeking: urge to consult medical professionals based on online findings. The eighth and final section consists of two subjective questions.

TABLE I. Questionnaire Items

Items	Question / Item Description	Mean	SD
1	Are you aged between 30–55 years?	NA	NA
2	Do you currently own a smartwatch?	NA	NA
3	Health metrics measured (HR, HRV, SpO2, etc.)	NA	NA
4	Which smartwatch brand do you own?	NA	NA
5	How long have you been using a smartwatch?	NA	NA
6	Chronic health condition (Yes/No)	NA	NA
7	Monitoring Frequency [Heart Rate]	2.80	1.07
8	Monitoring Frequency [HRV]	2.13	1.07
9	Monitoring Frequency [SpO2]	2.15	1.07
10	Monitoring Frequency [Sleep Score]	2.39	1.24
11	Monitoring Frequency [Stress Score]	2.15	1.28
12	C2ST [Must check health metrics daily]	2.57	1.11
13	C2ST [Feel unsettled if not checked]	1.93	1.32
14	C2ST [Influence on daily decisions]	1.98	1.11
15	C2ST [Check even if readings are normal]	2.24	1.12
16	C2ST [Compare current with past data]	2.20	1.20
17	C2ST [Loss of control if stopped for a week]	2.02	1.34
18	Worry Frequency (triggered by reviewing data)	2.28	1.09
19	Response to unexpected physical sensation	NA	NA
20	Feeling when reading is outside range	NA	NA
21	Ease of setting aside health concerns	NA	NA
22	Health-related worry interference	1.85	1.13
23	CSS [More anxious than reassured]	2.07	1.16
24	CSS [Checking tends to escalate]	2.17	1.08
25	CSS [Recheck reading after some time]	2.65	1.20
26	CSS [Monitoring more than intended]	2.26	1.12
27	CSS [Search online for meaning]	2.63	1.48
28	CSS [Need to consult doctor for reassurance]	2.26	1.41
29	Metric causing most concern (Qualitative)	NA	NA
30	Change in relationship with health (Qualitative)	NA	NA

A. Procedure

The study was conducted entirely online. The questionnaire was completed using a Google Form, and the survey link was spread across WhatsApp and email channels. Informed consent was obtained electronically before participants proceeded. The survey took approximately 7–8 minutes to complete. Data collection occurred over 19 days after which responses were compiled using Google Sheets. The dataset was cleaned to remove incomplete or inconsistent responses. The cleaned dataset was then analyzed using AI software (Claude and Gemini) to identify trends and patterns.

VII. DATA ANALYSIS

The dataset was analyzed using Claude and cross-checked with Gemini. Descriptive statistics were used to summarize the prevalence of most anxiety-inducing smartwatch health metrics.

A. Sample Demographics

The sample is dominated by long-term smartwatch users, primarily using premium brands. Apple tops the chart with 41.30% using it, while Noise stays at the bottom with 6.52%. 82.50% of users have been using a smartwatch for over a year. 21.74% of people have been diagnosed with chronic health conditions.

B. Metric Monitoring Frequency

Participants check Heart Rate significantly more often than more complex metrics like HRV or SpO2.

TABLE II. Metric Monitoring Frequency

Metric	Mean	Median	Mode	Std. Deviation
Heart Rate	2.8	3	3	1.07
Sleep Score	2.39	2	1	1.24
SpO2	2.15	2	1	1.07
Stress Score	2.15	2	1	1.28
HRV	2.13	2	1	1.07

C. Commitment to Self-Tracking Scale

The overall scale mean of 2.16 indicates a low-to-moderate commitment to self-tracking data.

TABLE III. C2ST Scale Results

Item Description	Mean	SD
1. Must check metrics every day	2.57	1.11
2. Feel unsettled if not checked	1.93	1.32
3. Influence on daily decisions	1.98	1.11
4. Check even when readings are normal	2.24	1.12
5. Compare current data with past data	2.20	1.20
6. Loss of control if usage stopped	2.02	1.34
Overall C2ST Scale Mean	2.16	1.20

D. CSS-12

The scale mean of 2.34 indicates mild cyberchondria tendencies, with “rechecking” and “online searching” being the most common behaviors.

TABLE IV. CSS-12 Scale Results

Item Description	Mean	SD
1. Feeling more anxious than reassured	2.07	1.16
2. Checking escalates to more metrics	2.17	1.08
3. Rechecking the reading after some time	2.65	1.20
4. Monitoring longer than intended	2.26	1.12
5. Online searching for meanings	2.63	1.48
6. Need to consult a doctor for reassurance	2.26	1.41
Overall CSS Scale Mean	2.34	1.24

E. Key Findings

1) Heart Rate is the most frequently monitored metric, suggesting it is the primary “anchor” for health status in this demographic.
2) Most users are deeply habituated to their devices; however, commitment (C2ST 2.16) remains moderate, not yet fully compulsive.
3) When an unusual reading occurs, the most common behavioral outcome is delayed rechecking rather than immediate panic.
4) A plurality of users uses the smartwatch specifically as a tool to mitigate anxiety through immediate reassurance.
5) While generally mild, the link between “unusual readings” and “online searching” (2.63) presents a potential pathway to cyberchondria.

VIII. STATISTICAL ANALYSIS

This study employed a two-stage analytical approach. Spearman Rank Correlation (r_s) served as the primary inferential test, with Multiple Linear Regression conducted as a secondary exploratory analysis.

A. Primary Analysis: Spearman Rank Correlation

All independent variables (monitoring frequency for HR, HRV, SpO2, Sleep Score, and Stress Score) and the dependent variable (health anxiety) are measured on ordinal Likert scales (1–5). Spearman Rank Correlation is the appropriate non-parametric alternative, operating on the rank ordering of values rather than their absolute magnitudes [14]. Five Spearman correlations were computed per outcome variable, yielding fifteen tests in total. A Bonferroni correction was applied, producing an adjusted significance threshold of $p < 0.01$.

B. Secondary Analysis: Mann-Whitney U Test

Two independent group comparisons were conducted using the Mann-Whitney U test. The first compared health anxiety and self-tracking outcomes between

Apple Watch users (n=19) and non-Apple users (n=27). The second compared outcomes between users with a diagnosed chronic health condition (n=10) and those without (n=36). Shapiro-Wilk normality tests confirmed that all outcome composites violated normality (all $p < 0.01$), confirming Mann-Whitney U as the correct test.

C. Tertiary Analysis: Multiple Linear Regression

Multiple Linear Regression was conducted as a secondary analysis to examine whether the bivariate associations observed in Spearman hold when all five metrics are entered as simultaneous predictors. This addresses a limitation inherent in running five separate Spearman correlations: each tests a metric in isolation and cannot account for the shared variance between correlated predictors. With N = 46 and five predictors, the model is underpowered relative to the conventional rule of thumb of 10–20 observations per predictor [15]. These results are therefore treated as exploratory and directional rather than confirmatory.

D. Composite Score Construction

The SHAI composite was constructed as a 5-item mean comprising two original Likert items (worry frequency, Col 18; life interference, Col 22) and three categorically converted ordinal items. The CSS-12 partial composite was computed as the mean of 6 items (Cols 23–28), and the C2ST partial composite as the mean of 6 items (Cols 12–17).

IX. RELIABILITY ANALYSIS: CRONBACH'S ALPHA

Internal consistency of all Likert-scale instruments was assessed using Cronbach's Alpha (α) before proceeding to hypothesis testing. Values above 0.70 are considered acceptable, above 0.80 good, and above 0.90 excellent [16, 17].

TABLE V. Cronbach's Alpha by Scale (N = 46)

Scale	Items (k)	α	Interpretation
Metric Monitoring Frequency	5	0.876	Good
C2ST — Self-Tracking Commitment	6	0.905	Excellent
SHAI — Health Anxiety (proxy)	5	0.815	Good
CSS-12 — Cyberchondria	6	0.895	Good–Excellent

All four scales demonstrated good to excellent internal consistency. The C2ST scale achieved the highest alpha (α = 0.905). The SHAI composite warrants specific discussion: the 5-item proxy achieved α = 0.815, indicating good internal consistency despite the mixed item format. However, two important caveats apply: the scoring conversion introduces an element of researcher judgment, and the 5-item proxy does not carry the full construct validity of the validated 18-item SHAI.

X. NORMALITY OF DATA

Normality of all study variables was assessed using the Shapiro-Wilk test, which is recommended for samples of $N < 50$ [14]. Results indicated that all eight variables violated the assumption of normality (all $W < 0.93$, all $p < 0.01$). All distributions were positively skewed (skewness range: +0.36 to +1.03). Kurtosis values were near zero or mildly negative across all variables. These results confirm that the use of non-parametric statistical tests — Spearman Rank Correlation and Mann-Whitney U — was appropriate throughout this study.

XI. INFERENCES AND FINDINGS

A. Composite Score Descriptives

The 5-item SHAI composite yielded a mean of 1.95 (SD = 0.75, range: 1.0–4.0), indicating low-to-moderate health anxiety in the sample. The CSS-12 partial composite produced a mean of 2.34 (SD = 1.01, range: 1.0–5.0), reflecting mild cyberchondria tendencies. The C2ST partial composite yielded a mean of 2.16 (SD = 0.99, range: 1.0–5.0), indicating low-to-moderate self-tracking commitment.

B. Primary Analysis — Spearman Correlations

TABLE VI. Spearman Correlations: Metric Monitoring Frequency vs Outcome Composites (N = 46)

Metric	SHAI r_s	p	Sig*	CSS-12 r_s	p	Sig*	C2ST r_s	p	Sig*
Heart Rate	+0.303	0.041	Marginal	+0.163	0.281	NO	+0.473	0.001	YES
HRV	+0.491	<0.001	YES	+0.454	0.002	YES	+0.599	<0.001	YES
SpO2	+0.426	0.003	YES	+0.520	<0.001	YES	+0.521	<0.001	YES
Sleep Score	+0.485	<0.001	YES	+0.434	0.003	YES	+0.727	<0.001	YES
Stress Score	+0.516	<0.001	YES	+0.501	<0.001	YES	+0.668	<0.001	YES

**Bonferroni corrected threshold: p < 0.01. All correlations are positive in direction.*

Eight of ten primary anxiety correlations (SHAI + CSS-12) are statistically significant after Bonferroni correction. Heart Rate is the single metric that does not reach significance against either anxiety outcome.

C. Metric Anxiety Hierarchy

Metrics were ranked by their average absolute Spearman coefficient across the two primary anxiety outcomes — SHAI and CSS-12. C2ST correlations are excluded from the hierarchy ranking, as C2ST measures self-tracking commitment rather than anxiety.

TABLE VII. Metric Anxiety Hierarchy (Primary Finding)

Rank	Metric	SHAI-5 r_s	CSS-12 r_s	Avg $\lvert r_s \rvert$	Classification
1	Stress Score	+0.516	+0.501	0.509	High-Stress Trigger
2	HRV	+0.491	+0.454	0.473	High-Stress Trigger
3	SpO2	+0.426	+0.520	0.473	High-Stress Trigger
4	Sleep Score	+0.485	+0.434	0.460	Moderate Trigger
5	Heart Rate	+0.303	+0.163	0.233	Low / Informative

Classification thresholds: avg $\lvert r_s \rvert$ > 0.47 = High-Stress Trigger; 0.30–0.46 = Moderate Trigger; < 0.30 = Low/Informative.

D. Secondary Analysis — Multiple Linear Regression

DV: SHAI Composite — $R^2 = 0.564$, Adjusted $R^2 = 0.510$, $F(5,40) = 10.35$, $p < 0.001$

TABLE VIII. MLR Results — DV: SHAI Composite

Predictor	β (standardised)	p	Significant ?
HRV	+0.396	0.021	YES
Sleep Score	+0.292	0.088	NO
SpO2	+0.253	0.091	NO
Heart Rate	−0.220	0.111	NO
Stress Score	+0.067	0.712	NO

DV: CSS-12 Composite — $R^2 = 0.443$, Adjusted $R^2 = 0.374$, $F(5,40) = 6.37$, $p < 0.001$

TABLE IX. MLR Results — DV: CSS-12 Composite

Predictor	β (standardised)	p	Significant ?
SpO2	+0.448	0.010	YES
Heart Rate	−0.319	0.044	YES
HRV	+0.243	0.200	NO
Stress Score	+0.163	0.428	NO
Sleep Score	+0.092	0.630	NO

The regression models explain 56.4% and 44.3% of variance in SHAI and CSS-12 respectively. At the individual predictor level, most metrics lose significance once shared variance is controlled — a direct consequence of the moderate-to-high inter-correlations between predictors (Sleep × Stress r_s = 0.752) and the limited sample size.

E. Reconciling Spearman and Regression

The most important finding that emerges from running both analyses is the divergence of Stress Score between the two tests. In Spearman, Stress Score ranks first in the anxiety hierarchy (avg $\lvert r_s \rvert$ = 0.509). In regression against SHAI, its standardised beta collapses to +0.067 and becomes non-significant (p =

0.712). This is a statistical consequence of shared variance: Stress Score and Sleep Score correlate strongly with each other (r_s = 0.752). Stress Score's association with anxiety is real and significant in isolation but is not uniquely separable from Sleep Score's contribution when both are modelled together.

F. Additional Analyses

Brand Comparison — Apple vs Non-Apple (Mann-Whitney U): An exploratory comparison did not yield statistically significant differences on any outcome measure (SHAI-5: U=332, p=0.090; CSS-12: U=306, p=0.268; C2ST: U=305, p=0.283). Apple Watch users showed consistently higher median scores across all three outcomes but differences do not reach statistical significance at the current sample size.

Chronic Condition vs Anxiety (Mann-Whitney U): Users with a diagnosed chronic health condition (n=10) did not report significantly different anxiety or self-tracking scores compared to users without a diagnosis (n=36) on any outcome measure (SHAI-5: U=176, p=0.925; CSS-12: U=172, p=0.852; C2ST: U=227, p=0.214). This null finding suggests that clinical vulnerability does not significantly amplify smartwatch-induced anxiety in this sample.

Inter-Scale Correlations (Spearman): All three inter-scale correlations were statistically significant and strong in magnitude.

TABLE X. Inter-Scale Correlations (Spearman)

Pair	r_s	p
C2ST vs SHAI-5	+0.705	<0.001
C2ST vs CSS-12	+0.554	<0.001
SHAI-5 vs CSS-12	+0.714	<0.001

The strong correlation between C2ST and both anxiety outcomes indicates that self-tracking commitment co-occurs with health anxiety and cyberchondria in this sample.

G. Key Findings Narrated

H0 is rejected. Eight of ten primary Spearman correlations reached significance after Bonferroni correction, confirming that monitoring frequency is significantly associated with anxiety across four of five metrics.

H1 is not supported as stated. H1 predicted HRV and SpO2 as the strongest correlates of health anxiety. The data shows Stress Score as the strongest associate (avg $|r_s|$ = 0.509), a metric H1 explicitly classified as a weaker, aggregated comparator. HRV and SpO2 ranked second and third, directionally consistent with H1, but the hierarchy predicted was not confirmed.

Finding 1 — Stress Score: Stress Score monitoring frequency produced the highest average correlation across both anxiety outcomes (avg $|r_s|$ = 0.509), with r_s = +0.516 against SHAI-5 ($p < 0.001$) and r_s = +0.501 against CSS-12 ($p < 0.001$). The Stress Score is an algorithmically derived, brand-specific composite index with no standardised clinical reference range. This opacity appears to be driving anxiety.

Finding 2 — Sleep Score: Sleep Score monitoring frequency produced r_s = +0.485 against SHAI-5 ($p < 0.001$) and the highest single correlation in the entire analysis, r_s = +0.727 against C2ST ($p < 0.001$). Sleep is checked every morning, embedding a recurring biometric judgement into the daily routine.

Finding 3 — HRV and SpO2: HRV and SpO2 produced equivalent average anxiety correlations (both avg $|r_s|$ = 0.473). HRV had the lowest mean monitoring frequency (Mean = 2.13) yet co-occurred with high anxiety scores among those who do check it frequently. SpO2 produced the strongest individual correlation against CSS-12 (r_s = +0.520).

Finding 4 — Heart Rate: Heart Rate had the highest mean monitoring frequency (Mean = 2.80) yet produced a marginal correlation with SHAI-5 (r_s = +0.303, p = 0.041) that did not survive Bonferroni correction, and a non-significant correlation with CSS-12 (r_s = +0.163, p = 0.281). Cultural familiarity appears to function as a buffer against anxiety.

Finding 5: For four of five metrics, higher monitoring frequency is consistently associated with higher anxiety. This directly contradicts the assumption that biometric monitoring functions as a reassurance mechanism for this demographic.

XII. MANAGERIAL IMPLICATIONS

The findings carry direct implications for product managers, UX designers, and health technology companies responsible for smartwatch health dashboard design. The following recommendations are oriented toward proactive risk mitigation.

Implication 1: Redesign Stress Score and Sleep Score display as the highest design priority. Both are currently displayed as single composite scores without contextual thresholds, trend baselines, or plain-language interpretation. Designers should introduce contextual anchors such as rolling averages, population percentiles, and direct explanations of what a given score means in practice.

Implication 2: Apply progressive disclosure to HRV and SpO2. Product managers should consider a layered display design: a simple descriptive indicator by default (Normal, Slightly Low, Low), with detailed numerical values accessible only on deliberate tapping.

Implication 3: Use Heart Rate's display design as the benchmark for other metrics. Heart Rate's non-significant anxiety association is not accidental. It has clear units (BPM), intuitive reference ranges (Resting, Active, Elevated), and decades of cultural familiarity. This design approach should be deliberately replicated when redesigning the display of more complex metrics.

Implication 4: Introduce personalised threshold design rather than display customisation alone. The introduction of dynamic personalised thresholds calibrated to the individual's historical baseline would address the interpretive gap that the current data suggests is the primary driver of monitoring-induced anxiety.

Implication 5: The 40–55 urban professional demographic warrants targeted design attention. This cohort sits at the intersection of high health awareness, high device literacy, and elevated baseline health concern. Features that build interpretive confidence and reduce unnecessary checking behaviour are likely to serve this group better than those designed purely to increase time spent on health screens.

Implication 6: Treat self-tracking commitment as a design risk signal. The strong correlation between C2ST and both anxiety outcomes ($r_s = 0.705$ with SHAI-5, $r_s = 0.554$ with CSS-12) suggests that smartwatch platforms could track monitoring behaviour itself as a proxy risk signal, triggering gentle contextual interventions when escalating checking frequency is detected.

XIII. THEORETICAL IMPLICATIONS

Contribution 1: Preliminary evidence for extending the cyberchondria framework to wearable devices. The significant correlations between smartwatch metric monitoring frequency and CSS-12 composite scores (r_s range: 0.43–0.52, all $p < 0.01$) provide preliminary evidence suggesting that the cyberchondria construct may extend to wearable-generated biometric data [12, 13].

Contribution 2: Questioning the reassurance hypothesis of self-tracking for this demographic. The consistent positive directionality of almost all significant correlations suggests that the reassurance hypothesis does not hold when users lack clinical literacy to interpret complex biometric outputs.

Contribution 3: Introducing preliminary evidence for a metric anxiety hierarchy as a theoretical construct. The hierarchy — Stress Score > Sleep Score > HRV > SpO2 > Heart Rate — maps directly onto the gradient of clinical complexity and interpretive ambiguity of each metric.

Contribution 4: Self-tracking commitment as a potential antecedent construct. The strong inter-scale correlations (C2ST vs SHAI-5: $r_s = 0.705$; C2ST vs CSS-12: $r_s = 0.554$) suggest a sequential pathway from behavioural commitment to cognitive anxiety to cyberchondric escalation.

Contribution 5: Informing the Theory of Psychologically Safe Health Dashboard Design. This study adds an empirical cautionary note: persuasive health technology designed to maximise engagement with biometric data may produce net psychological harm for certain populations.

A. Limitations

First, the sample size of N = 46 limits generalisability and constrains the regression analysis. Second, the SHAI composite is based on 2 original Likert items supplemented by 3 categorically converted items — a study-specific proxy rather than the validated full-scale instrument. Third, the CSS-12 and C2ST scales were each administered at half their validated length. Fourth, the cross-sectional design precludes causal inference. Fifth, the brand comparison analysis was severely underpowered with only 19 Apple Watch users.

XIV. CONCLUSION

This study set out to examine which smartwatch health metrics are most strongly associated with health anxiety among middle-aged urban professionals aged 40–55 and to propose an evidence-based metric anxiety hierarchy to guide psychologically informed dashboard design.

The primary finding is that four of five smartwatch metrics — Stress Score, HRV, SpO2, and Sleep Score — show statistically significant positive associations with health anxiety and cyberchondria outcomes after Bonferroni correction. The proposed metric anxiety hierarchy ranks Stress Score first (avg $|r_s| = 0.509$), with HRV and SpO2 jointly second (both avg $|r_s| = 0.473$), followed by Sleep Score (0.460). Heart Rate, despite being the most frequently monitored metric, did not reach significance against either anxiety outcome.

A critical qualification must be carried through: the anxiety and cyberchondria scores in this sample are low-to-moderate. No metric is producing clinically alarming distress in this population at present. The concern this study raises is a measurable directional association embedded in the structural design of health dashboards. As smartwatch adoption deepens and monitoring frequency increases, the conditions this study has identified may intensify rather than resolve on their own.

The secondary regression analysis confirmed that the five metrics collectively explain a meaningful proportion of variance in anxiety outcomes ($R^2 = 0.564$ for SHAI-5, 0.443 for CSS-12). Stress Score, the strongest bivariate associate of anxiety, loses individual significance in the regression model due to its high overlap with Sleep Score. This divergence is a substantive finding demonstrating that the anxiety-related associations of individual metrics cannot be understood in isolation.

From a design perspective, the study argues for a reorientation in how health dashboards are conceived — from maximising data visibility to optimising interpretive clarity. Heart Rate achieves the design standard the other metrics do not: it gives users enough context to interpret a reading without distress. Applying this same interpretive clarity to Stress Score, Sleep Score, HRV, and SpO2 is the central design challenge this study identifies.

Future research should address the limitations directly: a larger sample for generalisability, a longitudinal design to resolve the directionality question, full validated versions of the SHAI, CSS-12, and C2ST scales, experimental studies testing whether contextual design interventions reduce monitoring-induced anxiety, and qualitative follow-up to illuminate the psychological mechanisms.

REFERENCES

[1] Triantafyllidis, A. K., & Velardo, C. (2024). Smartwatches in healthcare. Journal of Medical Internet Research.

[2] Köhler, A., et al. (2024). Wearable adoption among adults aged 40–55. Digital Health, 10.

[3] Henriksen, A., et al. (2018). Using fitness trackers and smartwatches to measure physical activity in research: Analysis of consumer wrist-worn wearables. Journal of Medical Internet Research, 20(3), e110.

[4] Vemuri, S., & Prusty, S. (2025). Biometric data misinterpretation in wearable health devices. Health Informatics Journal.

[5] Masoumian Hosseini, S. T., et al. (2023). Smartwatch health metrics: A systematic survey. Sensors, 23(5).

[6] Kerner, D. A., & Goodyear, V. A. (2017). The motivational impact of wearable healthy lifestyle technologies. American Journal of Health Education, 48(5), 287–295.

[7] Perry, R., & Steel, Z. (2025). Somatic surveillance and health-related internet use. Journal of Anxiety Disorders.

[8] Jo, E., et al. (2024). Association between heart rate variability metrics from a smartwatch and self-reported depression and anxiety symptoms: A four-week longitudinal study. Frontiers in Psychiatry, 15. https://doi.org/10.3389/fpsyt.2024.1371946

[9] Hickey, B. A., et al. (2021). Smart devices and wearable technologies to detect and monitor mental health conditions and stress: A systematic review. Sensors, 21(10), 3461. https://doi.org/10.3390/s21103461

[10] Baron, K. G., et al. (2017). Orthosomnia: Are some patients taking the quantified self too far? Journal of Clinical Sleep Medicine, 13(2), 351–354.

[11] Jahrami, H., et al. (2024). Prevalence of orthosomnia in a general population sample: A cross-sectional study. Brain Sciences, 14(11), 1123. https://doi.org/10.3390/brainsci14111123

[12] Schenkel, K., et al. (2021). Conceptualizations of cyberchondria and relations to the anxiety

spectrum: Systematic review and meta-analysis. Journal of Medical Internet Research, 23(3). https://doi.org/10.2196/24876
[13] Odeh, V. A., et al. (2024). Recent advances in the wearable devices for monitoring and management of heart failure. Reviews in Cardiovascular Medicine, 25(10). https://doi.org/10.31083/j.rcm2510386
[14] Field, A. (2018). Discovering statistics using IBM SPSS statistics (5th ed.). SAGE Publications.
[15] Green, S. B. (1991). How many subjects does it take to do a regression analysis? Multivariate Behavioral Research, 26(3), 499–510.
[16] Nunnally, J. C. (1978). Psychometric theory (2nd ed.). McGraw-Hill.
[17] George, D., & Mallery, P. (2003). SPSS for Windows step by step: A simple guide and reference (4th ed.). Allyn & Bacon.
[18] Salkovskis, P. M., et al. (2002). The Health Anxiety Inventory: Development and validation. Psychological Medicine, 32(5), 843–853.
[19] McElroy, E., et al. (2019). The CSS-12: Development and validation of a short-form version of the cyberchondria severity scale. Cyberpsychology, Behavior, and Social Networking, 22(5), 330–335.
[20] Sullivan, G. M., & Artino, A. R. (2013). Analyzing and interpreting data from Likert-type scales. Journal of Graduate Medical Education, 5(4), 541–542.
[21] Swan, M. (2013). The quantified self: Fundamental disruption in big data science and biological discovery. Big Data, 1(2), 85–99.
[22] Lupton, D. (2016). The quantified self. Polity Press.
[23] Starcevic, V., & Berle, D. (2013). Cyberchondria: Towards a better understanding of excessive health-related internet use. Expert Review of Neurotherapeutics, 13(2), 205–213.

CRediT AUTHORSHIP CONTRIBUTION STATEMENT

Ayushi Saxena: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Visualization, Writing – original draft.

Pranjal Sinha: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Visualization, Writing – original draft.

Prashanth Cherian Kochuveetil: Methodology, Supervision, Writing – review & editing.